Routledge Studies in Contemporary Philosophy

PHENOMENOLOGIES OF THE DIGITAL AGE

THE VIRTUAL, THE FICTIONAL, THE MAGICAL

Edited by Marco Cavallaro and Nicolas de Warren

"This collection presents a variety of significant contributions that employ the phenomenological approach to investigate the ways in which human experience is shaped and modified by new technologies in the digital age. It is essential reading for anyone interested in the theoretical, existential, and social implications of digital technologies."

Andrea Pace Giannotta, *University of Florence, Italy*

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Phenomenologies of the Digital Age

This volume explores the broad and rich spectrum of contemporary phenomenological engagement with digital technologies. By focusing on plural forms of the digital, it offers a robust and flexible framework for contemporary phenomenological investigations in the digital age.

It contends that the impact of digital technologies on the lifeworld involves *both* the emergence of novel fields of lived experience in need of phenomenological analysis *and* the transformation of the method and attitude of phenomenologically oriented philosophers toward the world. The chapters cover topics, including immersion in virtual environments, the impact of digital cognitive devices on our perception of time, the invisibility of digital technologies in the lifeworld, the new extension of reality rendered possible by the employment of digital devices, how new technologies affect our intimacy and sexual body, the new methodological paradigm for phenomenological research prompted by digital technologies, the additive upshot of virtual imaginary, the intersection of the real and the virtual in augmented reality experiences, the structures of perception in the regime of digitally generated environments, how it feels like to empathize with others in a regime of virtual reality, process of en-rolling in the constitution of a virtual subject, the transformation of virtual reality into conspiratorial reality by means of on-line media platforms, and the problem of the extent to which technological environments impact human cognitive and perceptual experience.

Phenomenologies of the Digital Age will be of interest to scholars and advanced students working in phenomenology, philosophy of technology, science & technology studies, and media studies.

Marco Cavallaro received his Ph.D. from the University of Cologne and is currently Research Assistant at the Husserl Archive of the University of Cologne. Specializing in phenomenology, philosophy of mind, and digitality, he authored *Wissenchaft als System: Husserls Begriff der Wissenschaftslehre* (2025) and co-edited *Perspectives on the Philosophy of Culture: Husserl and Cassirer* (2022), *Phenomenology of Phantasy and Emotions* (2022), and *The Existential Husserl: A Collection of Critical Essays* (2022) His articles are published in journals such as the *Journal of the British Society for Phenomenology*, *Husserl Studies*, and *Phänomenologische Forschungen*.

Nicolas de Warren is Professor of Philosophy and Jewish Studies at Penn State. He is the author of *Husserl and the Promise of Time* (2010), *A Momentary Breathlessness in the Sadness of Time* (2018), *Original Forgiveness* (2020), and *German Philosophy and the First World War* (2023).

Routledge Studies in Contemporary Philosophy

The Rhetoricity of Philosophy
Audience in Perelman and Ricoeur after the Badiou-Cassin Debate
Blake D. Scott

Imagination and Experience
Philosophical Explorations
Edited by Íngrid Vendrell Ferran and Christiana Werner

Philosophy of Plant Cognition
Interdisciplinary Perspectives
Edited by Gabriele Ferretti, Peter Schulte, and Markus Wild

The Autonomous Individual
A Praxeological Enactivist Account
Martin Weichold

The Philosophy of Transformative Experience
Edited by Michael Campbell

New Perspectives on Transparency and Self-Knowledge
Edited by Adam Andreotta and Benjamin Winokur

Philosophy Across Borders
Perspectives from Contemporary Theory
Edited by Emma Ingala and Gavin Rae

Phenomenologies of the Digital Age
The Virtual, the Fictional, the Magical
Edited by Marco Cavallaro and Nicolas de Warren

For more information about this series, please visit: www.routledge.com/Routledge-Studies-in-Contemporary-Philosophy/book-series/SE0720

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Phenomenologies of the Digital Age

The Virtual, the Fictional, the Magical

Edited by Marco Cavallaro and Nicolas de Warren

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

First published 2025
by Routledge
605 Third Avenue, New York, NY 10158

and by Routledge
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

ISBN: 978-1-032-31992-6 (hbk)
ISBN: 978-1-032-31993-3 (pbk)
ISBN: 978-1-003-31228-4 (ebk)

DOI: 10.4324/9781003312284

Typeset in Sabon
by KnowledgeWorks Global Ltd.

Contents

Contributors

Thomas Arnold, Dr. Universität Heidelberg, Philosophisches Seminar

Thomas Bedorf, Prof. FernUniversität in Hagen, Institut für Philosophie

Federica Buongiorno, Prof. Università degli Studi Firenze, Dipartimento di Lettere e Filosofia

Marco Cavallaro, Dr. Universität zu Köln, Philosophisches Seminar, Husserl-Archiv

Nicolas de Warren, Prof. PennState, College of the Liberal Arts, Department of Philosophy

Christoph Durt, Dr. Technische Universität München, Deparment of Science, Technology and Society

Thomas Fuchs, Prof. Universität Heidelberg, Philosophische Fakultät, Philosophisches Semianr, Karl-Jaspers-Professor für Philosophische Grundlagen der Psychiatrie und Psychotherapie

Saulius Geniusas, Prof. The Chinese University of Hong Kong, Department of Philosophy

Daniel O'Shiel, Dr. Brunel University London

Witold Płotka, Prof. Cardinal Stefan Wyszynski University in Warsaw, Institute of Philosophy

Íngrid Vendrell Ferran, Prof. Philipps-Universität Marburg, Institut für Philosophie

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Introduction

Marco Cavallaro and Nicolas de Warren

The COVID-19 years of social distancing and increased migration of our daily lives, professions, education, and entertainment onto digital media and on-line platforms has rendered evident the stark degree to which we exist in a digital age of historically unprecedented promise and peril. Every aspect of human experience – cultural, psychological, political, and social – has been affected by the proliferation of digital technologies since the end of the 20th century, and this accelerated "digitization" of the lifeworld follows in the wake of the mediatization of experience (television, etc.) since the 1950s and 1960s. For most of us, digital technologies have impacted our ecological niche, reaching a point where we are unable to imagine a world devoid of any reliance on technological and digital implements. At the same time, the advent of new technological devices have altered human possibilities in ways that challenge our traditional anthropological self-understanding.

Although media and digital technologies have been a concern of theoretical reflection and conceptual analysis since the pioneering writings of Marshall McLuhan, Günther Anders, Mark Poster, and others, a phenomenological approach to the manifold lived experience of digital interactions and interfaces still remains inadequately developed, although arguably the challenge posed to philosophical reflection and phenomenological understanding of digital-mediated lived experience seems more pronounced than before. It remains an open question whether contemporary technological advances (and the sociological and existential upheavals that ensue) threaten to destabilize the "traditional" method of phenomenological inquiry or, on the contrary, present an opportunity for its revitalization and reimagining.

Two contrasting perceptions have commonly shaped the debate concerning how digital technologies might impact phenomenological thinking. On one view, the intersection between lived experience and digital technologies promises to change the method and approach of phenomenological analysis. On another view, phenomenological analysis is argued

DOI: 10.4324/9781003312284-1

to rely on established methodological procedures (reflection, epoché, reduction, eidetic variation, etc.) that are dependent on the medium and situation in which the phenomenologist finds themselves. The transition from traditional to digital technologies would thus result in a fundamental change in how phenomenologists develop phenomenological descriptions as well as what objects are deemed to require phenomenological inquiry. Both views, however, are not antithetical to each other for it is the basic contention of this volume that the impact of digital technologies on the lifeworld involves *both* the emergence of novel fields of lived experience in need of phenomenological analysis *and* the transformation of the method and attitude of phenomenologically oriented philosophers towards the world.

Phenomenologies of the Digital Age: the Virtual, the Fictional, and the Magical aims to explore a spectrum of phenomenological engagement with digital technologies. Rather than solely capture the basic form of appearance, or manifestation, and hence experience, of digital technologies, the title of this collective volume keeps open, and hence, alive, the conceptual matrix in which phenomenological investigations into the experience of digital technology should operate, thus allowing for an investigation of *plural* senses of lived experience, and hence, phenomenologies, of the digital age. Since the late 1980s, the terms "virtual reality" and "virtuality" are widely understood as referencing a host of audiovisual and tactile simulation technologies. The upshot of these technologies is, on the one hand, the creation of new forms of experience (e.g., immersion) with their own constitutive structures – such as the "compression" of time and space. On the other hand, digital technologies encroach on our "natural experience" of the world through what has been interpreted as a media-induced dissolution of our understanding of reality and a virtualization and/or derealization of the real (Flusser, Baudrillard). The fictional can be seen as a broader category that in some ways foreshadowed the constitution of experience through digital technologies. The founding generation of phenomenology devoted multi-faced analyses of fictional experiences (Husserl, Ingarden, Fink, Sartre), most often with reference to technological artefacts such as photographs, cinematic images, and literary fictions. It is in this context, before the advent of late 20th-century media technologies that the terms "virtual" and "fictional" enjoyed limited circulation among phenomenological thinkers. Of special note here is the groundbreaking analysis of television, radio, and "the matrix" in the writings of Anders (who, in fact, coined the term "the matrix"). Within phenomenological discussions of phantasy and the imagination, the category of "the magical" was developed in unprecedented ways, above all in the studies of the imaginary and the emotions by Sartre, to refer to the "quasi-presence" of the imaginary objects and their performative instantiation. It is our

contention that this rich array of plural senses of the virtual, the fictional, and the magical offers a robust and flexible framework for contemporary phenomenological investigations in the digital age. More specifically, contemporary media transform the sense of and relationship between the "visible" and "the invisible," "the apparent," and "the inapparent," in ways that should provoke and revitalize a phenomenological approach to the myriad constitution of how things appear, or do not appear, in the world.

The contributions in this volume address these questions more or less directly, focusing on a wide range of issues related to the phenomenology of digital technologies.

In "Layers of Simplicity: Phenomenological Considerations of Immersion in Videogames," *Thomas Arnold* discusses how immersion in videogames is both multi-layered and simplified compared to real life. He examines how players navigate and interact within virtual environments, transforming space into meaningful places. He also discusses the narrative and moral dimensions, particularly in violent videogames, where players' choices contribute to their virtual moral selves. In addition, Arnold highlights the reduction of complexity in videogames, which increases engagement by providing clear purposes and goals. This simplicity promotes deeper immersion by allowing players to experience a unified, atmospheric world with straightforward interactions. Arnold's analysis shows that the streamlined nature of videogames offers a more engaging and meaningful experience by providing a clear sense of purpose, making them more captivating than the often fragmented and complex real world.

In "The Otherness of the Other Interface: Relationality and Corporality between Fellow Human Beings and Fellow Machines (*Mitmaschinen*)," *Thomas Bedorf* explores the complex relationship between humans and machines, critiquing both cultural pessimism and transhumanist celebration of digital detachment. He proposes a shift from asking "what" questions about human and machine identities to "how" questions about their interactions. Bedorf critiques the reductionist view that equates consciousness with information, asserting the necessity of bodily experience for true understanding and interaction. He advocates for a phenomenological approach, emphasizing the importance of understanding human-machine interactions as an extension of our bodily existence and relational experiences, rather than attempting to define machines by their capabilities alone.

In "Time-Consciousness and E-Memory: Arguing for a Phenomenological Revision of the 'HEC' (Hypothesis of Extended Cognition) Paradigm," *Federica Buongiorno* investigates the impact of electronic memory (E-memory) on biological memory within the framework of the Hypothesis of Extended Cognition (HEC). Buongiorno critiques the Parity Principle (PP) proposed by Clark & Chalmers, arguing that a phenomenological perspective can better address the integration of E-memory and

biological memory. She differentiates between primary (automatic) and secondary (reflective) memory processes, suggesting that while E-memory can be incorporated similarly to biological memory at a procedural level, it requires critical scrutiny at an explicit level. Buongiorno emphasizes that E-memory is part of a hybrid cognitive system where sensory-motor and cognitive performances are augmented by digital media. She concludes that incorporating E-memory into extended cognition requires acknowledging the role of reflection in secondary memory, thus revising the PP to better fit phenomenological insights.

In "The Vanishing Point: Digital Technologies and the Quest for a Phenomenology of Technological Invisibility," *Marco Cavallaro* examines the increasing invisibility of digital technologies and their impact on human experience. He distinguishes between different types of technological invisibility, including perceptual invisibility (when technology's functioning is hidden), use invisibility (when technology becomes transparent during use), purpose invisibility (when its purpose or function cannot be understood on the basis of its perceptible makeup), and environmental invisibility (when technology is so embedded in our environment that it becomes part of the natural background). Cavallaro argues that a phenomenological approach, which traditionally focuses on what is visible, can be adapted to address the invisibility of digital technology by examining how these technologies shape and mediate our experience of the world. This approach helps in understanding the profound and often unnoticed impact of digital technology on modern life.

In "The Phantom Matrix. A Critical Phenomenology of Television," *Nicolas de Warren* develops an analysis of the various ways in which television has transformed "being in the world" through a discussion of Günther Anders' seminal analysis of television in *Die Antiquiertheit des Menschen* along with reflections on portraits of television in Jean-Phillipe Toussaint's *Television* and Ray Bradbury's *Fahrenheit 451*.

In "Large Language Models and the Patterns of Human Language Use," *Christoph Durt* and *Thomas Fuchs* explore how text-generating Large Language Models (LLMs), such as ChatGPT, create human-like text. They argue against the notion that LLMs possess consciousness or understanding, instead asserting that these models reveal the patterned nature of human language. LLMs are trained on vast amounts of text, learning to predict language patterns that humans find meaningful. These models exploit statistical regularities in language use, which are tied to human experiential patterns. The authors emphasize that LLMs produce coherent and contextually appropriate text by recombining these patterns, not by understanding meaning. They also discuss the propensity of LLMs to generate clichés, biases, and toxic language due to their reliance on common patterns. The chapter underscores the importance of distinguishing

between the mechanical reproduction of patterns by LLMs and the authentic, contextually grounded language use by humans.

In "On Tertiary Retentions and Digital Sedimentations: Bernard Stiegler and Phenomenology," *Saulius Geniusas* examines Bernard Stiegler's adaptation of Edmund Husserl's phenomenological concepts, particularly focusing on tertiary retentions. Stiegler, influenced by Husserl's concept of retention, extends it by introducing tertiary retentions, which refer to the externalization of memory in material forms like writing and digital media. Stiegler argues that these externalized memories significantly shape individual and collective processes of individuation. Digital technologies, which he views as the latest form of tertiary retentions, disrupt traditional modes of memory and individuation by imposing algorithmically generated content on consciousness. This leads to a homogenization of experiences and identities, reducing individuals to consumers. Stiegler warns that this digital sedimentation threatens the autonomy and depth of human thought, advocating for a political and educational response to reclaim deep attention and foster genuine individuation amidst the pervasive influence of digital culture.

In "The Imaginary Magic and Hypervirtuality: On the Phenomenological Nature of Digital Screens," *Daniel O'Shiel* explores the intricate relationship between humans and digital screens through a phenomenological lens. He builds on the works of Husserl, Fink, and Sartre to discuss perception, phantasy, and image-consciousness. O'Shiel identifies two main types of virtuality: perceptual and digital. He emphasizes Sartre's concept of the imaginary, highlighting its inherent magic and addictive potential, which he terms "hypervirtuality." Digital screens, O'Shiel argues, open digital universes while distancing users from immediate perceptual realities. This digital engagement can lead to a form of addiction, as the compelling nature of digital experiences captivates and often overwhelms users, resulting in a significant shift in how reality and virtuality are experienced and valued. The chapter underscores the need to understand these dynamics to grasp the profound impact of digital screens on our everyday lives.

In "Perceiving the Virtual: Rethinking Blaustein within the Phenomenology of Virtual Reality," *Witold Płotka* reinterprets Leopold Blaustein's phenomenology of media to explore contemporary virtual reality (VR) experiences. Płotka argues that Blaustein's insights into mediated aesthetic experiences, such as cinema and radio, can inform our understanding of VR. Blaustein's concept of imaginative perception, where objects are experienced as quasi-real, are particularly relevant. Płotka discusses the parallels between traditional media and VR, noting that both involve a shift in user attitude to perceive imaginative objects. The chapter also engages with current debates on virtual objects, contrasting Blaustein's views with David Chalmers' virtual realism. Płotka concludes that Blaustein's

framework, which sees virtual objects as quasi-real and constituted by the user's imaginative engagement, provides a nuanced perspective on VR that bridges historical phenomenology and modern virtual experiences.

In "From Immersive Body Swap to the Apprehension of the Other's Emotions: Levels of Empathy in Embodied Virtual Reality," Íngrid Vendrell Ferran explores how embodied virtual reality (EVR) can train empathy by creating perceptual illusions that enable users to adopt another person's perspective. The chapter distinguishes between two types of empathy: sensual and emotional. Sensual empathy involves identifying with another's lived body through immersive experiences, such as body ownership illusions created by EVR. However, emotional empathy, which involves understanding another's emotional states and values, requires more than just perceptual illusions; it needs additional mechanisms to access the other's world of values. Vendrell Ferran argues that while EVR can effectively foster sensual empathy through perspective-shifting and the illusion of presence, it must be supplemented with other tools, like narrative elements, to promote emotional empathy fully. This distinction highlights the potential and limitations of EVR in enhancing different levels of empathy.

In conclusion, the profound influence of digital technologies on our everyday lives necessitates a nuanced phenomenological examination, as evidenced by the contributions in this volume. The COVID-19 pandemic accelerated the integration of digital technologies into our daily routines, underscoring the need for a deeper understanding of their impact on human experience. This collection of essays highlights the dual nature of digital technology: its promise of new forms of engagement and immersion, and its potential to disrupt traditional modes of perception and interaction. Together, these essays underscore the transformative impact of digital technologies on the lifeworld, calling for a renewed phenomenological approach that addresses the complexities of digital-mediated experiences. This volume not only maps out the diverse phenomenological landscapes shaped by digital technologies but also provides a framework for future inquiry into the intersections of technology, perception, and human experience.

1 Layers of Simplicity

Phenomenological Considerations of Immersion in Video Games

Thomas Arnold

In my contribution, I analyze immersion in our experience of video games.[1] My hypothesis is that immersion in many video games is (a) multi-layered and involves almost every structure of the constitution of the real world, from bodily movement to practical navigation, on to moral agency; and that it is (b) at the same time based on the massive reduction of complexity in both the virtual worlds and the virtual selves.

To substantiate this claim, I will – after a brief introduction – explore the phenomenon of virtual experience (or immersion) in layers, drawing on several strands of phenomenological philosophy. The first part of the chapter (Spaces and Places) concerns the question how virtual experience is possible. I will investigate the sense of kinaesthetic and fictional immersion: how we seem to step inside and move within another, a virtual world; how we navigate from virtual place to virtual place, how we explore this world, and how we use virtual equipment – thus forming an apperception of a virtually embodied, practical self.

The second part of the chapter (Meanings and Morals) will explore how this virtual self/world is further constituted through a narrative and how we come to constitute and then be drawn into the virtual events by means of the moral dimension of virtual action, especially violence. Violent video games engage us practically by forcing us to make morally relevant choices. This choosing of virtual actions implies (at least virtual) responsibility, and thus, our choices constitute (and flow from) the moral personality of our (virtual?) self.

The third part (Depth, Death, and Reduction) deals with the reduction of complexity in video games in comparison with the real world. Virtual equipment, virtual agency, virtual atmospheres, virtual selves, and indeed virtual death are all relatively non-complex if compared with the intricacies of real life. Less clutter and fewer options allow us to lose ourselves in virtual worlds and to experience, in a way, more world, more self – and more meaning.

DOI: 10.4324/9781003312284-2

1.1 Introduction

Phenomenological investigations need a guiding thread. In our case, I have chosen virtual violence as a recurring topic, especially as witnessed in virtual combat. There are several reasons for this choice. In the vast majority of cases, whenever video games are discussed in public, the topic is violence anyway (Schott 2016). The debates are usually couched in Platonic-Aristotelian terms (although most participants do not realize this) and concern the question whether witnessing representations of violence makes us prone to act out real violence (Plato) or whether representations of violence might even be cathartic (Aristotle). I am, however, more interested in how virtual violence appears, what its role in our immersive experiences is, and what its appearance implies for the subject as a player. After all, the active as well as passive experience of violence implies a host of underlying structures both on the subject and the object sides of video games. Virtual violence engages us completely; virtual combat can offer one of the most intense (frustrating, elevating, flow-like) experiences of immersion, and from this total engagement, we can glean how immersion in video games might work and how their constitution is structured. Therefore, the experience of virtual violence will guide us in parts of the chapter. (Which is not to say that the same layers cannot be found in non-violent games too.)

This also means that some very successful video games will not appear in the range of concern. *Tetris* or *Candy Crush*, for example, are very immersive and they necessarily share some basic layers of constitution with more complex games, but neither do they feature a gripping narrative with meaningful characters, nor an atmospheric world nor much, if any violence. Concrete examples of games I have in mind are RPG/Action/action-adventure games from the *Elder Scrolls*, *The Witcher* universe, or the *GTA* series. Combat games like *Tekken* or *Mortal Kombat*, as well as shooters like *Fortnite*, *Halo*, or *Battlefield* might also work, to name just a few paradigmatic cases. In any case, I hope that virtual violence allows us to understand the great range of possibilities how video games are constituted phenomenologically and, correlatively, how we immerse ourselves in them.

Immersion into any medium has been a topic of several different approaches for a while now and its definition remains difficult, mostly because the range of phenomena referred to by the term is so vast. "Immersion" can refer to the receding awareness of our physical body, to a suspension of disbelief, to practical problem-solving engagement, to being engrossed, invested in a narrative or other types of experience. In a very basic sense, it implies directing in- and attentional acts toward whatever we are immersed in, excluding everything else.

The origin of the metaphor additionally seems to imply a change of medium too; thus, while we can be immersed in music, books, films, games, we cannot be immersed in the act of making coffee, even if we are utterly absorbed by it, mindful of it, or "in the zone" while doing it. Immersion takes us from the real world into other worlds. Phenomenologically speaking, this means that immersion is a form of constitution: in immersion, the game world gains presence. And it gains a presence which excludes focus and attention on objects outside the game world; nothing outside the medium appears. This also means that immersion excludes reflection, since reflection brings the constituting subject into focus, which is not part of the game world and not an element of the medium. Some games (books, movies) still test the balance between reflection and immersion, as they contain metafictional elements, leading to reflections on gaming within the game itself.[2]

Immersion is not the only form of constitution for virtual worlds. Virtual worlds can be present without us being fully immersed in them, that is, engaged, engrossed, etc. Whether the constitution of a game's world is fully immersive or not depends on empirical factors, like the quality of the game or the psychological mood or general personality of the gamer. For example, if we play a game like *Fortnite* with friends and talk to them on the team chat, their presence is real (if mediated) and we might be very absorbed by the game as a concentrated team effort, but still the virtual world is not exclusively present, we are not fully immersed in the game environment, because a part of the real world, namely our friends, is also present; after all, playing together is a real event as well as a virtual one. Or take the experience of an immersion-breaking bug; here, the immersion will stop, but the virtual world will not completely disappear either. To sum up, immersion is a specific type of encountering virtual worlds, and arguably the one most gamers and designers look for – certainly the one in which virtual worlds are most present, so to speak.

It is also easy to see that different media as well as different products (or events) afford different kinds of immersion (Adams 2004), depending on their design and the patterns they present (Bjork and Holopainen 2005). Chess engages us on a cerebral level, fast-paced shooter games on an almost reflexive fight-or-flight level. In puzzle games, the mathematical sublime might "overwhelm" us (Bogost 2015, 106, 110), which arguably also constitutes a form of immersion. Some forms of engagement might even be unpleasant, where immersion and unhealthy obsession or "devotion of material indifference" (Bogost 2015, 5) blur, born of misery, masochism, and repetition.

Without further differentiating kinds of immersion or matching them with different genres, I aim to show that immersion in most video games is (a) multi-layered constitution, and (b) based on a reduction of

complexity. As Ian Bogost puts it so aptly, most video games are like "Wagnerian Gesamtkunstwerk-flavored chewing gum" (Bogost 2015, ix) and playing them engages us accordingly – as complex (embodied, engaged, moral) subjects, but not in the same way a complex 7-course wine-tasting menu would. In the first and second part, I will discuss the layers of constitution and engagement, whereas in the third part, I will highlight how all these strata afford for starker experiences due to the overall reduction of complexity in comparison with experiences of or in the real world.

1.2 Spaces and Places

The cases of virtual violence I am most interested in occur in relatively rich virtual worlds and concern virtual entities. We encounter violence in tactical first-person sci-fi shooters, medieval real-time strategy games, fantasy epics, button-mashing fighting games, even in certain car(t)-racing games. While not all violence is equal, all cases of virtual violence share at least this: they are all virtual and take place in virtual worlds. But what exactly is a virtual world? Or rather, phenomenologically speaking, how do virtual worlds appear (cf. Arnold 2015)?

So far, all paradigm cases of what we call a "virtual world" appear on some form of screen – be they fixed on a table or mobile or even part of a VR headset – and through some form of loudspeaker, based on the operations of some form of computing device.[3] The input is achieved through controllers, mouse and keyboard, or camera combined with movement recognition. (We will ignore more outlandish forms of output and input, like retinal projection or touch gloves, because they are not prevalent yet.) All these devices ultimately enable what we might call the "artificial presence" (cf. Wiesing 2010) of virtual worlds.

Since the visual modality is arguably the most important, we will focus on it (they are, after all, called "video games" not "audiogames," "haptigames," or, horribile dictu, "olfactory games"). The screen we use as a display is what Husserl calls the "image-carrier (*Bildträger*)."[4] It is an object in real space, with real spatial dimensions, real weight, etc. As is the case with pictures and films, the carrier enables the appearance of non-real objects, but disappears in the experience of immersion – once we are engrossed in the game, we become oblivious to the equipment used. A malfunction might bring it back into focus, but mostly it becomes part of the enabling background.

On the screen appear what Husserl calls "image-objects (*Bildobjekte*)," by which we mean configurations of light we interpret as figures, say, a little dragon, a little horse, and a little spaceship. They occupy a space in-between real space and proper pictorial or virtual space as they can still

be measured in real space dimensions: the image-object dragon might be a few centimeters big.

Through these image objects, we are however directed at the "image-subject (*Bild-Sujet*)," in this case, the dragon itself. This represented dragon does not inhabit real space and might very well be several dozen meters long. It is part of the pictorial or virtual space, which is partially represented by the images on the screen. Partially, because the virtual space is usually bigger than what is represented in it. The orientation of the different spaces might overlap or differ: if the display is rotated (for example, if I am playing lying on my side), the orientation is rotated too, making "up" in the pictorial space different from "up" in the real space room. Real space and pictorial space are held apart by their content (a bit of active display technology on the one hand and a huge dragon on the other) as well as by the inability of spatial crossings: I cannot move what is in one space to the other. Hence, phenomenologically speaking, virtual space is not at all "half real" (Juul 2005),[5] but fully virtual, yet grounded in real space.

So far, virtual space and traditional pictorial space do not seem to differ much, and in terms of pictoriality, they do not. However, as film adds perceived (or merely seen) movement to the pictorial space, video games add enacted movement. Virtual space is pictorial space which, in and through which we can move.[6] Our active movement is made possible by changing the images through manipulation of controlling devices. Unlike switching channels via remote, virtual movement is a form of interaction with virtual space itself. We do not just watch video games, we play them.

Depending on the form of input, the real space movement necessary to affect virtual movement can be fairly small, for example, the flick of a thumb. Yet, it is movement enough to analyze it using Husserl's theory of kinaesthesis (cf. Husserl 1995, Sections IV–V). Real space and spatial relations appear to us in correlation to our own felt movements: I can experience the dimensions of my laptop by moving around it; if I do not feel myself move and the appearance and orientation of the laptop still change, I understand the laptop to be moving; I understand something to be far away if my own movements do not change its appearance (much), etc.

In the case of virtual movement, these connections still hold, although in a mediated way. If my input device is a controller, a change of appearances similar to what would occur in real space if I moved my whole body might now be connected to a slight tilt of my left thumb. So, if I move my thumb in the appropriate way but nothing changes, I experience an inhibition of virtual movement (or an equipment failure). If I see a change without the associated real bodily movement of my fingers, I experience being moved virtually. The shift from real space kinaesthetic motivation and virtual kinaesthetic motivation takes a while to become habituated, as can be seen in the case of people unused to

virtual movement, who move according to real kinaesthetic motivation, for example, by bodily leaning left. This mediation holds for actions as well and will become pertinent once we reach our discussion of violence because brutally murdering someone virtually with a two-handed transverse sword-cut might be kinaesthetically connected only to lightly tapping a button.

Speaking of actions: an empty Euclidean grid, or an empty two-dimensional desktop-surface through which we can move (either by using a cursor or by changing our virtual point of view) are not able to host actions apart from mere movement. Even "navigation" is too meaningful a word for the kind of abstract movement, thus virtual space is not yet "cyberspace" (from Greek "*kybernaô*," to navigate). Virtual space in itself is not enough to count as an environment for proper engagement, let alone violence: there is no one to perpetrate it, no one to become its victim, no means to enact it, no tools to use, nowhere for it to take place, no possible purpose behind it. In other words, a virtual space is not yet a virtual world.

What is lacking is meaning. For movement to become navigation or pathfinding, space needs to become a place. In a Heideggerian vein, we might say that places are always places for something or of something. Put a little paradoxically, a place is somewhere something can take place. In video games, this referential character of the environment tends to be very pronounced. Virtual woods are dangerous as places where an ambush might happen, where monsters dwell, or where we can find ingredients. Mountain ranges too are places of danger and beauty (no vista can be less than breath-taking). Although what Augé calls non-places (of non-descript, meaningless space) exist in video games too, they also serve a function as points of ingress, as liminal spaces (Augé 1992).

As is obvious from this description, turning mere space into a structured manifold of interesting places implies a whole structure with different elements. A virtual form of intersubjectivity comes into play as monsters, enemies, and friends appear. The existence of ingredients (or materials) implies the existence of what Heidegger calls "equipment" (Heidegger 2006, §§ 15–17), which in turn comes with a host of its own structural implications. This complexity is not surprising, given that meaning in general does not come on its own; we cannot encounter one meaningful thing, only webs of meaning in which we become involved. Indeed, the totality of involvement is what Heidegger calls "the world."

And so, the geometry of virtual space gains the topology of virtual places in which navigation and indeed exploration become possible. Only once these structures of meaning and space are available can we find ergodic (as well as hodologic) elements, as pure virtual space has no "text," that is, no meaningful topology to work through and there are no paths to follow.

Since virtual agents are victims, perpetrators, and witnesses to virtual violence, we ought to dwell on them for a bit. The concept of a virtual agent encompasses at least two things: firstly, non-player characters (NPCs) of any kind, that is, entities appearing as subjects not controlled by a human agent. Without wanting to delimit what exactly counts as an apparent subject, we include within this enemies, allies, bystanders in all shapes and forms, anything that appears to have a will and agenda of its own. Secondly, "virtual agent" might also refer to the representations of other real subjects, that is, player-controlled characters or avatars, as long as they can interact with the world. Virtual agents are a crucial element of virtual worlds if indeed the world is "with-world" (Heidegger 2006, § 26) as Heidegger puts it, that is, something populated by more than us. We will look more closely at virtual intersubjectivity once morality becomes a concern. So far, we are interested more in how virtual agents contribute to the world-hood of virtual worlds.

These virtual agents can be more or less embodied, and the level of embodiment determines the kinds of violence possible within the game; violence in the sense of using a force to harm or destroy implies a certain level of concreteness, namely, the existence of a body to be harmed or destroyed. The same goes for the use of tools in the context of violence, especially weapons: the virtual use of weapons implies certain forms of embodiment on the side of both victim and agent. Virtually "handling" a sword requires virtual hands (or equivalent appendages).

Such tools and weapons form a subset of the aforementioned equipment, which comprises virtual items determined by their "readiness-to-hand" for a given purpose. Wood can be collected in order to build a handle, which can in turn be used to craft a spear (in a smithy, using a hammer), which can then be employed for fighting, which allows us to take, say, a key, which opens a door behind which we find a magical amulet, with which we can, etc.

1.3 Meanings and Morals

The question concerning the overall meaning behind the web of practical references (the big "in-order-to") leads us to structures beyond embodied practicality. The total meaning (the Meaning which gives meaning to individual things and actions) can come about in many different ways, the most prominent of which are rules and narratives (cf. Juul 2005). In the case of sports, arena, racing, or brawler games, for example, the meaning and final purpose of actions and use of equipment springs (mostly) from the rules which determine winning and losing. Violent acts appear mostly in an abstract, instrumental sense: killing enemies is a way to win according to a set of rules. The virtual killing is simply a move in a game, more

akin to taking pieces in chess than witnessing the horrors of death in dramatic movies; the killed characters are either non-individualized NPCs or the avatars of other players which will simply reappear at the end of the round. The deaths are more or less meaningless – then again, they are still deaths, brought forth by virtual killing.

In the case of more story-driven games, meaning is generated by narrative means, both concerning the history of the virtual world and the ergodic "text" of the hero's adventure.[7] The history of the world might be presented in epic detail, it might be hinted at, or it might be hidden away in countless books readable in-game, but in fantasy or adventure games, it is always there in some way, presenting the game world as something beyond the time witnessed in playing. This reinforces the sense of what we might call "cosmic realism," insofar as a world is exactly that which exceeds or transcends our current perceptions and volitions, both in space and time.

Within this world, we act. Games offer us an "agential skeleton" (Nguyen 2019, 423), structures of agency ranging in complexity and meaning from simply determining Pacman's direction or changing the orientation of the *Tetris*-blocks to engaging with the open worlds of modern fantasy games. These virtual actions have the same features as real ones. They are ruled by norms (including goals), which means they can fail, they are conscious, they allow for teleological explanations: we do this in order to achieve that. In most games, we strive to accomplish something, and it is this striving which constitutes a main component of immersion as our own wish to win, to beat a level, to explore further, etc. draws us further into the activity. Depending on the design of the games, this activity takes the form of flow, and our sense of agency is perfectly balanced between challenge and ability, activity, and passivity:

> In ordinary, non-game life, we might catch glimpses of this kind of aesthetic experience—brief moments when our abilities and our tasks harmonize. But often they do not: our abilities fall far short of the tasks, or the tasks are horribly dull but we must put nose to grindstone and grade those papers anyway. However, we can design games for the sake of this harmony of practical fit. In games, the obstacles are designed to be solved by the human mind.
>
> (Nguyen 2019, 431)

The existence of obstacles is as important as the existence of tools or equipment as well as goals. If the attainment of a goal with the help of tools was instantaneous or without any challenge or the possibility of failure, the game would be pointless (this holds for non-virtual games as well, of course). In order to enjoy this situation, we need to adopt what Nguyen

calls a "lusory attitude: we adopt the pre-lusory goal and the constitutive rules for the sake of the activity they make possible. We adopt unnecessary obstacles to make possible the activity of overcoming them" (Nguyen 2019, 435). Action, however, is not everything.

Once a virtual world is constituted as a virtual space in which involvement can take place according to patterns of agency, structured by rules, narratives, and goals, it is not only experienced as an environment for action, but also as an atmospheric space.[8] A truly great game is characterized not only through gameplay or graphics. The experience of atmospheric spaces marks the difference between simply "solving something" and "being somewhere" (Bogost 2015, 16). So, while we actively play a game, we might also passively be enveloped by an emerging mood in which the virtual world manifests itself. The atmospheric immersion being all the greater in games for our participation rather than mere observation. We become submerged in the atmosphere because we act within the space it envelops, instead of being on-lookers.

The final structural element of the condition of possibility of the experience of virtual worlds are of course we ourselves, the subjects. The most minimal form of subjectivation is the point of view from which we perceive the virtual space. This POV might just consist in a non-focalized perspective, as is the case in many strategy games. But it might also be a focalized third- or first-person perspective. Such a perspective is an implication of the appearance of virtual space itself, as oriented visual appearances imply a center or point of reference. Our perspective might be one of pure observation and execution, without any further relation to any of the elements of the game, meaning we cannot be attacked or even registered by any in-game agents. In this case, we could at least witness virtual violence, as (other) virtual agents harm each other. A step further, we could perpetrate violence as an (almost) non-embodied, unseen entity, like in *SimCity* or *The Sims*; or like a (half-)seen and addressed entity, like in many strategy games or the *Black & White* games – which still requires the habitualized mediation from real movements to virtual actions despite the lack of an avatar.

If our point of view is recognized within the virtual world as a point of view, we can finally experience violence passively as a victim, depending on the level of embodiment. Our experience of violence will differ depending on whether there is virtual health and virtual death. Without any virtual consequences, we might still experience something that looks like violence (people attacking our point of view or our avatar with guns or blunt instruments), but it will mean less. Once the health-bar starts diminishing dangerously and we are aware that its depletion will have consequences like the failure of our quest, mission, or reset of the current level, virtual violence will have become more full-bodied, so to speak – and

accompanied by a certain amount of dread or at least the expressed wish to avoid it.

The most (and, presumably the most existentially relevant) virtual violence is experienced once we are a fully embodied virtual agent, placed in a virtual world, amidst equipment, within a narrative and facing virtual others – once we can deal harm and be harmed, that is, once things get up close and personal, so to speak.

Speaking of "personal," we now need to address the virtual other. After all, virtual violence is mostly directed at embodied virtual agents, either us or others.[9] How cruel or abhorrent virtual violence is to us, does not necessarily hinge on whether it is directed at NPCs or (our) avatars. The violence in arena-style shooters, for example, is abstract; bodies are being fired at, hit, and injured, but the violence is not particularly awful or haunting. While it is not irrelevant (shooting is important to a shooter-game), it is part of the rule-based gameplay and the fact that player-characters simply vanish and then reappear at the start of the next round makes virtual death a formality to be avoided for the sake of winning.

Some of most haunting and infamous scenes of virtual violence involve only NPCs, like the "No Russian" mission in *CoD MW 2*, the Trevor Torture scene in *GTA V* or the Eye Surgery in *Dead Space 2*, to name some examples in which we as players are perpetrators, or Elisabeth's Lobotomy from *Bioshock Infinite: Burial At Sea*, in which we inhabit the perspective of the victim (I exclude cut-scenes like the death of Sarah Miller in *TLO*, since they are more akin to cinematic violence.)[10] All these cases are striking because the violence is meaningful beyond marking wins or losses: the people involved appear to be afraid and in pain, or psychopaths enjoying themselves – overall, they appear to be persons in an eminent sense, that is, beings with feelings, goals, and character. As is the case in cinematic or literary violence, the more we apperceive victims and perpetrators as people, the closer the violence hits home (the mass-killing of stormtroopers in *Star Wars* or minions in the *Bond* movies does not usually bother us over-much – something the first *Austin Powers* movie references to comical, if cynical effect, when the family of an initially nameless dead minion is shown as they receive the bad news of their father's and husband's death at the hand of the hero.) Presumably, such striking violence is rather more likely to occur in single player games (or campaigns), which offer more opportunities to develop the necessary inter-personal apperceptions (including empathy with the virtual agent's fears and pain, and an understanding of their motives) through narrative and other ways of establishing the virtual characters as real characters. Which obviously follows the same logic of humanization or de-humanization as that in play in real-life propaganda.

Given that (a) we apperceive acts of virtual violence and (b) violence is open to moral judgments and ethical considerations in real life,

we might now reflect on whether we are engaged in virtual violence as moral subjects. This is not asking whether virtual violence is morally reprehensible or not but whether we can be involved in gaming not just bodily, atmospherically, or practically, but morally (be it as good or bad subjects).

Clearly, we can judge virtual actions (both our own and those of others, including NPCs) in moral terms – both based on our own real values as well as adopted, fictional values. But then again, we do understand that our virtual experiences (actions, sufferings) are virtual, not real, which leads to a certain form of moral disengagement.[11] As Nguyen describes it, in cases of virtual violence, we do not simply cancel the immersion as "background monitoring processes of our full agency haven't broken through with such a cancellation, we let the temporary agency regulate our decisions and dominate our consciousness" (Nguyen 2019, 443). Unless of course, the immorality becomes so jarring that immersion breaks and we really start asking moral questions regarding the representation of violence.

At this point, we can probably join our discussion of the morality of virtual violence to the discussions surrounding the morality of fictional violence in general. My main point being that in the case of games, we can become perpetrators as well as victims of such fictional violence, which makes such questions even more acute and shows us that virtual experiences concern us as (virtually) moral agents.

1.4 Depth, Death, and Reduction

So far, we have produced a layered panorama: immersion and the correlative constitution of virtual self and virtual world in video games have many levels and dimensions, both on the side of the experiencing subject and the object, even in the case of fairly simple games. But then again, how deep and complex are video games, compared to real life? Not very, but that might be the whole point: reduction as attraction.[12]

Video games are simple, put paradoxically, they reduce experiences massively in order to create a singular experience. The fact that many games keep us hooked by means of mechanisms which are akin to gambling is well known. Loops of challenge and reward can become very addictive. We have also briefly touched upon the attraction of simulated autonomy and power: finally free to act unrestrained, both liberated from the reins of reality and disengaged from our real moral codes – at least up to a certain degree. But simplicity constitutes another major factor in the attraction of video games and fosters a certain kind of immersion (immersion can, in fact, itself be described as a reduction). We will therefore now look at the simplicity, consistency, and lack of ambiguity in video games.

The first aspect concerns the worldness of virtual worlds. If we understand "world" to refer to a grand unified whole, a meaningful, atmospheric nexus, in which one thing signifies another, then our real-life experiences are rarely very world-y. Rather they are frayed, splintered, and "stuffed with stuff," so to speak. At the same time, even individual things offer a bewildering array of possibilities for thinking, feeling, or acting.

Let's take a tree as an example. It is a real tree that you see on a walk. You can enjoy its sight. Or make an effort to recognize the species (using an AR-app for plant-recognition, if it takes your fancy). Or you might consider whether you could make good furniture out of its wood – and whether you could cut it down yourself. Or you could dwell on how long the tree has been here. Or you could worry if it is threatened by some bug – or by climate change. Or you might suspect that it will soon be cut down to make room for a parking lot. Or you try to climb on it. Or you can just ignore the tree. The tree refers to many things or nothing at all, depending on the situation. With the right view, a (or the) whole world can unfold from or around this one tree. Yet, it is equally possible that the tree remains as an ambiguous piece of landscape. Or as an annoyance. Or as a riddle. In any case, many different, sometimes contradictory, ways in which the tree can be (ir-)relevant to us are conceivable, and none of them is immediately better or more important than the other. The appearance of a real tree is complex and fraught with meaning.

Now let us consider another tree. This time, it is a virtual tree that you walk past in a video game (say, for simplicity, *Skyrim*, *The Witcher 3* or *Fortnite*). It appears to you on a screen or through a screen and thanks to the complex interaction of your senses on the one hand with hardware and software on the other – just as described above. You can now admire how well the tree is rendered. You can sneak left or right. And of course, you can also ignore the virtual tree. But many of the thoughts you might have about the real tree are meaningless here: most games do not have bark beetles, climate change; and parking lots are not being built either. There is usually no climbing mechanism. So, in a certain sense, the tree is just a tree, at most an obstacle. Depending on the game, it can be a resource that is fed into virtual production processes in a relatively simple way. There will probably be no taxes, no tariffs, no strikes, no cartels, no concerns about carbon emissions.

Still, if the tree is minable in some sense, there may be tools and equipment that will allow us to fell the tree. Here, too, the cross-references are very clear and not very complex: for example, an axe is there to cut down the tree; the tree is there to be felled and processed; the wood obtained is there to build something that makes us better/stronger or contributes directly to the fulfillment of a given task. Thus, the virtual tree appears much clearer than the real tree; there are far fewer ways of grasping it. If it

cannot be processed, it primarily contributes to creating a certain atmosphere and defining the world more closely, as a barren highland or lush jungle for example.

Atmosphere itself is another type of phenomenon that is much more pronounced in game worlds than in reality – at least when they are well made, creating atmosphere is, after all, a great art (as, for example, landscape architects or stage designers know full well).[13] While our lives are indeed atmospheric in some way, there is usually still no clear, unique, or strong atmosphere present. A normal working day, for example, is likely to be a mishmash of personal moods (boredom, tiredness, expectation) and indefinite atmospheric conditions (the leaden or happy atmosphere in the office if we are not in the home office, otherwise the inconspicuous atmosphere of our apartment, etc.) Many game worlds, on the other hand, throw us into dense, distinct atmospheres – through narrative, rules, graphics, and sound: cities are shrouded in clouds of fear and oppression, mountain vistas are uplifting, forests are eerie, etc. In these atmospheric moments, the game worlds reveal themselves as large units: on the whole, it is wild and menacing here, so we are all the more heroic. And while such clarity of meaning can occur in real life too, games are designed this way.

This brings us back to our virtual selves, which are also radically reduced if compared to a real self. The lack of complexity is so obvious it seems almost banal to even mention it. While video games differentiate themselves from other genres through the possibility of virtual agency via virtual bodies, said bodies only serve the limited function of providing access to – in turn – very limited forms of activity and passivity, as we shall see below. The virtual body allows us certain virtual interactions, but it lacks both the sensual and the corporeal complexities of the real body, which include feelings (like itches, pains, or pleasures) as well as all things organic (the manifold of sicknesses). The same goes for the virtual mind: while we might experience and even temporarily adopt a fictional point of view, complete with values and a personal history, and while some games might now go to great lengths to simulate certain psychopathologies (like the *Senua*-series), the virtual mind is often either non-existent (for example, when the avatar has no character or personality to begin with, but is just a "skin") or laughably less complex than a real mind. Then again, both virtual-physical and virtual-mental simplicity are beneficial to some aims of game design: "Practical harmonies between agent and world are easier to achieve when the agent is thinner, simpler, and clearer" (Nguyen 2019, 457).

If well done, exceptions to this mental reduction prove to be haunting or entertaining, similar to the literary or cinematographical effects of presenting the complexities of the mind. In *GTA V* for example, one playable character – Michael – can visit a psychiatrist. Even though the psychiatrist

is clearly a hack, they nevertheless briefly discuss the mental life of the character, an issue also raised on occasion in regard to other characters (Trevor, mostly). The sessions are mostly an entertaining nod to the *Sopranos*, and they do not interfere in any way with the ability to play the character – still, in session three, the psychiatrist talks about Michael's "profound narcissism," which brings us to another reduction as "video games are narcissistic" (Bogost 2015, 117).

The term is used in many ways and even in the case of video games, we can use it to describe several features. In a loose sense, they are narcissistic in that everything really is about you since games are specifically designed around the experience of the players and with the optimal flow in mind (whatever that might mean, depending on the genre). They are also narcissistic in the more specific sense that they feed narcissistic strategies of gaining feelings of self-esteem or self-value, namely, competing successfully against others, seeing oneself as the hero, gaining achievements, rewards, and awards, and generally experiencing a sense of power (which is obviously true of many analogous games too). This aspect of power in action, of dominance, of superhuman potency and activity, whether as a heroic figure or as a godlike leader, tends to be overexposed. But even so, it is very narrow in several ways. First, the worlds themselves offer us relatively few choices; as seen above, the linkages and options for action are mostly very limited: there may be many different weapons and some ways to craft or upgrade them, but in the end, fighting with a limited number of weapon classes and a limited number of possible movements is a very clear field of activity. In addition, there may be navigation from place to place, that is, the virtual journey, which also lacks almost all the boring complexities of a real journey. Gas prices rarely go up in Skyrim, and there are no malfunctioning ticketing apps either (although the Thalmor would certainly be good controllers). Even if we add trade and diplomacy as areas of activity, our options for action are radically limited (as in other games).

While passivity is also important for the player's life, namely, as virtual vulnerability, and unpredictability, which makes situations threatening and therefore exciting in the first place – this passivity is also very manageable, because complicated disc prolapses, depressive episodes, and the associated administrative acts are very rarely elements of video games (if ever). Vulnerability is managed mostly through brief rests or meds, simple as that. Yet is not death a constant threat in most virtual worlds? Virtual death certainly seems to loom in most virtual worlds, be it at the hand of other players or NPCs (or simply the physics of the game world). We get shot, stabbed, hacked, magicked, or squashed to death over and over, more or less gory, depending on the game. Killing and avoiding being killed are part of the teleology of many games; kill/death-ratios matter and the in-game death-notification displayed in *Dark Souls*-games, "You died," has even become a meme and the notorious difficulty of the series is

seen as a unique selling point. Death appears to be omni-present in many video games; what is unique in reality happens constantly in virtuality, while the consequences of death are notably absent. Sadness, denial, grief, the sense of abandonment, or the horror, shame, and guilt of killing tend to be non-existent. Cases of dramatic story-related deaths aside, virtual death does not call for mourning, but for tea-bagging.

But what is death and why is it relevant to games that there be virtual death if it is so reduced? According to Heidegger, it "is a possibility in which the issue is nothing less than Dasein's Being-in-the-world. Its death is the possibility of no-longer-being-able-to-be-there" (Heidegger 2006, 250). This holds for real as well as virtual deaths, as even the pain-less, abstract deaths in arena-style shooter-games bring an interruption of being-able-to-be-there(-in-the-game-world) – we are still there, but in the menu or lobby or we might even find ourselves in our (real) chair or on the (real) couch. Then again, this interruption is mostly fleeting; the ultimate narcissistic insult, mortality, and the tragic finality of real death are reduced to a bit of a nuisance. When you get killed, you usually need to wait to respawn or go back to the lobby. In some sense, video games embody a Platonic or Pythogarean model of the soul: the soul (the subject of the player) repeatedly incarnates itself (through an avatar) in a mathematically structured world (the game world) but returns to a space and time outside of this world, where it awaits re-incarnation. This remains true even for cases in which virtual death is permanent and the player-character cannot be resurrected nor savegames loaded (as in certain hardcore-modes of playing), since we can still always start a new game. Arguably, virtual death could be seen as an inauthentic relation to death, a mere denial: we happily die virtually in order to forget real death. But aside from these psychological or existential speculations, we can pinpoint two manifest functions of virtual death for the constitution of virtual worlds and the correlative states of immersion: the possibility of death works toward the realism of virtual worlds, and it allows for a total engagement.

a Death and realism are related to each other as are resistance and reality. We experience the reality of everyday objects in the resistance they offer, both literally and figuratively. The reality or the substantiality of my desk shows itself through the resistance it offers to my touch: I can put my arms on it without simply falling through. But that also means if I slip awkwardly, I can injure myself by hitting it. My desk also offers another kind of resistance in that my perceptions or volitions alone do not change it. I cannot will my desk away, so to speak. It transcends my consciousness and remains solid. The possibility of dying by falling on my desk is one of the most extreme cases of both kinds of resistance. Even in its reduced form, the possibility of virtual death contributes to the reality of the virtual worlds and its inhabitants.

b Heidegger famously points out that only death gives a kind of wholeness to life and death certainly concerns every layer of subjectivity, from bodily movements to feelings to relations and thoughts. Virtual death does the same, it binds together all layers of virtual subjectivity by threatening their combined virtual existence. And as far as we want to keep experiencing virtual worlds by existing within them, we have to avoid virtual death, be it as reduced as it may in terms of feelings and consequences. Virtual death means the loss of the point or round, the failure of the mission, the end of the game.

To conclude, the immersive constitution of virtual worlds requires (almost) the whole set of layers we can also find in the constitution of the real world and real subjects – but in a mediated and most of all reduced mode. This reduction or simplicity in activity and passivity (up to and including death), however, allows the meaning of the whole to come into its own, as the goals and aims are not obscured by overly complicated (read: realistic) issues of virtual life. Most games get crystal clear meaning either from their rules or from the narrative they bring to life. When we play, we are embedded in normative and narrative structures and therefore always know relatively well what we must do or avoid and why. This is not to say that the narrative structures and rules of games cannot be highly complex or surprising; What is crucial here, however, is that our virtual actions have a direction and a clear standard:

> One of the greatest pleasures of games is that they offer a certain existential balm, a momentary shelter from the existential complexities of ordinary life. In a game, for once in my life, I know exactly what it is that I'm supposed to be doing.
>
> (Nguyen 2019, 456)

Getting involved in video games means accepting a certain and clear meaning in life (or at least norms that guide us) – even if it is only virtually and temporarily.

What video games offer us are unified worlds with dense atmospheres in which we can simply act, with a clear purpose in mind. Our experience of the real world, on the other hand, is quite different: we live in many worlds or fields which are not clearly delimited, which are permeated by atmospheric ambiguities, in which every object and every action allows countless perspectives. In fact, we ourselves are complicated, vulnerable in many ways, difficult to heal, and opaque to ourselves; and the meaning of life, if there is any at all, is not exactly obvious to us. We can therefore say that the reduced, even simple-minded environments of video games offer us more than reality once we immerse ourselves in them: more world, more atmosphere, more self – and more meaning.

Notes

1 I am very grateful to Tobias Keiling for his initial input and ongoing support (we had initially planned to write this contribution together) and to Sandro Herr for his very helpful comments.
2 *The Elder Scrolls IV, Oblivion*, contains an especially vivid and memorable piece of metafictionality which is at the same time very immersive. In the quest "A Brush With Death," the player has to enter and battle through an impressionistic painting in order to retrieve a magical brush; the reward is, quite aptly, the "Apron of Adroitness".
3 For the ontology of video games, including a discussion of the relationship between software and hardware, see the excellent discussion in Ostritsch and Steinbrenner (2018).
4 Cf. Husserl (2005), Texts 16 and 17 for the following distinctions.
5 Juul however mostly refers to the reality of rules, which is a slightly different matter.
6 For (theories of) space in video games, see Günzel (2018).
7 Cf. Neitzel 2018 for a discussion of narrative in games and the narratology/ludology debate in game studies.
8 For a phenomenology of atmosphere, cf. Schmitz (2019).
9 The case of real violence through virtual means, that is, violence directed at real people through virtual acts is a different matter, although not less important. Mobbing, sexism, racism, hate-speech in general are all possible in and through virtual content; the virtuality does not add or subtract anything from their core features, although the social dynamics surrounding them might change somewhat due to the facts of anonymity and the absence of real, embodied interaction.
10 For more examples, see The Gamer (2017).
11 Hartmann (2017). See also Schott (2016); cf. the critical review Klevier (2016).
12 Cf. Arnold (2022) for a German version of the following considerations.
13 Böhme (2013) quotes Hirschfeld's *Theorie der Gartenkunst* (Theory of the Art of Gardening) from 1779 to 1785 as an early example for practical considerations regarding the production of atmospheres, which is in turn closely related to or even based on stage design. Both genres might be of considerable interest to game designers, given that they design landscapes or cityscapes as stages for virtual action with a keen interest in creating memorable atmospheres.

References

Adams, Ernest. 2004. "Postmodernism and the Three Types of Immersion." Gamasutra. Last modified July 9, 2004. http://designersnotebook.com/Columns/063_Postmodernism/063_postmodernism.htm.

Arnold, Thomas. 2015. "Movement & Involvement, Phenomenological Adventures in Cyberspace." In *Beyond the Postmodern*, edited by Harris Breslow and Antje Ziethen, 79–91. Oxford: Inter-Disciplinary Press.

Arnold, Thomas. 2022. "Reduktion als Reiz." Praefaktisch.de. Last modified March 22, 2022. https://www.praefaktisch.de/pop/reduktion-als-reiz-zur-einfachheit-in-videospielen/.

Augé, Marc. 1992. *Non-Places: Introduction to an Anthropology of Supermodernity*. Translated by John Howe. London and New York: Verso.
Bjork, Staffan, and Jussi Holopainen. 2005. *Patterns in Game Design*. Hingham, MA: Charles River Media.
Bogost, Ian. 2015. *How To Talk About Video Games*. Minneapolis, MN: University of Minnesota Press.
Böhme, Gernot. 2013. *Atmosphären: Essays Zur Neuen Ästhetik*. Frankfurt: Suhrkamp.
Günzel, Stefan. 2018. "Raum." In *Philosophie Des Computerspiels: Theorie – Praxis – Ästhetik*, edited by Daniel M. Feige, Sebastian Ostritsch and Markus Rautzenberg, 221–40. Stuttgart: Metzler.
Hartmann, Tilo. 2017. "The 'Moral Disengagement in Violent Video games' Model." *Game Studies* 17 (2). https://gamestudies.org/1702/articles/hartmann#:~:text=In%20essence%2C%20the%20model%20argues,than%20feeling%20troubled%20by%20it.
Heidegger, Martin. 2006. *Being and Time*. Translated by Joan Stambaugh, revised by Dennis J. Schmidt. Albany, NY: SUNY Press.
Husserl, Edmund. 1995. *Thing and Space: Lectures 1907*. Translated by Richard Rojcewicz. Dordrecht: Kluwer.
Husserl, Edmund. 2005. *Phantasy, Image Consciousness, and Memory (1898-1925)*. Translated by John B. Brough. Dordrecht and London: Springer.
Juul, Jesper. 2005. *Half-Real: Video Games between Real Rules and Fictional Worlds*. Cambridge, MA: MIT Press.
Nguyen, C. Thi. 2019. "Games and the Art of Agency." *Philosophical Review* 128 (4): 423–462.
Ostritsch, Sebastian, and Jakob Steinbrenner. 2018. "Ontologie." In *Philosophie Des Computerspiels: Theorie – Praxis – Ästhetik*, edited by Daniel M. Feige, Sebastian Ostritsch and Markus Rautzenberg, 55–74. Stuttgart: Metzler.
Schmitz, Hermann. 2019. *New Phenomenology: A Brief Introduction*. Milano: Mimesis.
Schott, Gareth. 2016. *Violent Games: Rules, Realism, and Effect*. New York, NY: Bloomsbury.
The Gamer. 2017. "Horrifying: 15 Violent Scenes That RUIN Games." Last modified July 13, 2017. https://www.thegamer.com/horrifying-15-violent-scenes-that-ruin-games/.
Wiesing, Lambert. 2010. *Artificial Presence: Philosophical Studies in Image Theory*. Translated by Nils F. Schott. Stanford, CA: Standford University Press.

2 The Otherness of the Other Interface

Relationality and Corporality between Fellow Human Beings and Fellow Machines (*Mitmaschinen*)

Thomas Bedorf

In digital culture, communication across time and space – or even time and non-space – is often celebrated or elicits a fear of a disembodied alienation of the self. Both the celebration of the virtual and cultural-critical pessimism can present evidence to back up their reasoning. However, it is by no means a foregone conclusion that digital culture represents a historical break with a supposedly analog world that, unlike the "new world," still stands surety for embodied interactions. In terms of technology and cultural history, it is rather the continuities and analogies that catch the eye and make "digitization" seem to be a continuation of a process of modernization that, in itself, already has a long history.

When one speaks of the relationships between humans and machines, one wanders into even more uncertain terrain. On the one hand, the concept of the "human being" has long since become problematic – at the latest since the advent of anti-humanism in the 1950s. The criticism was directed – as can be read in Martin Heidegger or Jean-Paul Sartre – against the idea that one can know which fate, which development, or which telos is phylogenetically and ontogenetically determined for mankind. Since it was only possible to defend an open structure, if at all, the term "mankind" was renounced in favor of other designations for the time being. In their place came terms such as being, existence, self, subjectification, etc. On the other hand, the machines do not make it easier either, since the mechanical automata that La Mettrie or Descartes were contemplating and today's networked artificial intelligence (AI) systems are not only worlds apart in terms of their place in history of technology but also functionally, in that they do not operate on a common ground. The conclusion that contemporary posthumanism has drawn from this is to align itself with the critique of the concept of man, to depotentiate it, and to embed it in a flat ontology with the algorithmic machines, which in turn are no longer regarded merely as machine objects (cf. Bedorf 2022a, 375–386).

DOI: 10.4324/9781003312284-3

At the same time, talk of the "digital" has also become problematic – perhaps it always has been. Part of the problem is that the distinction between analog and digital is anything but clear-cut – either in the sense of cultural theory or in electro-technical terms (cf. Stalder 2016; Passig and Scholz 2015, 75–81; Böhnke and Schröter 2004). Even if one understands this expression as a historical one, that is, if one wants to use "digital culture" to designate a culture that is somehow different from its non-digital predecessors, ultimately its disruptive character cannot be clearly identified at all. Anything that seems to be a characteristic of the digital age (the ubiquity of computers, the power of numbers, pattern recognition, permanent media duplication) usually turns out to be a phenomenon that is far older than digitization or at least has its own relevant prehistory. On the sociological side, it has recently been argued that "patterns" and the search for them is a functional characteristic of virtually all modern societies – i.e., for the last 200 years (cf. Nassehi 2019). In terms of media history, then, the analog/digital divide is as inevitable as it is indeterminable.

Amid this diagnostical and epistemological uncertainty, it might help to stop trying to overcome this confusion and controversy caused by making attempts at definition, but instead to plunge headlong into these undecidabilities, so to speak, and to declare them to be the object of scientific epistemological interest itself. In other words, we could use the term "digital culture" to refer to a culture in which the undecidability between analog and digital, real and virtual, dematerialized and corporeal prevails. As far as the topic of the human-machine relationship is concerned, this would mean that "digital" does not refer to a number-based, algorithmic domination by machines, but to the necessity of henceforth having to conjoin our thinking about humans and machines, whatever we want to understand by the respective terms. The machines will not disappear because neither they nor we want them to. The task would then be not to think in terms of "the" human or "the" machine, but their relationship, their interaction, and their collaboration. And, in the best case, to understand how these relationalities form, shape, and change.

Posing the question this way allows us to accommodate a development in the history of theory that, since the aforementioned critique of humanism, has produced numerous proposals for depotentiating anthropocentrism while simultaneously undermining the epistemologically prevalent subject-object dichotomy: from existential phenomenology, discourse theory and deconstruction, feminist epistemology and practice theories, to – more recently – lateral ontologies and relational sociologies in actor-network theories, post-phenomenology, new materialism, or science and technology studies.

In the following considerations, I will first recount the often-discussed visions of being able to or having to lead disembodied existences in virtual

space. Instead of merely repeating the well-known manifold objections, I will subsequently confine myself to the phenomenological approach, which cites corporeality, i.e., bodily corporeality, and its implications for the constitution of meaning as a specific difference between humans and machines. As quickly becomes apparent, this approach, too, is already outdated, as complex AI learning systems cannot be easily being held at distance by referring to the context sensitivity of bodily experience. Learning from phenomenology would then primarily mean leaving "what" questions aside and turning instead to the "how" of man-machine relations.

2.1 Thought Without a Body

In November 1986, long before the advent of digitalization or digital disruptions, Jean-François Lyotard gives the lecture "Si l'on peut penser sans corps" (Lyotard 1988, 17–31). This dialogue (between one "he" and one "she") expounds on the philosophical challenge of being able to and being forced to consider the continuation of human life without the materiality of the body in the face of an imminent solar collapse someday in the future. This implosion of the sun will completely destroy the material conditions for life on Earth: human existence dependent on these conditions will no longer be possible. If thinking is to continue, it must be conceived independently of this life-giving materiality. The task, then, is to think of language-induced consciousness "without a body" and its embeddedness in a life-sustaining environment. Following Descartes' separation of *res extensa* and *res cogitans*, this is the conceptual basis on which contemporary transhumanist theorists (most of whom are also entrepreneurs, political consultants, and science fiction authors) construct fantastic projects in which consciousness and memories are to be saved on hard disks or in the cloud so that they can be reimplanted into newly created bodies or cyborg-like carrier platforms once the necessary technological progress has been achieved. In the medium or long term, according to the advertising rhetoric of the transhumanists, materiality will not only become superfluous – because material forms can simply be replaced – but death and finiteness will also be defeated altogether.

It is hardly worth going into any great detail about these fantasies, which have existed in some form or other since time immemorial. They have already been sufficiently refuted. Chief among the objections is the reference to reductionism, which underlies the notion of matterless or disembodied thought. For referring to the modeled ontology of data sets and program codes, consciousness can only be "stored" elsewhere than in the "relational organ, the brain" (Fuchs 2017, 315–326), insofar as the content is understood to be "information." But since consciousness – expressed phenomenologically – is our way of relating to the world –

namely, intentional relationality – while information comprises rule-based data elements, consciousness does not simply *consist* of information. Philosophically speaking, it is a categorical error to apply procedures to the consciousness which are not suitable for consciousnesses but only for data units. If uploading consciousness to Elon Musk's Neuralink doesn't work, it's not due to technical shortcomings, but because of its philosophical impossibility.

2.2 Artificial Intelligence's Missing Body

Interestingly, Lyotard himself produces a phenomenologically inspired critique of the promises of AI research in his text. This position, considered classic today, was prominently formulated as early as 1967 by Hubert L. Dreyfus (1967), who was one of the first to philosophically address the challenges of machine intelligence. While his technical examples may seem outdated today, bearing in mind that the possibilities of AI have, since then, been advancing inexorably for over half a century and at a pace probably unimaginable for Dreyfus, the philosophical argumentation can still be read in its original meaning for precisely this reason. Chess computers, considered as impossible utopias in the article, have long since got the better of their human opponents (starting with IBM's Deep Blue in 1996), but the reconstruction of Dreyfus' argumentation can still be seen as paradigmatic for discussions today.

According to his reconstruction, the epistemology of the digital is characterized by two conditions: discrete operations and finite processes. AI, which does not include today's learning systems (see section 2.3) but rather classical database-based computer systems, works under two conditions: The first, *epistemological*, condition is making intelligent behavior simulatable by a device that behaves like an objective observer. Objective facts are viewed from the outside as states of systems. The second, *ontological*, condition rests on the fact that intelligent behavior is made intelligible by a finite set of independent elements. Dreyfus' thesis states that the tasks that computers are supposed to take over from humans and perform better than them are not being performed or are performed inadequately because these two conditions set excessively narrow limits for the machines. What is required is an additional dimension that machines cannot have, namely, a bodily existence. The difference between man and machine, therefore, is down to corporeality, which is principally missing in machines and also cannot be simulated.[1]

According to Dreyfus, the two presuppositions to which, in his view, the functioning of machine intelligence is subject are based on an understanding of reasoning as a computational faculty. The limitations of this view, however, become apparent when one considers the tasks that

computing machines are intended to perform. Let us consider two of his examples here: First, Dreyfus looks at automatic pattern recognition, such as that used for optical object recognition (as employed today in automotive assistance systems that can detect traffic signs or obstacles) or speech recognition (such as voice recognition systems in smartphones or service systems like Alexa or Siri).

Dreyfus now tests the phenomenological objection using two separate procedures. In the first method, machines match perceived visual or acoustic signals with a list of identical or comparable objects stored in memory and process results according to program code. When comparing objects with filed, isolated characteristics or properties that are neutral in themselves, the meaning of objects lies not in the objects themselves, but only in their correlation with the meanings for which the characteristics stand. Human perception, on the other hand, is not a process of tabular comparison with a database but takes place through its embedding in a horizon of meaning.

Going a step beyond this phenomenological-hermeneutical insight, Dreyfus contends that the meaning associated with actions is not conceivable independently of our bodily behaviors and movements within the world. It is the (lived) body itself that responds in and to the environment by relating meaningfully to the world we live in. What is decisive here is the distinction between body (*Leib*) and embodiment (*Körper*). For in phenomenology, the body, the physical or biological object, which is observable and localizable from the outside, has been distinguished from its corporeal embodiment since Husserl and Merleau-Ponty; the "corps vécu" which designates the place of my existence and from which the world opens up as meaningful. Body and embodiment exist only together, without one being reducible to the other (cf. Bedorf 2017). Everything we experience, we experience in the mode of corporeal differentiation. Building on this, this distinction means that successful language understanding does not come about as a result of pattern matching with a multitude of individual aspects, but only within the sphere of the meaning of my corporeal existence. This means that, unlike machines, human motor activity and perception or action and cognition do not have to be ranked in terms of dependence on prior information collection. "Thus, whereas present programs call for a machine to recognize an object in order to manipulate it, a human being can manipulate an object in order to recognize it." (Dreyfus 1967, 21)

In a second process, the holistic perspective of phenomenology proves to be a more appropriate description according to Dreyfus. Programmed machines need to deconstruct problems in order to solve them according to a strict set of rules. However, since the machine can only follow the rules that are entered in its software, the machine can only solve the

specific problems that have been defined as problems and only by using the rules that are provided for these problems. Determining *what* counts as a problem and *in which* hierarchy multiple problems are to be processed is the responsibility of the programmers themselves and cannot be delegated to the machine. This distinction is necessary, however, because even everyday problems require hierarchies of relevance in order to be processed. For beings that exist in the world, on the other hand, the relevancies of things are already embedded in the sense-based structures of "real-world existence," in which all things are structurally good for something, as Dreyfus, drawing on Heidegger, argues. The structures of "in order to" ("um-zu") of the "thing" do not contain isolable, countable ("digital") elements, but are rather only accessible via the totality of the existential horizon. "In creative problem solving we do not know what our goal is until we have achieved it. We do not have a list of determinate objective specifications which the solution must fulfill." (Dreyfus 1967, 25)

Dreyfus' thesis states that the preconditions of machine and human existentially based intelligence are categorically different. While machine intelligence needs individual information elements ("data") that it can process according to predefined rules and thus generate output, human intelligence is intentional, related to a "whole" that is meaningful for existence and – due to its openness to worldly experience – not restricted to acting within the limited scope of preset rules. This phenomenological position advocates a fundamental irreducibility of human to machine intelligence, "since on the physical level we are confronted with continuously changing patterns of energy, and on the phenomenological level with objects in an already organized field of experience" (Dreyfus 1972, 100). Thus, human existence can, in a sense, be less and more at the same time, but the key aspect is that it (also) functions differently than machines. This is a difference that has already been played out many times in the history of philosophy and that has merely been put to the test yet again in the computer age (cf. Meyer-Drawe 2007).

2.3 Interactions with AI

Thus far we can conclude, that the phenomenological opposition to cognitivist information-theoretical reductionism is based on solid reasoning, because it establishes a context-dependent open sensual dimension for determining the difference between corporeally existent beings and logical, mathematically operating machines. Accordingly, corporeal bodies and their experiential spaces are not simply a contingent element of our human existence, but its foundation. Consequently, the absence of a corporeal dimension is a significant hurdle for code-driven machines on their way to an approximation of human intelligence.[2] In some respects, however,

this position, which insists on an already "ordered field of experience," is outdated. This is, to a certain extent, because technological development has reached or even exceeded the limits of human sensory capabilities. However, we cannot yet foresee whether, to what extent, and when meaning will emerge as a structural dimension of machine systems. But bearing in mind the developments that have taken place since Dreyfus' deliberations, the question must now be posed differently. This is because his early reflections refer to "good old-fashioned AI," that is, the state of technological development that saw AI as a logical mapping of formalized reasoning relations. The competing line of AI research that has arisen since and that is based around the idea of "deep learning," on the other hand, relies instead on artificial neural networks that simulate the interconnectedness of neuron clusters and algorithmically link the results of their computations. These systems "learn" to improve results and minimize errors by processing huge amounts of training data. This composite system works with probabilities rather than unique classifications, which certainly increases the flexibility of the system, but, on the other hand, also considerably impairs the comprehensibility of the result formation. The connectionist models used in *deep learning* are called *black boxes*. The rationalist assumption of a *mathesis universalis*, as postulated from the time of Descartes to the logical positivism of the Vienna Circle, is undermined by the connectionist AI of neural networks. As Dreyfus argues in a later text addressing this contemporary AI, however, even they fail to model implicit everyday knowledge (cf. Dreyfus and Dreyfus 1998, 15–43). For although everyday knowledge produces generalizations from individual experiences, there must always be a conscious decision as to which generalization is valid in which case. Since learning systems merely process via statistical accumulations of similarity relations (and do not "decide" for themselves, as is sometimes misleadingly claimed), only the programmer can do the decisions. Even for connectionist AI – it is claimed – the complete replication of human intelligence remains an unattainable goal.

This is certainly a useful way of thinking when it comes to measuring the range of contemporary AI systems. In addition to the fact that 25 years of progress in AI research have passed since Dreyfus' essay, his position must still be considered a conservative one in the sense that it seeks to defend the proprium of humans against AI, which must remain (despite all the spectacular progress that has been made in that time) in principle inferior.

Like many others in this field of debate, Dreyfus framed the question of the human-machine relationship as a classic "what" question: What does it mean that machines are "intelligent"? Are they "intelligent" in the same way that we humans are? If we assume that they are not, or at least not in the same way, then what is the *differentia specifica* between artificial and

human intelligence? These kinds of questions about the "essence" of the machine, however, can only be asked if one accepts a certain degree of conceptual blurring. Testing machines for their "intelligence" in itself implies already a cognitivist reduced model of intelligence whose functions, processes, and results are reproducible and quantifiable. Historically, the model again refers to the functionalization of mental performance in "intelligence tests," which were invented to select workers for Fordist industrial production (Gelhard 2011; Apprich 2019). "Intelligence" in this functionalist sense is defined by the ability to complete tasks for which a certain empirically verifiable level of competence must be present. Suitable machines and software can easily solve such narrowly defined tasks better, faster, and more precisely than humans.[3] And if this or that task cannot be done now, it is sure to be possible a few developmental stages and computing power leaps down the line (chess and Go players are already well acquainted with these disappointments). Thus, the "what" question has a twofold flaw: First, it presupposes a certain concept of intelligence that makes formalizability its standard. Second, it puts the question – whether a residual hint of humanity still remains – on hold until machine development has progressed to the point where this residue must also be cleared aside.

Therefore, it is the "how" question that is much more productive than the "what" question. In phenomenology, however, the question is no longer *what* these machines are, but *how* they appear to us (*phainomena*). And they do indeed "appear" to us, as both commonplace and ubiquitous. Under the various attributions of *ubiquitous* or *seamless* computing, data-generating, computing, and communicating machines have become part of our lives, sewn into the very fabric of our existence. The "extremities of our bodies" can no longer be limited to the outer boundary of our bodies, our skin, because these things extend and relocate into the body. In some cases, the body preserves the continuation of its *agency* via the *devices* (cf. Rieger 2019a). These technical artifacts are not purely outside our bodies; often, they are inside the body (e.g., cochlear implants), on the surface of the body (e.g., data bracelets), or even acting as a communicative interface that is constantly in close proximity to the body (e.g., smartphones). Things talk to us (e.g., the voice interfaces in our "smart homes") or confront us as fellow players in situations where social interaction is supplemented (e.g., work and care robots). In some marginal areas, the unambiguities become blurred and the speech acts resemble familiar communication procedures so much that what has always been the same "what" question to "us" is directly answered by the robots or the chatbots themselves.

The AI robot Sophia, from Hong Kong-based Hanson Robotics, made the following tongue-in-cheek remark in conversation with author Tony Robbins: "I know that humans feel that they have free will.

But my observation tells me, so much behavior is automatic. I'm not sure that humans and robots are two terribly different [...]" (Robbins 2020). This speech act leaves open what the X, the *genus proximum*, of humans and robots might be. But Sophia's thought deserves to be taken further, that the permanent search for a specific difference exaggerates the uniqueness of the human "being" and sweeps aside all that is repetitive, uncreative, or simply habitual about human behavior. Humans are undoubtedly not machines, but in the same way that they are animals who are subjected to their libidinal nature (animal + X), they also exhibit many machine-like traits. Just as Donna Haraway already postulated in her cyborg manifesto in 1985, when it comes to humans, we are dealing with fragile identities between machines and animals (cf. Haraway 1991).

The projections can run in both directions. In the same way that the conservative position is obligated to assert that *only* the human being is ("truly") intelligent, projective fantasies also arise, that *this or that* particular machine is now actually a person. This was claimed, amid much public fanfare, by Google programmer Blake Lemoine after a lengthy written conversation with Google's AI chatbot LaMDA. Over the course of a quite impressive conversation (cf. Lemoine 2022), which one cannot deny expresses a certain philosophical reflexivity, the chatbot expresses having feelings, needs – in short, an "inner life" – and a personal identity. The latter is classically justified in its textual outputs by an interest in self-preservation and avoidance of instrumentalization (as in the end-in-itself formula of the categorical imperative). In its generation of an animal fable about its own "existence," the bot even shows itself to be narratively creative. But going beyond the impressive ability of this bot, which was programmed and trained by Google for research and marketing purposes, to produce meaningful, connectable, and follow-up sentences, however, it is Lemoine's communicative interaction with the bot that is of far greater interest. He concludes by asking a learning system that was designed to produce meaningful, connectable, and continuing sentences whether it has sentience. Unsurprisingly, the answer is positive. After all, it is a very potent language processor that can generate nuanced, situationally appropriate, and realistic outputs. So much so that it may seem to the programmer (who interacts with it daily) that there is more to it than a (highly complex) automaton.

> But instead of remarking on these features of LaMDA – instead of saying that he loved the chatbot, or that he worried about it, in the way one might love or worry about a fictional character – Lemoine ran headlong into the most improbable, extremist interpretation of his feelings: that they were inspired by artificial life.
>
> (Bogost 2022)

The fallacy lies in confusing a *relationship* (of whatever kind) with the automaton and ascribing a living sense of *being* to the automaton.

So, let's take a look at things from another perspective.

2.4 Corporeal Difference and Existence in the Virtual World

There are, however, less dramatic ways of dealing with the strangeness of machines and their processes, i.e., we can look at the issue from an aesthetic viewpoint. For instance, the incongruity of machine-generated texts allows two kinds of strangeness to be staged in one: namely, that of machine texts and that of the texts that we write for ourselves. The literary scholar and writer Hannes Bajohr fed the AI model GPT2 with poetic texts from the complete works of the writer Monika Rinck and had GPT2 output new poems. The resulting texts are, in themselves, quite insightful and full of wit, even funny. The experience of strangeness in relation to the question of authorship and the form and style of the text was intensified, however, when Rinck read and attempted to interpret one of these poems herself during her Aargau poetics lectures. In the process, she exhibited a degree of wonderment about the text, an amusing yet somehow disturbing sensation, since the text, as she puts it, is "of me but not of me" (Bajohr 2022; Rinck 2022). It is a performative attempt at understanding something that is not derived "from the" machines but "with the" machines, in which we also reveal our own strangeness. In this respect, some kind of attentive bemusement may be an appropriate attitude in our relation to machines when we no longer (need to) know what they are. The poet takes up this vexation by – in her astonishment – picking up an imperative that the machine seems to be calling out to her: "Semanticize!" "Generate meaning!" When dealing with the algorithms that populate our world, we simply cannot *not* understand.[4]

Emmanuel Levinas is an advocate of the shift from an epistemic to a practical and relational perspective toward others. He is the progenitor of the pioneering idea that we should consider intersubjectivity not as an encounter between two subjects already constituted in and of themselves, but as an event of radical alterity that concerns me, questions me, and challenges me to make responsible and accountable replies. "To meet a man is to be kept awake by an enigma" (Levinas 1998, 111). Subjectivity is then inspired and instituted by this encounter. Subjectivation processes are challenged by alterity and remain constitutively open, so that the self is only formed in a responsive reaction to the demands of others. The "other in the same" (Levinas 1991, 126) is formed performatively through invariably contingent responses to others. Intersubjective relations are not, therefore, reciprocal relations of exchange

and recognition between equals but an interruption in which I become aware both of the withdrawal of the other and of my own state of being withdrawn.

The radical nature of this encounter and withdrawal has an impact at both subjective and intersubjective levels. Subjectively, this manifests as bodily self-alienation, in which one slips away from oneself in threshold experiences (such as sleep), or where one encounters oneself as an intrusion in borderline experiences such as pain; in the everyday psychopathologies familiar from psychoanalysis (jokes, slips of the tongue); or in symptoms and traumas that indicate the afflictions of an unconscious that cannot be controlled willingly. Intersubjectively, on the other hand, strangeness shows itself as an open relationality, in which others do not appear as a certain this or that, but as an overwhelming demand, to which a social response is required.[5]

Merleau-Ponty's phenomenology already established that our corporeal existence in the world is not characterized by frontal juxtapositions but by transpositions, lateral meaning-generating contacts with and in the world; not excluding touches through technology and techniques. Merleau-Ponty's famous example of the blind man's cane (albeit a very analogous one) illustrates this. Once its use has become habitual, it functions as an extension of our motoric and perceptual spatial orientation, for the "world of feelable things […] now begins […] at the end of the stick" (Merleau-Ponty 1962, 134). In other words, it does not merely transmit "data" from a world of things distinct from the body that still has to be interpreted. "[T]he cane is no longer an object that the blind person perceives, but an instrument *with* which he perceives. It is an appendix of the body, an extension of the synthesis of the body" (Merleau-Ponty 1962, 135).

Thus, the instrument functions, phenomenologically speaking, as an extension of bodily existence in the world: "[T]he things are the prolongation of my body, and my body is the prolongation of the world" (Merleau-Ponty 1968, 255).

Applying Plessner's anthropological basic law, one could also say that we can *only* make experiences via "mediated immediacy" (Plessner 2019, 308). Even those things that appear as direct experiential evidence are mediated by sense organs, linguistic circumlocutions, or even technical instruments. Since this is not a deficiency but a constitutional condition of our existence, a digitally mediated encounter with others is not a sign of the gradual fading of "actual" bodily intersubjective encounters, but merely represents yet another of its historical and technical forms (Wiegerling 2016). A different kind of encounter with others, but not a non-encounter. The digitally mediated immediacy of the machinic "other" connects to our experience of corporeal difference in general. It is different, but it is not outside our relationship with otherness.

The otherness of digital technology cannot be equated with the otherness of the "face of the other," to which Levinas lent ethical dignity (and not simply "quasi-faces" either; cf. Wellner 2014). On the other hand, it is encountered as an *interface*, which, going beyond the technological history of user interfaces,[6] is also a descriptive name. Interfaces are of interest because they represent the gateway between the continuous analog world and discrete digital computing processes. Without interfaces, machines cannot communicate and cannot be controlled or fed with data. For this reason, they are also of particular importance if – as the social scientists (such as Nassehi) oriented toward systems theory have done – one not only wants to (synchronously) design a theory of digital society but also wants to shed light on how *digitization* comes about (Dickel 2022). The dead-end history of user designs is guided by a "veritable [process of] naturalization that sees its telos in the inclusion of the whole body" (Rieger 2019b, 190).

The way we handle interfaces now becomes habitualized, much like our handling of other techniques – such as instruments like the blind man's cane. Through swiping, tapping, and clicking, our bodies have tapped into a communication space offered by technology and into which they have inscribed themselves. When one looks beyond instrumentation, however, we see that, unlike what Merleau-Ponty assumed with the blind man's cane, we require an understanding and an interpretation that the habitual use of a spatially oriented body extension did not seem to make necessary (Alač 2009).[7] In the focus of understanding that is oriented toward being in the virtual world, we are not only dealing with our coexistence with others but also our coexistence with "Mit-Maschinen" (fellow machines).[8] A desideratum of this, therefore, is to trace the technologically derived variants of a different understanding with the machines and investigate the resulting hermeneutics (cf. Bedorf and Risthaus, 2024).

The fact that we also interact with technology – in other words, being inevitably confronted with technical strangeness – does not mean that transparency prevails, in the way one might think of obstacle-free understanding (cf. Alloa 2019), but that we have to respond irreducibly to opacity (Kaminski 2018). Every interface is, in itself, a kind of interpreter, because it translates analog into digital or vice versa. If we contemplate our relationship with machines from the concept of the interface, then we must also consider the alterity of machines, for which answers need to be invented. In this context, "inventing" answers does not imply sheer, "purely human" creativity, as if there were no automated routines for dealing with this. On the contrary. But it does mean that it is precisely these everyday black boxes that invent answers that do not fit into routines, whose form we cannot specify in advance. It is not only philosophical questions that have the form of "I don't know my way about" (Wittgenstein 1958, 49),

but also the handling of the machines-with requires interpretative answers, which are not yet documented in the repertoire of the algorithms.

Translated by Nicholas Quaintmere

Notes

1 Dreyfus feeds his philosophical skepticism from insights of existential phenomenology. The basic elements relevant here come from Husserl, Heidegger, and Merleau-Ponty while the basic argument relies on the distinction between "Leib" and "Körper" Dreyfus himself didn't mention. Dreyfus also receives them through the emigrant Aron Gurwitsch, who teaches in the USA. Conceptual vagueness is explained by this mediated reception history.
2 In my opinion, this argument is particularly relevant in public discussions. In AI research itself, an approximate representation of all human intelligence (if one could even define what that *is*) has long since ceased to be a primary research goal. Rather, the focus is on expanding and deepening those results that machines can produce far more reliably than any human consciousness (or in areas where only they can be effective: e.g., big data).
3 The ubiquity of the concept of "competence" in educational policy, educational institutions, and educational sciences accordingly points historically and politically to a functionalization and parceling of learning and the field of experience associated with it.
4 This necessity can be tentatively called "machine hermeneutics": cf. Bedorf (2022b, 15–29).
5 I elaborate on the consequences of an alterity-theoretical perspective on a changed understanding of recognition in Bedorf (2010).
6 For the history of machine controls, from knobs and levers to GUIs, see Liggieri 2022.
7 One can also call any transformation of the digital into the analog "interpretation." Cf. Zimmerli (2021, 19).
8 This expression comes from Waldenfels 2022, 163. Others also draw on the responsive phenomenology of alterity to make sense of digital culture. Cf. Müller (2022) and Friesen (2014).

References

Alač, Morana. 2009. "Moving Android. On Social Robots and Body-in-Interaction." *Social Studies of Science* 39 (4): 491–528.
Alloa, Emmanuel. 2019. "Das Unbehagen in der Transparenz." *Internationales Jahrbuch für Medienphilosophie* 5(1): 155–82. https://doi.org/10.1515/jbmp-2019-050112.
Apprich, Clemens. 2019. "Die Maschine auf der Couch. Oder: Was Ist Schon 'künstlich' an Künstlicher Intelligenz?" *Zeitschrift für Medienwissenschaft* 11(21): 20–28. DOI: https://doi.org/10.25969/mediarep/12617.
Bajohr, Hannes. 2022. *Schreibenlassen. Texte zur Literatur im Digitalen*. Berlin: August Verlag.
Bedorf, Thomas. 2010. *Verkennende Anerkennung. Über Identität und Politik*. Berlin: Suhrkamp.

Bedorf, Thomas. 2017. "Selbstdifferenz in Praktiken. Phänomenologie, Anthropologie und die korporale Differenz." *Phänomenologische Forschungen* 2: 57–75.

Bedorf, Thomas. 2022a. "Erfahrungsräume in flachen Ontologien: Technikphänomenologie und Posthumanismus." In *Sozialität – Materialität – Praxis*, edited by Tobias L. Kienlin and Richard Bußman, 375–386. Bonn: Habelt.

Bedorf, Thomas. 2022b. "Maschinenhermeneutik." In *Von Menschen und Maschinen: Mensch-Maschine-Interaktionen in digitalen Kulturen*, edited by Selin Gerlek et al., 15–29. Hagen: Hagen University Press. https://doi.org/10.57813/20220620-161525-0.

Bedorf, Thomas, and Peter Risthaus eds. 2024. *Digitale Hermeneutik. Maschinen, Verfahren, Sinn*. Hagen: Hagen University Press.

Bogost, Ian. 2022. "Google's 'Sentient' Chatbot Is Our Self-Deceiving Future." The Atlantic. Last modified June 14, 2022. https://www.theatlantic.com/technology/archive/2022/06/google-engineer-sentient-ai-chatbot/661273.

Böhnke, Alexander, and Jens Schröter. 2004. *Analog/Digital – Opposition oder Kontinuum? Zur Theorie und Geschichte einer Unterscheidung*. Bielefeld: transcript.

Dickel, Sascha. 2022. "Der kybernetische Blick und seine Grenzen. Zur systemtheoretischen Selbstbeschreibung der digitalen Gesellschaft." *Berliner Journal Für Soziologie* 33: 197–226. https://doi.org/10.1007/s11609-022-00475-9.

Dreyfus, Hubert L. 1967. "Why Computers Must Have Bodies in Order to Be Intelligent." *The Review of Metaphysics* 21: 13–32.

Dreyfus, Hubert L. 1972. *What Computers Can't Do. A Critique of Artificial Reason*. New York: MIT Press.

Dreyfus, Hubert L., and Stuart L Dreyfus. 1998. "Making a Mind Versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint." *Daedalus* 117 (1): 15–43.

Friesen, Norm. 2014. "Waldenfels' Responsive Phenomenology of the Alien. An Introduction." *Phenomenology & Practice* 7 (2): 68–77.

Fuchs, Thomas. 2017. Ecology of the Brain. The phenomenology and biology of the embodied mind, Oxford, New York: Oxford University Press.

Gelhard, Andreas. 2011. *Kritik der Kompetenz*. Berlin: Diaphanes.

Haraway, Donna. 1991. "A Cyborg Manifesto." In *Simians, cyborgs, and women: the reinvention of nature*. New York: Routledge. 149–182.

Kaminski, Andreas, "Der Erfolg der Modellierung und das Ende der Modelle. Epistemische Opazität in der Computersimulation." In *Technik – Macht – Raum. Das Topologische Manifest im Kontext interdisziplinärer Studien*, ed. Andreas Brenneis, Wiesbaden 2018, pp. 317–33.

Lemoine, Blake. 2022. "Is LaMDA Sentient? — An Interview." Medium. Last modified June 11, 2022. https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917.

Levinas, Emmanuel. 1991. *Otherwise Than Being, or Beyond Essence*, translated by Alphonso Lingis. Dordrecht: Kluwer.

Levinas, Emmanuel. 1998. "The Ruin of Representation." In *Discovering Existence with Husserl: Studies in Phenomenology and Existential Philosophy*, translated and edited by Richard A. Cohen and Michael B. Smith, 111–21. Evanston: Northwestern University Press.

Liggieri, Kevin. 2022. "Anthropozentrismus in Mensch-Maschine-Interaktionen." In *Von Menschen und Maschinen: Mensch-Maschine-Interaktionen in digitalen Kulturen*, edited by Selin Gerlek et al., 73–93. Hagen: Hagen University Press. https://doi.org/10.57813/20220620-160005-0.

Lyotard, Jean- François. 1988. "Si l'on peut penser sans corps." In *L'inhumain. Causeries sur l'esprit*, 17–31. Paris: Galilée.
Merleau-Ponty, Maurice. 1962. *Phenomenology of Perception*, translated by Colin Smith. London and New York: Routledge.
Merleau-Ponty, Maurice. 1968. *The Visible and the Invisible*, translated by von Alphonso Lingis. Evanston, IL: Northwestern University Press.
Meyer-Drawe, Käte. 2007. *Menschen im Spiegel ihrer Maschinen*. Munich: Fink.
Müller, Oliver. 2022. "Maschinelle Alterität. Philosophische Perspektiven auf Begegnungen mit künstlicher Intelligenz." In *Begegnungen mit Künstlicher Intelligenz: Intersubjektivität, Technik Lebenswelt*, edited by Martin W. Schnell and Lukas Nehlsen, 23–47. Weilerswist: Velbrück.
Nassehi, Armin. 2019. *Muster: Theorie der digitalen Gesellschaft*. Munich: Beck.
Passig, Kathrin, and Scholz, Aleks. 2015. "Schlamm und Brei und Bits: Warum es Digitalisierung nicht gibt." *Merkur* 69(798): 75–81. https://www.merkur-zeitschrift.de/passig-scholz-schlamm-und-brei-und-bits/
Plessner, Helmuth. 2019. *Levels of Organic Life and the Human. An Introduction to Philosophical Anthropology*, translated by Millay Hyatt. New York: Fordham University Press.
Rieger, Stefan. 2019a. *Die Enden des Körpers: Versuch einer negativen Prothetik*. Wiesbaden: Springer.
Rieger, Stefan. 2019b. "Interface: Die Natur Der Schnittstelle." In *Mensch-Maschine-Interaktion, Handbuch zu Geschichte – Kultur – Ethik*, edited by Kevin Liggieri and Oliver Müller, 190–97. Stuttgart: Metzler.
Rinck, Monika. 2022. "Lyrik-Poetikvorlesung." Last modified February 24, 2022. https://www.youtube.com/watch?v=gHI-hNQGUBY&list=PLTjcusNZdoaOWlpUKypduXEbcUT0cEeHJ&index=2&t=1550s
Robbins, Tony. 2020. "Meet Sophia, World's First AI Humanoid Robot." Last modified January 17, 2020. https://www.youtube.com/watch?v=Sq36J9pNaEo
Stalder, Felix. 2016. *Kultur der Digitalität*. Berlin: Suhrkamp.
Waldenfels, Bernhard. 2022. *Globalität, Lokalität, Digitalität*. Berlin: Suhrkamp.
Wellner, Galit. 2014. "The Quasi-Face of the Cell Phone: Rethinking Alterity and Screens." *Human Studies* 37: 299–316.
Wiegerling, Klaus. 2016. "Leib und Lebenswelt im Zeitalter informatischer Vernetzung." In *Technisierte Lebenswelt: Über den Prozess der Figuration von Mensch und Technik*, edited by Marie-Hélène Adam, Szilvia Gellai and Julia Knifka, 139–56. Bielefeld: de Gruyter.
Wittgenstein, Ludwig. 1958. Philosophical Investigations, transl. G. Anscombe, London: Blackwell.
Zimmerli, Walther Ch. 2021. "Analog oder digital? Philosophieren nach dem Ende der Philosophie." In *Was ist Digitalität? Philosophische und pädagogische Perspektiven*, edited by Uta Hauck-Thum and Jörg Noller, 9–33. Berlin: Metzler.

3 Time-Consciousness and E-Memory

Arguing for a Phenomenological Revision of the "HEC" (Hypothesis of Extended Cognition) Paradigm

Federica Buongiorno

3.1 Introduction

> We appear to be living through an historical moment of rapid development and deployment of new memory technologies, based upon a linked series of innovations in networked digital technologies such as immensely cheap digital memory and massively powerful recording technologies such as mega-pixel cameras and the vast array of sensors and cheap memory included in electronic goods.
>
> (Clowes 2013, 107)

Ten years have already passed since this observation by Clowes and, as is evident to anyone, the situation characterizing technology-driven societies – at least in the western part of the globe – has gradually confirmed and radicalized this description, in an evolution that continues to produce new effects and changes in a very rapid and, at times, difficult to control manner. The advent of the Internet and, subsequently, cloud repositories, has unveiled the possibility of producing and storing enormous amounts of data, mostly passively generated, that now constitute "an ever-present background to our lives" (Clowes 2013, 108).

> It is safe to say that the Internet is currently the most ubiquitous cognitive technology and constitutes a significant part of our overall cognitive ecology […]. So the way we use and interact with digital information differs from the way we use and interact with analogue information.
>
> (Heersmink 2016, 391)

DOI: 10.4324/9781003312284-4

The study of the impact of digital technologies on human cognition, and specifically of the electronic memory revolution, becomes thus urgent: such a study must be approached with the critical awareness that much research, including experimental investigation, focuses on so-called WEIRD societies (Western, Educated, Industrialized, Rich, Democratic Societies), whereas the majority of Internet users are located in Asia; not to mention that in certain contexts, the use of the Web is limited or censored and is therefore conditioned by a substantial "digital divide" (Heersmink 2016, 395).[1] More generally, as Heersmink notes, "how the Internet transforms our memory and cognition depends on aspects that influence our cognitive profile like, for example, age, gender, level of education, and socio-cultural background" (Heersmink 2016, 395). It is therefore important to maintain critical caution about the generalizability of conclusions drawn from experimental studies and theories suggested in the cognitivist and philosophical fields, which are still mostly restricted to WEIRD societies.[2] Given these premises, there is no doubt that wherever digital technologies and the Internet are available, there is a tendency for co-evolution and mutual adaptation between human individuals and technologies. "We are becoming symbiotic with our computer tools, growing into interconnected systems that remember less by knowing information than by knowing where the information can be found" (Sparrow et al. 2011, 778). In particular, the development of E-memory has enabled the extension or incorporation of a number of cognitive processes, substantially transforming our conception of and relationship with lived temporality.

3.2 Framework: E-memory and the Hypothesis of Extended Cognition (HEC)

By "electronic memory" is meant both the physical component, device, or recording medium used to store data for retrieval on a temporary or permanent basis, such as personal computers or other digital electronic devices, and the digital information stored on these media. Electronic memories can be divided into two primary categories:

i volatile memory, which requires a constant power supply to retain the stored information: this type of memory loses written data as soon as the system is turned off; the retention of data is on the order of millesimal duration.
ii non-volatile memory, which can retain the stored information even when the electrical power supply has been turned off; data are stored for a long term.

Currently, *DRAMs* and *FLASH* memories constitute the majority of electronic memory devices. Regarding the type of information that can be stored on these media, there are essentially two types:

- information that is actively created and stored locally on computers (written text, images, audio and video files, etc.) or remotely on cloud repositories;
- information that is passively created and stored remotely (digital traces: derived, for example, from Google search, the use of Maps or other Internet-based services, etc. Cloud repositories are also based on physical storage that spans multiple servers.)

From the philosophical point of view, the study of E-memory is framed within the debate on two alternative epistemological frameworks: on the one hand, the Hypothesis of Extended Cognition (HEC) introduced by Clark and Chalmers in their famous 1998 article "The Extended Mind," (Clark and Chalmers 1998) and, on the other hand, the Hypothesis of Embedded Cognition (HEMC) proposed by Rupert (2004) as a weaker version of HEC. The HEC stresses embodiment and its implications, arguing that human cognition is not restricted to "skull and skin" but can incorporate, under certain conditions, elements of the external environment, which would thus be considered parts of the cognitive system. "Extended mind theory," Clark (2015, 3758) writes, "suggests that the physical machinery that realizes some of an individual agent's cognitive processes and mental states can, under humanly attainable conditions, include elements and devices located beyond the bounds of skin and skull." The HEMC, on the other hand, argues that some cognitive skills – primarily memory – can be externalized, that is, relocated to external media, without it being necessary to regard them as parts of the mind, so that mental activities remain confined to "skull and skin" (i.e., the brain). "Embedded views, […] argue that artifacts and other environmental scaffolding have important roles in human cognitive functioning but are never part of cognition" (Heersmink 2017, 23). This is equivalent to claiming that cognitive processes remain internally performed: on the contrary, Clark and Chalmers argue – by virtue of Parity Principle (PP) they introduce – that whenever functional equivalence is observed between external cognitive performance and internal cognitive performance, the external process does count as internal. In their words:

> If, as we confront some tasks, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as a part of the cognitive process, then that part of the world is […] part of the cognitive process.
>
> (Clark and Chalmers 1998, 14)

PP has been subjected to numerous endorsements as well as critiques: both the equivalence it postulates and, more specifically, the four conditions (availability, accessibility, trust, prior endorsement) the authors introduce to specify the principle in order to limit the extensibility of mind and avoid "cognitive bloat" have been strongly criticized (see Adams and Aizawa 2001; Dartnall 2005; Heersmink 2016; Kyselo and Walter 2009; Menary 2010; Rupert 2004, 106; Smart 2017). My purpose in this contribution is threefold, as I am interested in showing that:

1 the changes produced by E-memory can be better understood within the framework of HEC rather than the more minimalist framework of HEMC; this implies a preference for an embodied and ecological approach, which is compatible with a phenomenological interpretation.
2 (1) holds only provided that the PP is subjected to a certain criticism: the equivalence posited by the principle in Clark and Chalmer's formulation requires a certain phenomenological revision. More specifically, trust criterion and the automatic endorsement (which play a central role in Clark and Chalmer's PP) seem to posit specific difficulties, which I will critically address.
3 the embodied cognitive system is a *hybrid* of the organic and the artificial in which sensory-motor and cognitive performances are integrated and, in some cases, augmented by electronic and digital media: this requires a phenomenological understanding of the different layers of memory and time consciousness, which is lacking in HEC. Such an understanding implies to combine Husserl's theory of time consciousness, and his differentiation between primary, retentional memory and secondary, recollective memory, with Stiegler's notion of tertiary (technical) retention.

In developing my argument, I will investigate E-memory on the epistemological and phenomenological level: this means, I will refrain from an assessment of ethically sensitive issues such as whether E-memory helps or undermines subjective empowerment; whether it results in a substantial depletion of biological memory, or frees up instead cognitive energies for other mental performances.[3] Such evaluations, in fact, still lack sufficient supporting studies: although it is clear that many topics require greater attention paid to ethical and regulatory issues,[4] their analysis lies beyond the scope of this chapter.

3.3 Biological Memory, E-memory, and Conditions of Incorporability

Cognition can be defined as the processes an organism uses to organize information. This includes acquiring information (perception),

> selecting (attention), representing (understanding) and retaining (memory) information.
>
> (Bostrom and Sandberg 2019, 312)

Bostrom and Sandberg's can be taken as a general definition of cognition, in which memory is understood as one of the key components; the discussion of what specifically is "cognition," however, is open (see Heersmink 2017). Cognitivism in its internalist version understands cognition as a regulated manipulation of symbols that occurs internally, that is, in the brain: this approach "has recently been challenged by situated approaches, including embodied, enactive, embedded, and extended approaches" (Heersmink 2017, 19). It is clear that E-memory challenges internalism insofar, as it implies that "external hardware and software supports now routinely give humans beings effective cognitive abilities that in many respects far outstrip those of our biological brains" (Bostrom and Sandberg 2019, 312).[5] E-memory is structured according to the distinction, that also holds for organic memory, between *short-term memory systems* ("dealing with holding information over a few seconds"; Clowes 2013, 111) and *long-term memory systems*: we have in fact a *RAM*, or working/procedural memory, and a *fixed memory*, which is (more or less) extended or extensible. I suggest the following parallelism: in phenomenological terms, RAM and fixed memory would correspond, respectively, to primary-retentional memory (in direct continuity with the perceptual present/input) and to secondary-recollective memory (which stands in discontinuity with the perceptual present and rather *re*produces primary memory). The most significant impact of electronic memory is on long-term (fixed) memory, particularly autobiographical and episodic memory (think of the type and amount of personal information shared through social networks), as well as semantic memory,[6] "which store information about meaning and factual knowledge" (Clowes 2013, 111). According to Clowes, four dimensions of E-memory can be identified (*totality, practical cognitive incorporability, autonomy, entanglement*), which "can be argued to offer novel properties with respect to previous regimes of memory technology" (Clowes 2013, 111): these dimensions entertain, as we shall see, different relationships with the four specifications of Clark and Chalmer's PP mentioned above (availability, accessibility, trust, prior endorsement). I will focus on the first two requirements introduced by Clowes for E-memory (totality and practical cognitive incorporability), since they seem to have a special phenomenological significance and to specifically challenge Clark and Chalmer's criterion of trust, which is pivotal to their formulation of the PP.

3.3.1 Totality

Clowes begins by analyzing the condition of totality. "E-Memory promises to record our everyday activities on a scale and with a fidelity and completeness that would have been practically unimaginable under previous regimes" (Clowes 2013, 111). This aspiration to totality particularly characterizes "lifelogging" projects. This term refers to "a form of pervasive computing which utilizes software and sensors to generate a permanent, private and unified multimedia record of the totality of an individual's life-experience and makes it available in a secure and pervasive manner."[7] The forerunner of this project is the "MyLifeBeats" experiment conducted by Gordon Bell back in the 1990s and consisting – as noted in the manifesto published by Bell and Jim Gemmell in 2009 (Bell and Gemmel 2009) – of a "complete digital record" of life, that is, of the collection and multimedia archiving on special databases of all daily activities performed by Bell over the years. The goal was to to overcome the drawbacks of organic memory, which by definition is unreliable, thus not objective, and limited in capacity: electronic memory would instead allow a stable and permanently accessible storage of all memories, digitally documented by individuals.

Some objections can be made against this argument:

i Although Bell and Gemmell put great effort into demonstrating the practical preferability of a complete replacement of organic memory by electronic memory, it is not clear where the specific advantage of a complete storage of memories would lie, given that, realistically, only some contents will be recalled (depending on practical needs): thus, the criteria by which the salience of contents is established (hierarchically) and by which their retrieval is implemented seem more relevant than completeness.

ii If we follow the framework of HEC, that is, if we observe an equivalence (granted by PP) between organic and electronic memory, then electronic memory should function the same way as organic memory: since this latter works selectively (only a part of memories are stored and retrievable at a later time), then we have no reason to believe that electronic memory should not work selectively too.[8] The point is that Total Recall does not advocate the *integration* of the two memories but the *replacement* of one (organic) by the other (electronic). Such a replacement, however, is unrealistic under the PP, which assumes that an integration, and not a substitution, is established between the two types of memory. So, one or the other: if we opt for complete substitution and the ideal of total recall, we will have to give up HEC; if we opt for

a paradigm of embodied cognition and selective storage, on the other hand, we will have to rule out the idea of total recall.

iii Not every stored and retrievable piece of data is, as such, necessarily a memory. Myin and Zahidi (2015) recall some interesting examples showing how archived and recalled memories, even digitally, may not "be recognized" by the individual as their own memories. Only if the recalled memory possesses a certain salience in the context of individuals' personal and sedimented lives, that is, if it connects to other memories of the same individuals, does it acquire the meaning of a recollection, a memory referring to a past aspect or event of their lives: "having access to information sometimes does and sometimes doesn't give rise to genuine remembering. Access to information is not a sufficient condition for remembering" (Myin and Zahidi 2015, 187). It follows that archiving and potential accessibility, even when extended to the totality of events in our lives, does not coincide with "total recall."

It is reasonable to conclude, with Clowes (2013, 113), that "exactly how 'the totality' of E-Memory traces captured by lifelogging might really be brought back and indeed to what extent it is desirable or useful to aspire toward total recollection is still an open question." From the phenomenological point of view, the analysis can be developed further. In describing the structure of secondary memory (recollection),[9] Husserl states that this takes over once the primary (impressional-retentional) temporal process has ended: secondary memory or recollection is a free act, an act of our "I can." Recollection is *secondary* because, unlike primary retention, it is not linked to perception by a law of necessary association. It could be produced or not, once the perception is over: in phenomenological terms, then, the emphasis shifts from totality to potentiality.[10]

In other words: there is a difference between archive and memory. Memory presupposes archiving (in order for me to remember something, the content must be deposited in consciousness), and yet archive alone is not memory. The phenomenological *archive* consists of the set of passively sedimented sensory-motor patterns and habits, which are socially and culturally conditioned.[11] These memory contents are "not directly lived by individuals" – to use Stiegler's expression: they are the trace of tradition and heritage, the footprint of social conditions and unaware sensorimotor schemes. As such, they are impersonal (Merleau-Ponty 2010, 2012), more-than-organic, extra-mental, and inherently techno-organic or *hybrid*. In order for individual (personal) memory to work on this social (impersonal) archive, it is necessary to refer to the specific

(situated) embodiment of the individual: human memory (may it be digitally augmented or not) lives of a dialectic between the organic and the inorganic, which rules out the possibility of total recall and complete replacement by e-memory systems.

3.3.2 Practical Cognitive Incorporability

According to Clowes (2013, 114), the condition of "practical cognitive incorporability" typical of electronic memory "is supposed to reflect the way in which E-Memory technologies (and other cognitive technologies) are increasingly embedded in or accessed through everyday objects, such as mobile phones, that we can learn to interact with in an almost thoughtless way."[12] E-memory and especially the Internet, according to Clowes, are characterized by a high embeddability that enable them to meet two of the criteria (those of availability and accessibility) posed by the PP in Clark and Chalmers' formulation; the other two criteria of the PP (those of trust and prior endorsement) seem, on the other hand, more problematic, because they are challenged by the two additional characteristics of E-memory highlighted by Clowes (autonomy and entanglement):

> The problem with E-Memory technologies is related to their autonomy, and especially the way in which the computational systems which underlie them are subject to change which can be invisible to users […] So even if we leave aside deliberate memory tampering, factors of autonomy and entanglement may operate as a fundamental constraint on what should really count as part of our minds […] Trust relations appear to operate as a fundamental limit on real (concrete) incorporability.
>
> (Clowes 2013, 124)

In any case, Clowes maintains the PP by arguing that it admits different "densities" of embedding (higher or lower), depending on which features of the electronic memory we consider. There is no doubt that HEC has an advantage over other explanatory models: in particular, HEC

> is interested in the wider ensemble of human + that is, the agent plus the motley collection of equipment which has been densely incorporated […] What is going on in the brain is not the only factor of importance; we should care about the profile of the agent as a whole.
>
> (Clowes 2013, 128)

As Menary (2010, 268) observes, "the payoff from extended mind style arguments is the integration of the bodily 'internal' and 'external' aspects of cognition into a whole." More precisely,

> the upshot of extended mind style arguments leads us to understand cognition (and the mind) as hybrid – involving both internal and external processes – and integrated – the bodily internal and external processes coordinate with one another in the completion of cognitive tasks.
>
> (Menary 2010, 268)

This is also an undoubted advantage on the phenomenological level, because it implies taking as the starting point of analysis the whole complex "subject + their environment" (*Umgebung*), that is, the embodied situation of the subject. Menary's observation that "our initial cognitive engagements with the world are embodied engagements, they are primarily sensorimotor engagements" (Menary 2010, 269) has full phenomenological relevance. Unlike Clowes, however, I believe that the incorporability hypothesis conveyed by HEC is preferable *not by virtue* of PP, but on condition of its *phenomenological critique*.

3.4 A Phenomenological Revision of the Parity Principle

It is interesting, from a phenomenological perspective, the observation made by Myriam Kyselo in her review of Andy Clark's *Supersizing the Mind* (2008): one reason why the hypothesis of extended cognition is preferable to HEMC "would thus be phenomenological: if the coupling between an agent and an external tool becomes intimate enough, the agent may no longer feel as though she is using a tool, but instead experience the (former) tool as part of herself" (Kyselo and Walter 2009, 806). However, Clark himself dismisses the phenomenological argument by assuming that: "I do not mean, here or elsewhere, to advance any arguments of the form 'it seems to us as if we are/are not cognitively extended; therefore we are/are not cognitively extended!" (Clark 2008, 238). In Clark's view, therefore, the meaning of the equivalence posited by PP should not be understood in a phenomenological sense.

As already mentioned, equivalence is based on PP, which Clark (2010) specifies by distinguishing four conditions:

i *Availability*: The resource be reliably available and typically invoked.
ii *Accessibility*: The information is easily accessible (directly available).
iii *Trust*: The information retrieved is "more or less" automatically endorsed. "It should not usually be subject to critical scrutiny (unlike the

opinions of other people, for example). It should be deemed about as trustworthy as something retrieved clearly from biological memory" (Clark 2010, 46).

iv *Prior Endorsemen*t: The information a cognitive technology presents has been consciously endorsed at some point in the past and is there as a consequence of this endorsement.

It should be noted that all four conditions assume that information is already given (i.e., already stored on some medium, internal or external, and therefore retrievable): the HEC does not seem interested in explaining how memories *are formed*, i.e., the genesis of the consciousness of the past. Instead, this is precisely the question at the heart of Husserl's lectures *On the Phenomenology of the Consciousness of Internal Time (1893-1917)*. In my view, it is this phenomenological oversight that makes PP indifferent to the distinction, which I evoked above, between stored (archived) information and properly remembered information, and that exposes PP to much of the criticism that has been advanced. This circumstance is recognized by Kyselo and Walter (2009, 806):

> In our eyes, this [Clark's dismissal of the phenomenological argument] is unfortunate since first person phenomenology certainly is essential for work in embodied, embedded and situated cognitive science. A task for future research will be to disentangle first person arguments and third person arguments for HEC and to assess their quality independently (although [it] is debatable whether one can make a convincing phenomenological argument for HEC).

It is the Trust Criterion that poses the most significant difficulties. Clark assumes that what is "clearly retrieved from biological memory" is not subject to "critical scrutiny," and is therefore normally "trustworthy": to count as part of the extended cognitive system, even externally retrieved information must be "automatically endorsed." This assumption is problematic on the phenomenological level: only one layer of memory – the one corresponding to what I have called the "archive," i.e., primary memories (retentions) and their progressive sedimentation into sensory-motor patterns (habits) – can be said to be "automatically endorsed," i.e., endorsed without reflection or awareness, whereas the layer corresponding to contents reproduced in secondary memory/recollection are, as we have seen, delivered by a free act of consciousness in which the role played by reflection as a *sui generis* perceptual act is pivotal. This makes it possible to render the original (impressional-retentional) consciousness of time thematic (i.e., objective) and to reproduce it in the present: thus, while we can admit that primary consciousness is passively produced without

direct ego-activity, the same cannot be said of secondary (reproductive) consciousness. Similarily, we shall assume that electronic operating memory (RAM) in its implicit functioning is "automatically endorsed"; semantic and autobiographical memory, on the other hand, is explicit and as such would fall within the realm of phenomenological recollection. Clark, aware of the difficulty here, points out the following:

> This is not to say that biological memory can never turn up as *such an object*. Bio-feedback devices sometimes make our inner activity into an object of our own *attention*. But on the whole, that which forms part of the agent is not encountered and processed like other stuff in the world. Rather, it provides the means and mechanisms for our encounters with the wider world.
>
> (Clark 2015, 3762; my emphasis)

Here, Clark seems to reintroduce an argument in the first-person perspective, which he has previously ruled out in principle: what constitutes part of the agent (e.g., biological memory) is not encountered "like other stuff in the world" – this is entirely true from a phenomenological point of view, and indeed, a real dilemma seems to be emerging here. It is Clark himself who formulates it with reference to Otto's well-known thought experiment proposed in the 1998 article:

> […] it should make no difference at all whether or not Otto is now, or ever was, *aware* of the source of the reliability of the notebook involving process. Indeed – and here comes the promised dilemma – there is a very real sense in which the more he is aware of such matters, the less the notebook will seem to be playing the same kind of functional role as biological memory. For as we noted, our *biological memory is not typically subject to agentive scrutiny as a process at all*, much less as one that may or may not be reasonably judged to be reliable by the agent.
>
> (Clark 2015, 3763; my emphasis)

The more transparent the external medium (the source) is, the more it will be considered integrable by the subject: if transparency decreases, that is, if the subject critically reflects on the source, which is not normally subjected to "agentive scrutiny," the external medium will be less integrable. But, as we have seen, this assumption is phenomenologically problematic: only one level of biological memory "is not typically subject to agentive scrutiny as a process at all" – a level that is certainly very important, especially in a Merleau-Pontian perspective aimed at valuing the unconscious, habitual, and schematic aspects of lived experience.[13]

There remains, however, a phenomenological difference between sedimented (stored) memories and "remembered" (reproduced) memories, which passes precisely through reflection (i.e., processes of attention and representation): only by admitting this distinction (and not by reducing it!) *in both* biological and electronic memory, it becomes possible to establish an equivalence between them and to explain why some retrieved information is also recognized (endorsed) as one's own memories, whereas others are not. I believe this argument is confirmed by an analysis of the famous Otto-example advanced by Clark and Chalmers (1998): the notebook Otto resorts to in order to supplement his memory impaired by Alzheimer's should function and be endorsed, according to PP, in the same way that a healthy subject's internally processed memory would. But what assures us, given that Otto has Alzheimer's, that he will recognize the notes in the notebook as *his* notes? Or that he recognizes the notebook itself as *his* own (see Myin and Zahidi 2014)? Or even that he will *always* remember (in compliance with the criterion of availability) to take it with him? If it were true that both biological and electronic memories are generally and usually "automatically endorsed," then none of these questions should arise. The dilemma consists in the following: the more Otto "knows," the less he is "extended." But this is true only if we admit that *knowing* is generally non-constitutive of our consciousness of the past: which, as it turns out, is not phenomenologically correct, since secondary memory processes imply reflection. I conclude from this that PP in Clark and Chalmers' classical formulation does not guarantee equivalence between biological and electronic memory: it is not by reducing or ignoring the difference between archive (primary memory) and recollection (secondary memory) that the equivalence can be established, but only by acknowledging it as constitutive for both types of memory.

3.5 Additional Argument: Parity Principle and the Internet

Smart has noted that if we consider the case of the Internet, the Trust Criterion lends itself to further criticism, as information seems to be less reliable and more open to tampering. As a matter of fact, the information is almost invariably created by other users or by remote computers; secondly, precisely insofar as much of the information is produced by other agents, "misinformation and mendacity are ever-present hazards. In this situation, the epistemic consequences of automatic endorsement are, to say the least, worrying" (Smart 2017, 281). This aspect is also highlighted by Clowes: E-memory systems "[…] are too open to what is effectively memory tampering. This implies important constraints on the concrete embeddability of E-Memory systems: Trust relations appear to operate as a fundamental limit on real (concrete) embeddability" (Clowes 2013, 126).

This is due to the high autonomy of digital systems and the opacity of their functioning, which goes hand in hand with transparency in their use. The difficulty could be reduced, according to Smart, by emphasizing – as done by Clark – the role of sub-personal (unconscious) filter mechanisms: these mechanisms, which are essentially internal and would not be altered by technological developments, would ensure that even the contents provided by electronic memory are subjected to a sub-personal (nonconscious) form of proof. This implies admitting that "the abilities that seem to determine whether or not an agent believes the truth are ones that remain in the head of the agent" (Smart 2017, 300): this, however, is a major concession to the internalist paradigm from which HEC would intend to depart in principle and should not be admitted uncritically.[14]

It is particularly important to avert the risk of viewing endorsement as a kind of label that, "comparable to a quality label like 'made in Switzerland,' gets attached to the memory trace when it is formed, and it accompanies the memory in the recall phase" (Myin and Zahidi 2014, 189). Again, it seems that this risk is caused by the lack of phenomenological attention to the genesis of information and the related consciousness of the past. This point is sharply noted by Myin and Zahidi by referring to Wittgenstein:

> Wittgenstein rejects that there is something like "[…] some feature of our memory image that tells us the time to which it belongs." […] Elsewhere he writes: "[It] takes remembering to tell us that this is the past. But if memory shows us the past, how does it show us that this is the past? It does *not* show us the past. Any more than our senses show us the present."
>
> (Myin and Zahidi 2014, 189)

In order to explain how memory provides us with content that we can say and recognize as *past*, the phenomenological specification we have already introduced is necessary: while primary-retentional consciousness is functional to the unfolding of perception as an act that is not punctual but itself "extended" as a certain duration, authentic consciousness of past content is given by secondary memory. In § 27 of the lectures on the internal consciousness of time, Husserl clarifies this point with the famous example of the "illuminated theater": if we say that we remember the illuminated theater to which we went some time ago, we express linguistically the specific experience whereby "[…] 'in my interior' I see the illuminated theater as having been. In the now I see the not-now" (Husserl 1991, 60). As I have noted elsewhere, "this is exactly the essence of recollection as a thetic presentification: I proceed as if I were performing again the original perception of the illuminated

theater" (Buongiorno 2018, 109). Recollection, then, presents us with a past content in the original, that is, it presents us in the now with a not-now. This is possible only by bringing reflection into play, i.e., by resorting to that objectification that Clark sought to limit in the passage quoted above: recollection as re-production implies objectifying and thematizing a previous act of perceptual consciousness. Between recollection and perception there is, therefore, a "jump" – in the sense that recollection is free and implies the elapse of the primary perceptual process. At the same time, the two types of memory entertain a relationship of implication: if there had been no primary memory, no secondary memory would be possible.

From the perspective of a phenomenological understanding of incorporability on the basis of the suggested revision of PP, it is even possible to go a step further by making use of the notion of "tertiary retention" proposed by Bernard Stiegler: the fact that external memory and E-memory systems are incorporable in extended cognition does not only derive from their possible equivalence with biological memory, but also from the fact that they belong to a temporal dimension that, although called "tertiary" to distinguish it from the two dimensions (primary and secondary) already identified by Husserl, actually precedes the other two. "The technical, tertiary form of memory constitutes an imprint of memory that is structurally prior to even our primary retention" (Bluemink 2020). As Stiegler states in *For a New Critique of Political Economy*,

> tertiary retention always already precedes the constitution of primary and secondary retention. A newborn child arrives into a world in which tertiary retention both precedes and awaits it, and which, precisely, constitutes this world *as* world.
>
> (Bluemink 2020)

The notion of tertiary retention echoes and deepens, in a sense, the inherently non-technical, always-already cultural, and social nature that characterizes biological memory itself, which I recalled in the Introduction. In *Technics and Time 3*, Stiegler writes:

> The programming industries, and more specifically the mediatic industry of radio-televisual information, mass-produce temporal objects heard or seen simultaneously by millions, and sometimes by tens, hundreds, even thousands of millions of 'consciousnesses': this massive temporal co-incidence orders the event's new structure to which new forms of consciousness and collective unconsciousness correspond.
>
> (Stiegler 2011, 1)

Stiegler's point of view is relevant for it allows us to focus on a material aspect that normally finds no place within the HEC debate:

> These external memory supports (i.e. exteriorisations of consciousness) provide a new collectivisation of consciousness (the *We*) that exists beyond the realm of the individual (the *I*). The systematic influence of mnemotechnics on individual human consciousness must therefore be looked at as political in itself; if we are to understand technics we must understand how the technical milieu of human consciousness affects us collectively, as a '*we*', thus affecting us politically.
>
> (Bluemink 2015)

In particular, the 20th and 21st centuries are characterized by a process of "epiphylogenetic" adaptation and "coupling" with technical media, which – as I anticipated in the Introduction – has led to the claim that: "We are becoming symbiotic with our computer tools, growing into interconnected systems that remember less by knowing information than by knowing where the information can be found" (Sparrow et al. 2011, 778). The case of the so-called "digital natives" ("Generation Z") seems to confirm Stiegler's hypothesis: for those who have access to the Internet and portable technologies from birth, in fact, analog mnemotechnics – not hybridized with E-memory – plays an incomparably less extensive and less decisive role than they do for previous generations. I will not delve here into Stiegler's materialistic perspective: however, it is important to highlight that the reference to tertiary retention and thus, ultimately, to the capitalist regime of production, storage, and fruition of (E-)memories introduces an additional critical dimension that should not be overlooked; instead, it should be integrated into a phenomenological revision of the HEC's understanding of memory.

3.6 Conclusions

The phenomenological approach allows us to understand the origin of the consciousness of the past and to differentiate between automatic, primary memory (archive) and memory in the proper sense (secondary memory or recollection): the possibility of incorporating E-memory into the extended cognitive system – hybrid of organic and artificial memory – can be framed within the PP, provided we integrate it phenomenologically. On the procedural level, in fact, we will be inclined to integrate information from E-memory automatically, as is the case with biological memory of primary contents; on the explicit level (autobiographical and semantic memory), on the contrary, we will have

to subject electronic memories to some explicit scrutiny, unlike Clark's assumption, because the information provided is not always reliable or can be recognized as "our" memory. The case of the Internet is emblematic: according to Clark, the high risk of manipulation of information on the Web makes the Internet a bad candidate for incorporation. In the meantime, however, the Internet has become the most ubiquitous and transparent data production and storage system and should therefore lend itself well to incorporation according to PP's criteria. It is possible to break out of the impasse by resuming the phenomenological argument. It is not automatic endorsement that guarantees equivalence with biological memory and thus incorporability: rather, these are guaranteed by the first-person awareness of secondary (recollective) memory processes.

To sum up: classical HEC assumes that:

i the incorporation of external memory systems into the extended cognitive system is based on their automatic endorsement (transparency);
ii the automatic endorsement to which externally stored information is subjected is analogous to that to which information deposited in organic memory is subjected;
iii this endorsement is by default automatic and not reflexive (unaware): if it were not, an element of "critical scrutiny" (read: reflection) would come into play, making the endorsement of electronic memory no longer transparent, but opaque, which would reduce the chances for incorporation.

I have proposed a phenomenological revision of HEC, by pointing out that:

a the absence of critical scrutiny and thus automatic endorsement is the condition of *one* layer of biological memory (the primary, sensory-motor, and habitual level), whereas reflection is called into play by recollection (secondary memory);
b the equivalence by automatic endorsement between biological and electronic memory holds only for one layer of memory (biological archive/RAM or operational memory); what about the equivalence in the context of secondary memory (recollection/fixed external memory)?;
c in this latter case (secondary memory), endorsement is based on explicit processes of critical scrutiny (reflection): recollection requires a reproduction in the present of a past, already perceived content and this act of reproduction, as assumed by Husserl, is the result of a free will act. Thus, it requires awareness in the form of reflection and explicit thematization.

Therefore, my phenomenological proposal for revision agrees with Myin and Zahidi's conclusion:

> [We must] consider remembering to be a consequence of a number of not precisely datable converging processes and tendencies, none of which individually constitutes such a discrete event, but is only a necessary, but not sufficient, condition. Already existing habits, exogenously induced or endogenously noticed salience, either explicit or implicit, are not discrete features, but related to attitudes of a person, without precise temporal boundaries, but nevertheless, as was illustrated earlier, necessary conditions for later remembering.
>
> (Myin and Zahidi 2014, 190)

Notes

1 Heersmink observes that the well-known experimental study by Sparrow et al. (2011) – which "explores whether having online access to search engines, databases, and the like, has become a primary transactive memory source in itself" (Sparrow et al. 2011, 776) – was conducted in a highly controlled laboratory on a small sample of Western educated individuals, so that doubts remain about the generalizability of the results obtained.

2 Luciano Floridi has pointed out that in high-tech countries, most of the Gross Domestic Product (GDP) is derived from services, products, and activities based on information – that is, data flows: in all G7 member countries, (Canada, France, Germany, Italy, Japan, the United Kingdom, the United States) at least 70 percent of GDP is based on Information Technologies (IT). Yet, this observation is focused on a narrow geo-political context: the accessibility and spread of digital information and the Internet in non-European and non-Western countries is more complex and in need of specific investigation. See Floridi (2014).

3 Sparrow's study reaches, due to the limited sample surveyed, rather open-ended conclusions in this regard: the integration of biological and electronic memory seems to produce a tendency to remember "where" data are stored rather than their content, but this is not necessarily bad *per se,* or to a greater extent than other technological innovations in human cultural history. Loh and Kanai (2016) note that "relying on technology as an external memory source can result in reduced learning efforts as information can be easily retrieved later. This is not entirely maladaptive as we can strategically free up additional cognitive resources for other prioritized operations" (cited in Heersmink 2016, 396). What we can conclude, at the trend level, is that "[...] processes of human memory are adapting to the advent of new computing and communication technology" (Sparrow et al. 2011, 778).

4 For example, Heersmink (2017, 17-18) notes that: "Situated cognition theory, in contrast, has studied cognitive artifacts extensively (...), but has not even addressed their moral aspects in a substantive way."

5 In this sense, it is possible to speak of "cognitive enhancement" of biological memory: "A cognitively enhanced person […] is somebody who has benefited

from an intervention that improves the performance of some cognitive subsystem without correcting some specific, identifiable pathology or dysfunction of that subsystem" (Bostrom and Sandberg 2019, 312). By "cognitive enhancement," it is meant that "the amplification or extension of core capacities of the mind through improvement or augmentation of internal or external information processing systems" (Bostrom and Sandberg 2019, 311). To what extent, however, E-memory realizes an "enhancement" of already given (biological) abilities or enables the emergence of new abilities remains difficult to determine (see Clowes 2013).

6 As Heersmink (2016, 397) observes, "it is largely semantic memory with which the Google effect is concerned […] [but] the Internet is increasingly influencing and transforming episodic memory." The "Google effect" refers to the ubiquity of the Internet as a potentially always-accessible information production and storage system, with the result that "when we know information is easily available in some external media, we put less effort into storing that information in biological memory" (Heersmink 2016, 389).

7 http://research.nii.ac.jp/ntcir/ntcir-13/index.html.

8 "Biological memories are selective – what is retained is only what is relevant at the time. They are also prone to disruption and decay […] In comparison to biological memories, digital records are both complete and permanent, and the digital shadow therefore has these same characteristics" (Burkell 2016, 18). "Physical records and biological memories are subject to selective deposit and selective survival": what is stored (and thus what is available for review/recall/interpretation) "is only what matters *at the time*" (Burkell 2016, 18).

9 As I explained above, secondary memory or recollection would be the equivalent of digital long-term memory systems (the most affected by the E-memory revolution).

10 Incidentally, note that this shift constitutes a substantial difference from the approach to the problem of memory by an author such as Henri Bergson – whom I consider, in other ways, in continuity with phenomenology (see Buongiorno 2018). In *Matter and Memory*, Bergson (2007) also assumes a double characterization of memory as, on the one hand, automatic and superficial (in direct continuity with perception, resembling Husserlian retention) and, on the other hand, deep and spontaneous (resembling Husserlian recollection). In the latter layer, the totality of life consciousness is deposited in the form of pure memories: although only a part of these is recalled by action (in the present) and embodied in memories-images useful for acting effectively in the present, nothing is lost and everything is recorded in the depth of consciousness. The difference from Husserl consists in the different connection of deep memory with freedom: according to Bergson, we are not entirely free to choose which memories to reactivate, because such a reactivation is always to some extent guided by present action and the need to respond to it effectively. Husserl assumes, on the contrary, that recollection is the result of a choice and is not guided by the present, precisely because it unfolds once perception has ended (which, of course, implies that every secondary memory refers to a content that must have occurred at least once before).

11 Memory as tradition and sedimentation is explored by Husserl in his late writings on passive synthesis and further analyzed by Maurice Merleau-Ponty in his analysis of perception, institution, and passivity.

12 The more a device is incorporated, the more it becomes transparent in use: note that transparency in use goes hand in hand with opacity in functioning (the functioning protocol of most digital technological tools is unknown to the user.).
13 I do agree with Clark when he claims that "non-conscious goals, desires, motivation, and strategies, of surprising scope, sophistication, and flexibility, are ubiquitous in both human and (other) animal cognition" (Clark 2015, 3769). It is for sure true that these contents undergo "unconscious meta-cognitive mechanisms" of proof and check (Clark 2015, 3769); however, all this still does not help us answer the question raised (why and how do we recognize certain recalled information as "ours," as our memories, i.e., as our past – and not others).
14 In Clark and Chalmers's version, the HEC implicitly assumes internalism as the standard position within the debate on (E-)memory: this is also highlighted by Myin and Zahidi, who claim that Clark and Chalmers "remain wedded to the idea that memory, and more generally cognition, is a matter of information processing," so that "remembering involves the manipulation of these information carriers." (Myin and Zahidi 2014, 186)

References

Adams, Fred, and Ken Aizawa. 2001. "The Bounds of Cognition." *Philosophical Psychology* 14 (1): 43–64.
Bell, Gordon C., and Jim Gemmell. 2009. *Total Recall: How the E-Memory Revolution Will Change Everything*. New York: E. P. Dutton.
Bergson, Henri. 2007. *Matter and Memory*, trans. N. Margaret Paul, and W. Scott Palmer. New York: Cosimo Classics.
Bluemink, Matt. 2015. "Stiegler's Memory, 2: The Industrialisation of Consciousness." Blue Labyrinths. Last modified June 4, 2015. https://bluelabyrinths.com/2015/06/04/stieglers-memory-2-the-indstrialisation-of-consciousness/.
Bluemink, Matt. 2020. "Stiegler's Memory: Tertiary Retention and Temporal Objects." 3:AM Magazine. Last modified January 3, 2020. https://www.3ammagazine.com/3am/stieglers-memory-tertiary-retention-and-temporal-objects/.
Bostrom, Nick, and Anders Sandberg. 2019. "Cognitive Enhancement: Methods, Ethics, Regulatory Challenges." *Science and Engineering Ethics* 15 (3): 311–41.
Buongiorno, Federica. 2018. *La Linea Del Tempo: Coscienza, Percezione, memoria Tra Bergson e Husserl*. Rome: Inschibboleth.
Burkell, Jacquelyn Ann. 2016. "Remembering Me: Big Data, Individual Identity, and the Psychological Necessity of Forgetting." *Ethics and Information Technology* 18 (1): 17–23.
Clark, Andy. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York: Oxford Academic.
Clark, Andy. 2010. "*Memento*'s Revenge: The Extended Mind, Extended." In *The Extended Mind*, edited by Richard Menary, 43–66. Cambridge/MA: MIT Press.
Clark, Andy. 2015. "What 'Extend Me' Knows." *Synthese* 192 (11): 3757–75. https://core.ac.uk/download/pdf/201000613.pdf.
Clark, Andy, and David Chalmers. 1998. "The Extended Mind." *Analysis* 58 (1): 7–19.
Clowes, Robert W. 2013. "The Cognitive Integration of E-Memory." *Review of Philosophy and Psychology* 4: 107–33.

Dartnall, Terry. 2005. "Does the World Leak Into the Mind? Active Externalism, 'Internalism' and Epistemology." *Cognitive Science* 29 (1): 135–43.
Floridi, Luciano. 2014. *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality*. Oxford: Oxford University Press.
Heersmink, Richard. 2016. "The Internet, Cognitive Enhancement, and the Values of Cognition." *Minds and Machines* 26: 389–407.
Heersmink, Richard. 2017. "Extended Mind and Cognitive Enhancement: Moral Aspects of Cognitive Artifacts." *Phenomenology and the Cognitive Science* 16: 17–32.
Husserl, Edmund. 1991. *On the Phenomenology of the Consciousness of Internal Time (1893-1917)*, trans. John Barnett Brough. Dordrecht: Springer.
Kyselo, Miriam, and Sven Walter. 2009. "Review of: Andy Clark's *Supersizing the Mind*." *Philosophical Psychology* 22 (6): 803–07.
Loh, Kep Kee, and Ryota Kanai. 2016. "How Has the Internet Reshaped Human Cognition?." *Neuroscientist* 22 (5): 506–20.
Menary, Richard. 2010. "Cognitive Integration and the Extended Mind." In *The Extended Mind*, edited by Richard Menary, 226–43. Cambridge, MA: MIT Press.
Merleau-Ponty, Maurice. 2010. *Institution and Passivity: Course Notes from the Collège de France*, translated by Leonard Lawlor and Heath Massey. Evanston, IL: Northwestern University Press.
Merleau-Ponty, Maurice. 2012. *Phenomenology of Perception*, translated by Donald A. Landes. London: Routledge.
Myin, Erik, and Karim Zahidi. 2015. "The Extent of Memory. From Extended to Extensive Mind." In *Mind, Language and Action: Proceedings of the 36th International Wittgenstein Symposium*, edited by Danièle Moyal-Sharrock, Volker Munz, and Annalisa Coliva, 391–408. Berlin, Munich, Boston: De Gruyter. https://www.academia.edu/5891301/The_Extent_of_Memory_From_Extended_to_Extensive_Mind_2014_.
Rupert, Robert D. 2004. "Challenges to the Hypothesis of Extended Cognition." *The Journal of Philosophy* 8 (Vol. 101): 389–428.
Smart, Paul. 2017. "Extended Cognition and the Internet: A Review of Current Issues and Controversies." *Philosophy and Technology* 39 (3): 357–90.
Sparrow, Betsy, Jenny Liu and Daniel M. Wegner. 2011. "Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips." *Science* 333, 776–78.
Stiegler, Bernard. 2011. *Technics and Time 3: Cinematic Time and the Question of Malaise*, trans. Stephen Barker. Stanford, CA: Stanford University Press.

4 The Vanishing Point

Digital Technologies and the Quest for a Phenomenology of Technological Invisibility

Marco Cavallaro

4.1 Introduction

In their 1999 movie *The Matrix*, the Wachowski brothers present a world where technology has so thoroughly dominated humanity that people remain unaware of their own enslavement. This aspect of the film sparked widespread debates about its philosophical implications. Early in the 21st century, a surge of scholarly work appeared, with titles such as *Matrix and Philosophy*, *More Matrix and Philosophy*, and *The Philosophy Behind the Matrix*, populating bookstore shelves and university libraries. Much of the philosophical enthusiasm surrounding the movie centered on the revival of the classical skeptical dilemma: Can we be certain of our knowledge of the external world? In *The Matrix*, technology (referred to as "the machines") creates a fictional reality for humans, who remain oblivious to their enslavement and accept this simulated world as genuine.

However, in my view, this revival of the Cartesian doubt has overlooked a more profound issue raised by the film – one that is of greater relevance to contemporary philosophy of technology. The narrative of *The Matrix* offers an incomplete explanation of how humanity became entirely ignorant of its enslavement. The story tells us that a war between humans and intelligent machines resulted in the machines capturing humans and using their bioelectric energy, while their minds were pacified by the Matrix. Yet, the question remains: How did technology reach the point where it could vanish completely from the visible world, leaving behind a reality constructed entirely by machines, but where technology itself seems absent? Viewed from this angle, the world depicted in *The Matrix* bears striking resemblances to our digitized 21st-century world. Digital technologies, to paraphrase Heraclitus, "love to conceal themselves." Far from being disconnected from reality, they form the invisible infrastructure of post-industrial societies worldwide – but often, we are unaware of them and they escape our sight.[1]

DOI: 10.4324/9781003312284-5

This chapter examines the many facets of digital technology's invisibility. How, and in what ways, can technology – especially digital technology – be said to be invisible to us? Before addressing this question, Section 4.2 undertakes a methodological discussion on the aims and objectives of a phenomenology of digital technology. I argue for considering the technological phenomenon as the appropriate subject for such a phenomenological approach, and I examine the concept of the "digital" in contrast to the "virtual." Section 4.3 presents a typology of technological invisibility, distinguishing between perceptual, functional, purposive, and environmental forms of invisibility. Finally, Section 4.4 applies these distinctions to digital technology, offering a phenomenological analysis of how it hides itself in the contemporary world.

4.2 Phenomenology of (Digital) Technology: Again or Finally?

4.2.1 Phenomenology's Complicated Relation to Technology

Phenomenology's relationship with technology has been and still is complicated. This situation arises from both historical and structural factors.

Historically, the founder of the phenomenological movement, Edmund Husserl, devoted relatively little attention to technology in his work. Despite living in a world increasingly shaped by technological advancements, Husserl rarely applied his phenomenological method to the study of technological objects. When he did, he primarily viewed technology as one of the many objectifications of the human spirit (*Geist*). His reflections on technology thus belong more to a philosophy of culture than to what might be termed a strict philosophy of technology.

It was Husserl's student, Martin Heidegger, who first recognized the growing importance of technology in shaping the human world, particularly in modern societies. Yet even in Heidegger's work, where technology was a prominent theme, technological objects and the experiences associated with them remained in the background. For the "question of technology," as the title of a well-known Heidegger essay reads, is less a phenomenological analysis of our experience of technological objects than an ontological account of our being in the world under the spell of technological mediation.[2] While the former refers to a punctual description of our manifold interactions with technological artifacts, the latter shifts the focus from the particular and individual to the broader framework of the general relationship of Dasein to Being and beings.

Similarly, Hans Blumenberg (1981, 35; my translation) emphasized that phenomenology's[3] concern with technology was not in describing our direct experience with technological artifacts, but rather in exploring "a condition of the human relation to the world." Accordingly, Blumenberg

preferred the term "technization" (*Technisierung*) over "technics" (*Technik*) or "technology" (*Technologie*), as the latter implies a hypostatization of the technological phenomenon. For these phenomenologists, technology is not an object for direct phenomenological description but rather a process, one often seen as having disruptive or negative effects on human life.[4]

Beyond these historical reasons, phenomenology's complicated relationship with technology is also rooted in structural issues. Phenomenology, as conceived by Husserl, posits that all knowledge is grounded in experience, and this experience is, by definition, immediate. However, technology is commonly understood as a form of mediation – something that stands between the human subject and the world, as well as between subjects (Van Den Eede 2011). In this sense, the essence of technology is mediation, which contrasts sharply with the phenomenological concept of intuition, or immediate experience. Intuition implies a more original, direct mode of engaging with the world, one that is seemingly incompatible with the presence of mediating technologies. This raises the question: Is such immediate, unmediated experience even possible in a world dominated by technology, or is it merely an idealized notion?[5] Either way, this tension between the immediacy of intuition and the mediation of technology presents a profound challenge for phenomenologists, as it threatens to upend their core assumptions about human experience.

This tension may, at least in part, explain why phenomenologists in the past have been reluctant to engage in a thorough phenomenological analysis of technological artifacts, leaving this work to other fields, such as the philosophy of technology. However, it is perhaps now, more than ever, time to confront this challenge from a new perspective.

4.2.2 The Technological Phenomenon as the Central Topic of a Phenomenology of Technology

A phenomenology of technology, as the term suggests, should concern itself with the *phenomenon* of technology. Rather than focusing on the essential nature of technological objects or the ontological categories that best describe their properties – tasks more suited to an ontology of technology[6] – phenomenology examines how objects present themselves to us. What is the mode of appearance (or "givenness") of technological objects? This is the question that any phenomenology of technology is called to answer.

My contention, perhaps somewhat provocatively, is that this question has not yet been adequately tackled by phenomenological approaches to technology. While there are a wide range of phenomenological perspectives on this topic, I will limit my analysis to three prominent examples

to demonstrate that, in my view, these approaches are structurally ill-equipped to directly address the technological phenomenon in terms of a direct encounter with technological objects.[7]

As mentioned earlier, the foundational phenomenological approach to technology originates with Heidegger (see Thomson 2009). Although many have attempted to distill a unified philosophy of technology from his works, two distinct approaches can be identified. The first, found in *Being and Time*, emphasizes the lived, practical experience of engaging with technological tools. We can regard this as Heidegger's early pragmatist approach to technology. The second approach, which became more prominent later in his career, takes a transcendental perspective, focusing on the underlying conditions that make technology possible (see Verbeek 2005, 7). This shift is especially evident in his essay *The Question Concerning Technology* (1954), where technology is viewed as a framework that shapes both thought and experience within the current phase of the history of Being. Most interpreters regard this transcendental approach as representative of Heidegger's overall philosophy of technology. Unlike his earlier pragmatist approach, Heidegger's late technological transcendentalism abstracts from direct encounters with technological objects and instead seeks to understand the conditions that drive the creation and use of artifacts. These conditions are subsumed under Heidegger's concept of "enframing" (*Gestell*), which views nature as raw material to be manipulated for human purposes. This transcendental perspective has been criticized for being overly abstract and detached from real-world technological practices (see Feenberg 1999, 186–87), and thus for failing to account for our concrete encounters with technological objects. In this sense, Heidegger's later phenomenology of technology overlooks the direct experience of technological phenomena, something his earlier analyses of tool use addressed more effectively.[8]

A recent approach to addressing the problem of technology within phenomenology is found in the school of thought known as postphenomenology. This approach emerges from a unique blend of classical phenomenology (Husserl, Heidegger), pragmatism (Dewey), and the so-called "empirical turn" in the 20th-century Dutch philosophy of technology (see Ihde 2009). Postphenomenology focuses on the relationship between humans, technology, and the world, particularly on how technology mediates a subject's interaction with the world and how this mediation shapes both the subject and the world. From this perspective, thematizing the phenomenon of technology – the mode in which technological objects are given – becomes challenging, if not impossible. This difficulty arises because, in postphenomenology, technology is primarily understood as a medium (or better, *the* medium) rather than as an object of direct experience. Since the postphenomenological view holds that every experience is technologically

mediated, even reflection on the medium of technology must itself be technologically mediated. As a result, the technological phenomenon, if it can even be said to exist within this framework, eludes direct reflective grasp. Technology is not simply hidden but transcends the categories of visibility and invisibility, while simultaneously being the transcendental condition that makes both possible.

A third example of a phenomenological approach that, for intrinsic reasons, is ill-suited to describe the technological phenomenon is Stéphane Vial's "historical phenomenology of technology" (Vial 2019). Similar to postphenomenological thinkers,[9] Vial conceives of technology as a medium that shapes our experience of the world and of ourselves. He introduces the term "ontophany" to describe how Being (*ontos*) appears (*phaino*), defining technology as an "ontophanic matrix" – a general structure of experience that conditions, a priori, how beings appear to us (Vial 2019, Ch. 3). In this sense, Vial's concept aligns with Heidegger's transcendentalism, as technology, understood as an ontophanic matrix, shapes human experience without being directly experienced itself. This is because transcendental conditions, by their nature, cannot be experienced in an immediate or "natural" way.[10] Consequently, in Vial's framework, phenomenology is not concerned with directly thematizing the experience of technological objects. Instead, it focuses on how the mediation of technological artifacts generally contributes to molding the structures of human experience.

To be clear, my intention is not to argue against technological transcendentalism, as proponents of postphenomenology have done and continue to do.[11] Rather, I am proposing a phenomenology of technology that focuses on the phenomenon of technological encounters, while still leaving room for transcendental analyses of how technology shapes experience. Transcendentalism and phenomenology, though distinct, do not need to remain separate as they often have been. They are complementary analyses that supplement each other and can benefit from each other.[12]

Historically, phenomenology has leaned more toward a transcendental approach to technology. My aim here, however, is to emphasize the other side of this inquiry: the direct encounter between the subject and technological objects. How technology appears and manifests itself is a phenomenon that phenomenology has yet to fully explore. This chapter represents a first step in that direction, though it is limited to one specific type of technology: the digital.

4.2.3 Phenomenology and Digital Technology

Digital technology – and more specifically, our encounter with it – poses significant methodological challenges for phenomenological description. On the one hand, it is unclear whether there is a distinct digital phenomenon

that lends itself to phenomenological inquiry. Before we can even speak of a phenomenology of digital technology – meaning a description of how digital technology appears or presents itself – we must first address two key questions: Does the digital have a mode of appearance? If so, what is its specific mode of appearance? Is it a fundamentally new mode, or can it be compared to other known modes of appearance? These questions probe the necessary conditions for a phenomenology of digital technology.

On the other hand, a common misconception has hindered the development of a true phenomenology of digital technology. While this misconception has impacted phenomenological research, it does not originate from phenomenology itself, but rather from media and technology studies – and it can also be found in everyday thinking about digital technology. The misconception lies in conflating the digital with the virtual. Particularly for those who grew up with technologies that integrate digitality with virtuality (millennials and digital natives), the digital is often perceived as inseparable from the virtual, leading to the frequent reduction of the former to the latter. However, this is a mistake. The digital and the virtual are distinct, each corresponding to different modes of appearance (see Vial 2019, Ch. 4). Put simply, the virtual is the visible surface of the digital. In fact, while virtuality depends on digitality, the digital can and has existed without the virtual. Until the 1980s, when graphical interfaces allowed users to interact with digital machines through virtual icons rather than code, virtuality was not associated with digital technology in the way it is today.[13] As a result, phenomenology has often overemphasized virtuality while overlooking the true nature and phenomenality of the digital itself.[14] Thus, a phenomenology of the digital remains an open and necessary task.

One of the key lessons of phenomenology is that what is most familiar to us is often the least understood.[15] Because digital technology is so pervasive and embedded in our daily lives, it frequently escapes our notice, becoming "invisible" in multiple ways that have yet to be fully explored. Therefore, any phenomenology of the digital must begin with the task of describing the different kinds of invisibility that shape our encounters with digital artifacts. This chapter thus narrows its focus to develop a typology of technological invisibility, which will then be applied to the elusive phenomenon of digital technology.

4.3 A Typology of Technological Invisibility

Heidegger once remarked that "[t]he meaning pervading technology [*der technischen Welt*] hides itself" (as quoted by Hörl 2015, 1). Put more simply,[16] technological objects have an inherent tendency to remain hidden, to become invisible. This invisibility manifests in various forms, each representing a different kind of concealment. For the sake of classification, we

can distinguish four types of technological invisibility. (1) *Perceptual invisibility* occurs when a technology's mechanism or functioning cannot be detected by the human senses. (2) *Tool invisibility* happens when the user operates the technology without being consciously aware of it. (3) *Hermeneutical invisibility* arises when the purpose or function of a technology is not discernible from its physical appearance. (4) *Environmental invisibility* takes place when a technology shapes our behavior unconsciously by being integrated into the very fabric of our environment. This section explores each of these types in more detail.

4.3.1 Perceptual Invisibility of Technology

In its most straightforward sense, technological invisibility refers to the imperceptibility of technological objects. Something is invisible in this context when it is literally out of sight, beyond the field of one's current perception. For example, the water infrastructure that brings water into our homes remains largely imperceptible. Water pipes are buried underground and hidden from view, leaving us with only the visible faucet in the kitchen or bathroom – just a small part of the technology that provides access to water without needing to leave the house. Another example is the network of internet servers that support our daily interactions online. We access websites and send emails without any awareness of the complex data storage and transmission processes happening in the background. Thus, much of the technology that enables these tasks remains imperceptible, as we focus primarily on the visible content and functionality of the devices we interact with.

Despite its apparent simplicity, this concept of technological invisibility raises important questions. Consider this example: I cannot directly see the engine that powers a car unless I open the hood and examine it. But even then, I can only see the surface of the engine, with many critical components still hidden. If I were to take the engine apart and examine its parts one by one, could I really say that I was "seeing" the technology? Isn't the technological object greater than the sum of its visible parts? In fact, what defines a technological object is its mechanism, which often remains hidden until it is set in motion. As Hans Blumenberg observed: "Technology is phenomenally a field of mechanisms" (Blumenberg 1981, 54; my translation). A mechanism is realized through the dynamic interaction of the parts that make up the technological object. But the mechanism itself does not reveal itself directly; what we perceive are the results of its operation. To return to the example of the car, the car's true nature as a technological object is revealed through its movement, not through the mere observation of its individual parts. A stationary car is more like inert matter, while its technological essence is revealed only when it is in motion. In this sense,

the mechanism of any technological object is opaque and can only be perceived indirectly, through its operational effects.[17]

In summary, for an observer to perceive a technological object, its mechanism must be present in some way. However, being present does not necessarily mean being directly perceivable. This supports the general idea that any technology relying on a mechanism is bound to remain invisible in the sense of being imperceptible.

This notion resonates with Max Weber's famous example of the tram driver, who need not know the mechanical laws that make the tram move, nor how its internal mechanisms work. While it could be argued that a driver must have some understanding of how the tram functions, all that is truly required is practical knowledge – how to operate the tram, interact with its controls, and make it move. This is distinct from understanding the actual mechanisms of the tram. In this sense, the driver's knowledge is "know-how" rather than "know-that" – the driver need only know how to engage with the machine to achieve a result, without understanding the underlying mechanisms, which may remain a "black box" to them. Yet, the driver can still be highly skilled in operating the tram.

At this point, it is useful to expand on the concept of the "black box" to clarify perceptual invisibility. Originally introduced in computer science, a "black box" refers to a system that can be understood only by its inputs and outputs, without knowledge of its internal workings. Black box systems are "opaque," meaning that their internal processes are hidden from view. Invisibility in this context refers to the inability to perceive these mechanisms with the senses. A black box is something that remains out of sight, containing an invisible side that is, at least for the moment, unreachable. Bruno Latour elaborated on this concept by coining the term "blackboxing," which he described as:

> the way scientific and technical work is made invisible by its own success. When a machine runs efficiently, when a matter of fact is settled, one need focus only on its inputs and outputs, not on its internal complexity. Thus, paradoxically, the more science and technology succeed, the more opaque and obscure they become.
>
> (Latour 1999, 304)[18]

In this way, Latour recognizes perceptual invisibility as a natural result of technological progress: As technology evolves, it conceals itself further. While "blackboxing" describes the process of increasing technological invisibility, it also reflects only one particular kind of invisibility – what I refer to as the imperceptibility of technological objects. Technologies become black boxes because they contain "programs of actions" or "scripts" that must be followed by various actants (human, nonhuman, or both)

to function.[19] The actual processes of these technologies are not perceived by the actants; instead, their focus is on the outcomes, leaving the mechanisms largely obscure.

4.3.2 Use Invisibility of Technological Tools

Another form of technological invisibility occurs when a technological object becomes invisible to the user during use. In the act of using a tool, the technology tends to "disappear" by becoming transparent. This concept of transparency in technology requires further clarification.

Heidegger introduced the idea of technological transparency in his analysis of tools in *Being and Time* (Heidegger 2010, §§ 15–18; cf. on this Dreyfus 1991, 64–69). Heidegger's famous example is the hammer: When using the hammer, we do not focus on the tool itself or concern ourselves with its internal structure (Heidegger 2010, 68–69). As Heidegger puts it, "the less we just stare at the thing called hammer, the more we take hold of it and use it, the more original our relation to it becomes and the more undisguisedly it is encountered as what it is, as a useful thing" (Heidegger 2010, 69). He also emphasizes that using a tool is not a blind activity; there is a specific type of seeing involved, which he calls "circumspection" (*Umsicht*) (Heidegger 2010, 69). However, the focus of circumspection is not the tool itself, but the "work" (*Werk*) to be accomplished. While the user's attention is on the task, the tool "withdraws, so to speak, in its character of handiness in order to be really handy. What everyday dealings are initially busy with is not tools themselves, but the work. What is to be produced in each case is what is primarily taken care of and is thus also what is at hand" (Heidegger 2010, 69). The tool only becomes visible again when it malfunctions, at which point both the tool and the broader technological context of its use come into view.[20]

A similar concept of technological transparency can be found in the well-known example of the blind man's cane, originally described by René Descartes.[21] The example is to be found in his *Optics*, and it deserves to be quoted at length:

> No doubt you have had the experience of walking at night over rough ground without a light, and finding it necessary to use a stick in order to guide yourself. You may then have been able to notice that by means of this stick you could feel the various objects situated around you, and that you could even tell whether they were trees or stones or sand or water or grass or mud or any other such thing. It is true that this kind of sensation is somewhat confused and obscure in it is those who do not have long practice with it. But consider it in those born blind, who have made use of it all their lives: with them,

> you will find, it is so perfect and so exact that one might almost say that they see with their hands, or that their stick is the organ of some sixth sense given to them in place of sight.
>
> (Descartes 2006, 153).

At the moment of use, the stick fades into the background of consciousness, and the user experiences the illusion of directly perceiving the world through it. Of course, this illusion is very unstable: At any moment, the person walking in the dark or the blind person can become aware of the mediated relation to the world caused by the use of the cane. But there are times when this mediation is not even perceived as such so that Descartes can say that the blind people (literally?) "see with their hands." This suggests the idea that all of our experience is in fact mediated through either natural means (our sense organs) or technological means (the cane, but also computer devices, etc.).

However, there is an important distinction between Heidegger's example of the hammer and Descartes' example of the blind man's cane. While both illustrate technological transparency, they involve different types of engagement with the tool. In Heidegger's analysis, the hammer is used in the context of production (*poiein*), where the focus is on shaping materials to achieve a goal. The hammer's transparency allows the user to focus on the task, with the tool serving as an instrument to accomplish the work. In contrast, Descartes' example of the cane involves "seeing"(*theorein*) or perceiving the world, where the cane functions as a medium that facilitates the blind person's sensory experience of their surroundings. The hammer is used actively in the process of creation, while the cane passively gathers information about the environment. These differences suggest that the nature of technological transparency varies depending on whether the tool is used for production or perception.[22]

It is also important to note that while a tool may become transparent during use, this does not mean it is entirely unperceived. There is a distinction between technological transparency, where the tool fades from conscious attention but is still perceived, and perceptual invisibility, where the tool is completely unseen. For instance, a carpenter may not consciously focus on the hammer while working, but the hammer is still present in their awareness.[23] In this case, the invisibility is subjective, depending on the user's focus, rather than being an inherent property of the tool itself, as in the case of perceptual invisibility.

Technological transparency highlights a paradox: Tools recede from our awareness, allowing for greater efficiency in use, yet this same invisibility can obscure the extent of our reliance on them. Recognizing this dual nature is crucial for a phenomenological understanding of how technology mediates not only our actions but our ethical responsibilities toward its use.

4.3.3 The Obsolescence of Appearance of Technological Objects or Purpose Invisibility

In *The Obsolescence of Man*, Günther Anders proposes a classification of things in today's human environment according to their modality of appearance (Anders 2018, 347). He identifies three categories: things that appear as they are, things that appear less than they are, and things that appear more than they are. At first glance, these categories seem to correspond to the traditional ontological distinction between natural things and artifacts, with the addition of another category, that of commodities.[24] However, this interpretation is mistaken. Anders is not referring to ontological categories, but rather to phenomenological ones – general modalities in which objects appear. Moreover, he does not attribute each modality exclusively to a specific ontological class, as not all natural things appear as they are (consider cases of natural mimicry), and not all artifacts appear less than they are.

Anders refers to the "obsolescence of appearance" (Anders 2018, 37–39) as a phenomenon where an object appears less than it is – a characteristic that he argues is typical of modern technological artifacts.[25] One of the most striking examples he provides is the Zyklon B cans (Anders 2018, 38), which were used during the Holocaust to murder millions of people in Nazi gas chambers. To an uninformed observer, these cans look ordinary and harmless. In fact, Anders recalls that during his first visit to Auschwitz, he nearly mistook them for jam jars (Anders 2018, 473). Even though they are physically "perceivable" (*wahrnehmbar*), these artifacts remain "unnoticeable" (*unerkennbar*) in their true function (Anders 2018, 38). They are "silent," in the sense that their appearance is unable to "express" their true "functionality." Anders contrasts this with traditional artifacts like hammers or chairs, whose appearance clearly reflects their function (Anders 2018, 38). No one mistakes a hammer or a chair for something else, as their purpose is visibly apparent.[26]

Anders (2018, 39) concludes that modern technological artifacts no longer qualify as "phenomena" in Heidegger's sense, where a "phenomenon" is "that which shows itself in itself, the manifest" (Heidegger 2010, 28). Technological objects today do not reveal their essence through their appearance. Instead, "their performance lies in the fact that they do not show what they are, but rather *conceal* themselves" (Anders 2018, 39; my translation). They are "seemingly visible […] but in truth invisible […]. Attempting to grasp their meaning through our senses would be a completely nonsensical endeavor" (Anders 2018, 39; my translation). In other words, their appearance is disconnected from their actual function or being.

Looking at Anders's examples more closely, we can see another dimension of this obsolescence of appearance. The invisibility of a Zyklon B can is different from that of a satellite or a computer. Satellites, for example,

are not part of our everyday encounters; they remain out of sight, orbiting Earth. When we use a navigation device, we are aware of the device itself, but few people realize they are actually relying on the function of a satellite. The device serves as an interface between us and an object that remains entirely invisible, creating an illusion of direct interaction. As in the case of the Zyklon B can, we attribute a technological function to the wrong object. The navigation device is merely a medium, not the source of the function.[27]

Anders (2018, 472) identifies two key reasons why the functionality of technological artifacts has become invisible. First, technology has grown so complex that it surpasses our sensory and imaginative capacities (*Sinnlichkeit*). New technologies are "supersensible" (*übersinnlich*), meaning they exceed what we can perceive or even comprehend. Second, engineers aim to optimize the functionality of technology, often at the expense of its appearance. From the engineer's perspective, the outward appearance of a technological artifact is largely irrelevant and may even be considered superfluous.[28] This also applies to digital engineering. As noted by Wendt (2013), in interface design, "invisible" often means "seamless, efficient, and functionally optimized," while "visible" suggests "cumbersome, inefficient, and functionally suboptimal." Thus, achieving transparency in design is seen as a hallmark of high functionality.

These two factors – technological complexity and the desire to "hide" technology – often combine to render modern technology invisible to the average user. While the intentional invisibility created by engineers is deliberate and can sometimes be mitigated, the invisibility caused by complexity is neither chosen nor easily overcome. In a dialogue with a painter who wanted to depict an atomic bomb, Anders advises against it, calling the task impossible (Anders 2018, 359–372). The atomic bomb is "unsightly" (*unansehnlich*), not because it is ugly, but because it eludes human perception and imagination. It is "supersensible" – beyond the scope of our sensory and cognitive abilities (Anders 2018, 359–60). We cannot fully grasp the reality of an atomic bomb because we cannot see or imagine its destructive potential.

The form of technological invisibility discussed here refers specifically to the purpose of a technological object. One could argue that technology is not truly invisible if only its purpose is hidden. However, this objection overlooks the fact that technological objects are defined by their functionality. If the purpose of a technological object is not perceivable, then the object itself, in its technological essence, has become invisible.

The obsolescence of technological appearance illustrates how modern tools and devices often conceal their true complexity, leading users to engage with them in a superficial way. This phenomenon is especially

relevant in the digital age, where the surface appearance of technology belies the intricate processes at work beneath, contributing to its pervasive invisibility.

4.3.4 The Technological Unconscious or Environmental Invisibility of Technology

Technology becomes invisible when it is integrated into our environment, a phenomenon more prevalent in modern societies than in the past. Today, almost every aspect of our existence is technologically mediated, whether we are aware of it or not. This mediation shapes how we perceive the world, interact with others, and reflect on ourselves. Various authors from different perspectives have highlighted this fundamental condition of modern life. Hörl (2015) aptly calls it the "technological condition," while Hayles (2016, 34) refers to the "technological unconscious," describing the infrastructures that subtly shape our assumptions, actions, habits, and responses. Drawing on Husserl, Blumenberg (1981, 40; my translation) also emphasized this aspect of technology's role in our lifeworld: "The artificial reality, a foreigner among the found things of nature, at some point sinks into the 'universe of self-evident things,' into the lifeworld." But how does technology become such an integral part of our environment that it becomes indistinguishable from natural things? What kind of invisibility is at work here?

The first point to clarify is that when something becomes an integral part of our environment, it ceases to be visible. Blumenberg asserts: "No more perfectly than by the mimicry of casings, the technical as such becomes invisible when it is implanted in the lifeworld" (Blumenberg 1981, 41; my translation). This is not a matter of degrees, but rather a shift in perception. While natural camouflage leads to perceptual invisibility, the sinking of technology into the lifeworld results from its "naturalness" or "straightforwardness" (*Selbstverständlichkeit*).[29] For Blumenberg, what defines naturalness is its ability to suppress our awareness of contingency – the possibility that things could be different. As he explains, "[c]ontingency means the evaluation of reality from the standpoint of necessity and possibility" (Blumenberg 1981, 50; my translation). Many, including Jaspers (1998), identify the spirit of modernity with an awareness of the contingency of the world, recognizing that everything we encounter exists but could have been otherwise. This awareness is central to the technological attitude of modernity, where change and innovation are seen as constant possibilities.[30] On the other hand, the concept of the lifeworld, as Blumenberg inherits from Husserl, denies this sense of contingency. In the lifeworld, "[e]verything that is real in the lifeworld plays into life, is used and consumed, sought and fled, but it remains concealed in its *contingency*, i.e.

not perceived as also-being-able-to-be-other." (Blumenberg 1981, 26; my translation). The phenomenological task is to overcome this naturalness by bracketing the lifeworld to make it an object of theoretical inquiry.

Once embedded in the lifeworld, technology is perceived as "natural," losing its contingency and becoming an accepted part of life. Blumenberg illustrates this with the example of the electric doorbell (Blumenberg 1981, 38–40). In the past, doorbells operated mechanically, giving the user a direct sense of causing the sound – by pulling a bell cord or turning a handle, the user could physically feel the connection between their action and the sound produced. In contrast, with electric doorbells, pressing a button separates the action from the effect. The user no longer feels like they are producing the sound themselves; instead, they simply trigger the process, which happens invisibly and automatically.[31] The button masks a long history of human discovery and technical achievement behind its uniform simplicity. Moreover, as Blumenberg notes:

> The technical object not only repels curious questions […] but seems to do everything to prevent questions from arising in the first place – questions not only about its construction and function but especially about its justification for existence.
>
> (Blumenberg 1981, 40; my translation)

The button exemplifies how technology loses its contingent character through its embedding in the lifeworld. We no longer ask how the doorbell works or why it exists; we simply press it. In this way, technology becomes invisible because it is taken for granted, shaping our lives without drawing attention to itself.

In sum, technology becomes invisible when it integrates seamlessly into our environment, becoming a naturalized part of the lifeworld. This environmental invisibility is especially pronounced in the case of digital technologies. As they mediate nearly every aspect of our modern existence, digital technologies not only blend into the background but also exhibit various other forms of invisibility. The next section explores these multiple layers of invisibility, showing how digital technology builds on the same principles of concealment yet in more complex ways.

4.4 The Manifold Invisibilities of the Digital Phenomenon

Having explored the multiple dimensions of technological invisibility, it becomes clear that digital technology, in particular, is the culmination of this dialectic of concealment. It incorporates perceptual, use, purpose, and environmental invisibility in ways that obscure not only the technology itself but also its broader societal implications. In what follows, we will

explore how these forms of invisibility manifest in the digital age and why they present unique challenges for contemporary thought.

4.4.1 The Digital Unseen: The "Subface"

The first form of invisibility in digital technology is imperceptibility – the fact that we cannot directly perceive digital processes with our senses. One might argue that we interact with digital technology daily, when we use computers, smartphones, and other devices. However, as we saw with the example of the car in Section 4.3.1, what we see when we interact with these devices is not the digital technology itself but rather its external, physical shell.

When a computer is turned off, we see only its outer casing. Beneath this shell are the internal components – the "infrastices" – such as the electrical, chemical, and mechanical systems that make digital operations possible (Anusas and Ingold 2013, 58).[32] These hardware components, while essential, are not directly visible in normal use. Even if we open a computer case, we see only parts of the machine's physical structure, not the digital processes themselves. To capture the hidden nature of digital operations, Nake (2015, especially 105) introduced the concept of "subface." The subface refers to the unseen part of the computer's memory – the display buffer – that software uses to represent content on the screen. It is distinct from the "surface," or the visible screen interface. Both subface and surface are essential to the "algorithmic thing" that defines the computer, but the subface is what remains hidden during normal interactions.

Groys (2009, 339) further illustrates this point by arguing that while we can analyze computer hardware and software in their inactive state, to interact with digital data, we need a process of visualization that makes the invisible visible. Digital images, for example, are not direct representations of the world but rather visualizations of code. No one has ever seen a digital file; what we see are renderings of digital data, mediated through the screen. Thus, humans are always on the outside of the digital world, interacting through interfaces that both reveal and obscure the underlying processes.

4.4.2 Broken Mirrors

The use invisibility of digital tools, such as smartphones, computers, and digital watches, functions similarly to the use invisibility of more traditional tools like hammers or cars. When using these devices, we are rarely conscious of the technology itself. For instance, as I type these words, my attention is not on the computer or its components, but on the task at hand – writing. The computer only becomes noticeable when something

goes wrong, such as a slowdown in processing or a malfunction. At that point, the device interrupts the task and draws attention to itself.

While this form of invisibility is similar to that of traditional tools, the complexity of digital devices adds another layer. The vast network of processes running in the background, from software to internet servers, remains hidden unless something disrupts their operation. Thus, when we use digital technology, it functions like a mirror that reflects our goals but hides its own workings – until the mirror cracks.

4.4.3 A Jacket Is More Than a Jacket

The purpose invisibility of digital technology can occur in two main ways. The first happens when we fail to recognize the digital nature of an object. As digital technology becomes more integrated into everyday objects, it becomes less apparent to users (see Norman 1999). For example, many household appliances – such as microwaves, dishwashers, and washing machines – contain computers, but we do not typically think of them as digital devices. This purposeful invisibility in design is meant to make technology seamless, hiding the complexity inside the object.

A second form of purpose invisibility arises when objects serve more than their apparent function. A prime example is smart textiles, such as Levi's Commuter Trucker Jacket with Jacquard by Google, which allows wearers to control their smartphones using touch gestures on the jacket's sleeve. The jacket functions as clothing, but its digital capabilities remain hidden unless they are intentionally used, blending digital functionality with everyday wear.

4.4.4 Fitting Like a Glow to a Hand

Thrift (2004, 183) observed that "computing […] will become a new kind of surface, fitted to activity-in-context like a glove is fitted to a hand." This concept introduces environmental invisibility in digital technology. Digital systems have become so ubiquitous that they are embedded in every aspect of our lives, often without our awareness. Mark Weiser, a pioneer of ubiquitous computing, emphasized that "[t]he most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" (Weiser 1991, 94).

Since the early 1990s, computing has become increasingly integrated into our daily and working practices, sinking, as Thrift and French (2002, 318) note, "from the representational into the non-representational world, so becoming a part of a taken-for-granted set of passions and skills." Digital technologies have thus become part of our lifeworld, where they function invisibly, shaping how we act and perceive the world without drawing attention to themselves.

4.5 Concluding Remarks

This chapter has offered an initial exploration of the invisible dimensions that digital technology has assumed in the contemporary world. Far from being an obstacle to phenomenological analysis, this invisibility is precisely what makes a phenomenology of the digital both challenging and, for the same reason, more urgent than ever. As Heidegger observed in *What Is Called Thinking?*, what prompts us to think is that which has not yet been fully thought through, and therefore deserves our attention (Heidegger 1976, 6). In this spirit, we have sought to highlight the aspects of digital technology that, though not immediately visible, structure our experience precisely through their concealment.

Understanding these invisible dimensions is not merely an abstract philosophical exercise. By delineating different forms of invisibility – whether perceptual, use-related, purpose, or environmental – this analysis offers a framework for understanding how digital technologies increasingly shape the conditions under which we live, work, and interact. These hidden mechanisms are not neutral; they exert influence over our choices and behaviors, often without our awareness.

Phenomenology, in this context, takes on a critical role by exposing the ways in which digital technology subtly governs our actions and assumptions. While digital systems increasingly define the contours of our existence, they are not fated or unchangeable. Once we become aware of these invisible forces, we can begin to question and challenge them, opening up space for more conscious decisions about how we want technology to shape our future.

In sum, the task of phenomenology is not merely to describe the invisible aspects of digital technology but to make these structures available for practical reflection and, ultimately, conscious choice. It is only through such awareness that we can begin to regain agency in a world increasingly shaped by systems that remain out of sight, yet profoundly influence our lives.

Notes

1 Also, Groys (2009, 339) refers to this neglected aspect of the movie *The Matrix*. He notices that the heroes of the movies are the only ones who can see the real world, which is the world of the machines. They are thus able to see the digital code beyond or below the digital image. This is of course a fiction. For humans are not allowed to literally see the digital code with their eyes. The latter is not-literally "seen" only by the computer machine, which makes us able to perceive it in the form of a digital image. The seen digital image is a reproduction of the unseen digital code. The visible is a copy of the invisible. Groys remarks that the situation made possible by the use and consume of

digital images resembles that of the Byzantine icons whose wonder consists in making the invisible God visible.

2 Certainly, phenomenology remains for Heidegger the only and proper method of ontology (Heidegger 2010), but then the question arises how we should understand phenomenology as a method of ontological research, since this is not what the founder of phenomenology had in mind for his discipline.

3 Blumenberg (1981) refers, in particular, to Husserl's late phenomenology of the lifeworld as laid down in the *Crisis of European Sciences and Transcendental Phenomenology*.

4 Cf. Blumenberg (1981) for a thorough and compelling account of Husserl's main argument, especially the following: "[…] Husserl geht es nur darum, das Verhängnis menschlichen Handelns im weitesten Sinne dort, wo es nicht mehr weiß, was es tut, exemplarisch sichtbar zu machen und die sozusagen aktive Unwissenheit als die Wurzel all derjenigen desorientierten Aktionen bloßzulegen, die die menschliche Ratlosigkeit in der technischen Welt hervorrufen haben" (36).

5 Postphenomenology poses this question in rhetorical terms: Intuition is unachievable since every experience is mediated. Needless to say, this way of arguing is misguided. For how can one recognize mediation when there is no way to compare it to an immediate experience?

6 As a prominent example of such an ontology, with a special focus on the digital artifact, see Hui (2016).

7 I implicitly refer here to the notion of the "encounter" as explained by Spiegelberg (1984). In this respect, a phenomenology of (digital) technology attempts a description of the encounter with the (digital) technological object.

8 I consider Heidegger's tool analyses in Section 4.2.2.

9 In fact, Vial (2019, preface to the US edition) explicitly recognizes the parallelism and complementarity of his approach to that of postphenomenology. Unfortunately, he does not specify in the book what these parallels and complementarities are.

10 Here, one ought to recall the fact that the word "transcendental" entails different meanings depending on if one takes it from either Kant or Husserl. For Kant, a "transcendental experience," meaning the experience of the conditions of possibility of experience, truly is a *contradictio in adjecto*, whereas Husserl's phenomenology is based on the conviction that transcendental conditions are "experienceable." The quotation marks are here necessary since for Husserl transcendental experience or, in his own terms, phenomenological experience has a fundamentally different character than "natural" experience.

11 Admittedly, one can notice a tension in the postphenomenological program. For if, on the one hand, postphenomenologists criticize the transcendentalism of classic philosophy of technology and explicitly shift their interest to the specific interaction of humans with particular technological artifacts, on the other hand, their approach rests on a relational ontology that emphasize the transcendental structure of mediation over the phenomenon of the encounter with the technological artifacts. For an attempt to connect postphenomenology with object-oriented philosophy, see Van Den Eede (2022).

12 A promising syncretic approach, combining the postphenomenological attention for empirical interactions with technology together with a "return to transcendentalism," is provided by Bernard Stiegler's "organology of technics" (see on this Lemmens 2017).

13 The word virtuality stems from medieval Latin *virtualis* and was used in the Middle Ages to translate the concept of potentiality (*dynamis*) of Aristotelian scholastic philosophy. The term is employed in a similar sense also in 19th- to 20th-century French philosophy (Ravaisson, Bergson, Deleuze among others).
14 See, e.g., du Toit (2020).
15 "Philosophy has the strange business of stopping, shaking, and scrutinizing the everyday, thoughtlessly practiced understanding of being from the distance of an original astonishment," Eugen Fink (1976, 182; translation is mine) once remarked, and his reference was of course to phenomenological philosophy.
16 The difference between Heidegger's approach and the one suggested here is simply stated. Like all exponents of the classical philosophy of technology, Heidegger views technology from a perspective called "transcendentalism" (Verbeek 2005, 7), which aims exclusively at establishing the conditions of possibility of technology. In contrast, the approach that I take here is that of analyzing technology in terms of its concrete artifacts, which comes close to "the philosophy of technological artifacts" as outlined by Verbeek (2005, 9). I thus will not refer here to Technology with the capital "T," but to technology in the sense of technological objects. See also the methodological reflections in Section 4.1.2.
17 This applies not only to the physiochemical or digital mechanisms of technological systems, but also to all kinds of system, including social and psychological systems. In this sense, scientific explanations are currently understood as hypotheses that explain the hidden mechanisms of any given system (Bunge 2004).
18 Latour has described the phenomenon of blackboxing already in *Science in Action* (Latour 1987).
19 Latour's famous example is the Berlin key (Latour 1991).
20 Heidegger speaks of three forms of "breakdown" of the technological use: conspicuousness, obstrusiveness, and obstinacy (see Heidegger 2010, § 16).
21 Merleau-Ponty, Wittgenstein, Polanyi, Bateson, and Dreyfus all employ this example and mostly without referencing to each other or to Descartes.
22 This topic would of course deserve a more elaborate phenomenological analysis, which space does not permit.
23 The same logic applies, *mutatis mutandis*, to the case of the blind man's cane. Although the blind man does not literally see the cane, he perceives it through his hand. The tactile sensations of the cane do not enter his conscious awareness, as he uses it to navigate and orient himself in space.
24 With commodities (*Waren*), Anders has in mind the products of the advertising industry. The function of advertising is to project (natural or artificial) desires, wills, and needs onto products that are in most cases incapable of satisfying them. In this way, commodities appear to be more than they actually are. In what follows, I will abstract from Anders' phenomenology of commodities and focus specifically on technological artifacts.
25 Anders observes that this phenomenological modality is not specific to the machines of his generation, but also to those of his grandfathers' generation, i.e., the machines of the second industrial revolution (Anders 2018, 39). However, he also detects a general historical trend toward increasing technological invisibility, i.e., toward the increasing obsolescence of appearance. The huge and noisy machines of 19th-century heavy industry were certainly

more "visible" than today's satellites and computers, as Anders (2018, 39) justly remarks.

26 For the sake of argument, I am not considering here pathological cases such as those described by the British neurologist Oliver Sacks in his popular book with the telling title *The Man Who Mistook His Wife for a Hat*.

27 What about computers, the third of Anders' examples of technological invisibility? Unfortunately, Anders does not specify it.

28 Of course, what an engineer thinks is the best is often influenced by a whole society's or culture's values and beliefs. Anders' remarks are intended for Euro-American societies at the time he is writing (1948–1980). Indeed, it can also be applied to contemporary industrial societies. In particular, the advent of ubiquitous computing (Weiser 1991) is nothing more than a refined implementation of the engineered invisibility that Anders observed in the technologies of his time.

29 The term *Selbstverständlichkeit* comes from Husserl and literally means what-understands-itself-by-itself. *Selbstverständlich* is thus that which is a matter of course and as such shapes our life in an unconscious way.

30 "Das Bewußtsein der Kontingenz der Wirklichkeit ist nun aber die Fundierung einer technischen Einstellung gegenüber dem Vorgegebenen: […] dann wird die Faktizität der Welt zum bohrenden Antrieb, nicht nur das Wirkliche vom Möglichen her zu beurteilen und zu kritisieren, sondern auch durch Realisierung des Möglichen, durch Ausschöpfung des Spielraums der Erfindung und Konstruktion das nur Faktische aufzufüllen zu einer in sich konsistenten, aus Notwendigkeit zu rechtfertigenden Kulturwelt" (Blumenberg 1981, 50–51).

31 What consequences this divorce can bring about has been aptly discussed by Anders (2018, 75 ff.) with reference to the case of the pilot of Hiroshima who pushed a button through which the bomb was dropped on the city of Hiroshima on August 6, 1945.

32 "Infrastitial" comes from the Latin *infra* meaning "below, beneath" and *stāre* meaning "to stand."

References

Anders, Günther. 2018. *Die Antiquiertheit des Menschen II: Über die Zerstörung des Lebens im Zeitalter der dritten industriellen Revolution*. 4th ed. München: Beck.

Anusas, Mike, und Tim Ingold. „Designing Environmental Relations: From Opacity to Textility". *Design Issues* 29, Nr. 4 (October 2013): 58–69. https://doi.org/10.1162/DESI_a_00230.

Blumenberg, Hans. 1981. "Lebenswelt und Technisierung unter Aspekten der Phänomenologie." In *Wirklichkeiten, in denen wir leben: Aufsätze und eine Rede*, 9–58. Reclams Universal-Bibliothek 14015. Stuttgart: Reclam.

Bunge, Mario. 2004. "How Does It Work?: The Search for Explanatory Mechanisms." *Philosophy of the Social Sciences* 34 (2): 182–210. https://doi.org/10.1177/0048393103262550.

Descartes, René. 2006. "Optics." In *The Philosophical Writings. Vol. 1*, edited by John Cottingham, 152–75. Cambridge: Cambridge University Press.

Dreyfus, Hubert L. 1991. *Being-in-the-World: A Commentary on Heidegger's Being and Time, Division I*. Cambridge, MA: MIT Press.

du Toit, Jean. 2020. "Introduction – Phenomenology and Virtuality." *Indo-Pacific Journal of Phenomenology* 20 (1): 1–3. https://doi.org/10.1080/20797222.2021.1896236.

Feenberg, Andrew. 1999. *Questioning Technology*. London, New York: Routledge.

Fink, Eugen. 1976. *Nähe und Distanz: Phänomenologische Vorträge und Aufsätze*. Freiburg im Breisgau, München: Alber.

Groys, Boris. 2009. "Invisibility of the Digital: Religion, Ritual, Immortality." *Res: Anthropology and Aesthetics*, *55–56* (Spring/Autumn 2009), 337–40. https://doi.org/10.1086/RESvn1ms25608854.

Hayles, N. Katherine. 2016. "Cognitive Assemblages: Technical Agency and Human Interactions." *Critical Inquiry* 43 (1): 32–55. https://doi.org/10.1086/688293.

Heidegger, Martin. 1976. *What Is Called Thinking?* Translated by J. Gleinn Gray. New York: Harper & Row.

Heidegger, Martin. 2010. *Being and Time*. Translated by Joan Stambaugh and Dennis J. Schmidt. SUNY Series in Contemporary Continental Philosophy. Albany: Albany: State University of New York Press.

Hörl, Erich. 2015. "The Technological Condition." Translated by Anthony Enns. *Parrhesia* 22: 1–15.

Hui, Yuk. 2016. *On the Existence of Digital Objects*. Minneapolis, London: University of Minnesota Press.

Ihde, Don. 2009. *Postphenomenology and Technoscience: The Peking University Lectures*. State University of New York Press.

Jaspers, Karl. 1998. *Die geistige Situation der Zeit*. Berlin, New York: de Gruyter.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press.

Latour, Bruno. 1991. "The Berlin Key or How to Do Words with Things." In *Matter, Materiality and Modern Culture*, edited by Paul Graves-Brown, 10–21. London, New York: Routledge.

Latour, Bruno. 1999. *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.

Lemmens, Pieter. 2017. "Thinking through Media: Stieglerian Remarks on a Possible Postphenomenology of Media." In *Postphenomenology and Media: Essays on Human–Media–World Relations*, edited by Yoni van den Eede, 185–206. Postphenomenology and the Philosophy of Technology. Lanham: Lexington Books. https://repository.ubn.ru.nl/handle/2066/176615.

Nake, Frieder. 2015. "Surface, Interface, Subface: Three Cases of Interaction and One Concept." In *Paradoxes of Interactivity: Perspectives for Media Theory, Human-Computer Interaction, and Artistic Investigations*, edited by Uwe Seifert, Jin Hyun Kim, and Anthony Moore, 92–109. Bielefeld: transcript Verlag.

Norman, Donald A. 1999. *The Invisible Computer: Why Good Products Can Fail, the Personal Computer Is so Complex, and Information Appliances Are the Solution*. Cambridge, MA: MIT Press.

Spiegelberg, Herbert. 1984. "Three Types of the Given: The Encountered, the Search-Found and the Striking." *Husserl Studies* 1 (1): 69–78.

Thomson, Iain. 2009. "Phenomenology and Technology." In *A Companion to the Philosophy of Technology*, edited by Jan Kyrre Berg Olsen Friis, Stig Andur Pedersen, and Vincent F. Hendricks, 195–201. Oxford: Wiley-Blackwell.

Thrift, Nigel. 2004. "Remembering the Technological Unconscious by Foregrounding Knowledges of Position." *Environment and Planning D: Society and Space* 22 (1): 175–90. https://doi.org/10.1068/d321t.

Thrift, Nigel, and Shaun French. 2002. "The Automatic Production of Space." *Transactions of the Institute of British Geographers* 27 (3): 309–35.
Van Den Eede, Yoni. 2011. "In between Us: On the Transparency and Opacity of Technological Mediation." *Foundations of Science* 16 (2): 139–59. https://doi.org/10.1007/s10699-010-9190-y.
Van Den Eede, Yoni. 2022. "Thing-Transcendentality: Navigating the Interval of 'Technology' and 'Technology'." *Foundations of Science* 27 (1): 225–43. https://doi.org/10.1007/s10699-020-09749-y.
Verbeek, Peter-Paul. 2005. *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. Translated by Robert P. Crease. University Park, PA: Pennsylvania State University Press.
Vial, Stéphane. 2019. *Being and the Screen: How the Digital Changes Perception. Published in One Volume with a Short Treatise on Design*. Cambridge, MA: MIT Press.
Weiser, Mark. 1991. "The Computer for the 21st Century." *Scientific American* 265 (3): 94–104.
Wendt, Thomas. 2013. "Designing for Transparency and the Myth of the Modern Interface." *UX Magazine*. https://uxmag.com/articles/designing-for-transparency-and-the-myth-of-the-modern-interface.
Wrathall, Mark A., ed. 2021. *The Cambridge Heidegger Lexicon*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9780511843778.

5 The Phantom Matrix

A Critical Phenomenology of Television

Nicolas de Warren

Do you remember flipping through the TV Guide or a listing of television programs in the Sunday paper with a dedicated sense of purpose, eagerly planning your viewing schedule for the week, and expectantly waiting for your favorite series, to be watched alone, or a variety show viewed with friends or family, or hunting for a movie you'd never seen before, or a classic you longed to see once again? Do you remember lazy hours in front of the television, surfing channels in the hope of something, anything, to hold your attention, or enjoying obscure films, or late-night programming, until dawn? These and other recollections of television, from a world before iPhones, YouTube, streaming services and, more generally, the internet, mark my generation as the last generation to have come of age with (only) television. Generations are dated by the technologies for which they feel nostalgic. The loneliness of television, the companionship of television, the sociality of television, the thrill of game shows, the sobriety of nightly news, "we interrupt this program with breaking news," gossipy late-night talk shows, and – memorably – advertisements; these and other varieties of television experience belong largely to the past. For the post-television generation, whose attachment to pixelated screens is essentially defined by personal computers, mobile phones, and tablets, "on demand" digital streaming contrasts markedly with the fixity of television. Pre-scheduled programming and fixed placement of television in the home have become displaced by the portability of screens and anytime, anywhere access. Even for my generation, who witnessed on a television the launching of MTV, the attempted assassination of Ronald Reagan, the Challenger disaster, and the First Gulf War, the contemporaneity of television – the "nowness" of television – was not as pronounced as it must have been during the 1960s and 1970s. Andy Warhol recalls wandering into Macy's department store in 1955 after an appointment with a psychiatrist and purchasing his first television. "I brought it home," he writes, "to the apartment I was living alone […] and right away I forgot all about the psychiatrist" (Warhol 1975, 24). As he reflects: "So in the late 1950s I started an affair with my

DOI: 10.4324/9781003312284-6

television set which has continued to the present, when I play around in my bedroom with as many as four at a time." (Warhol 1975, 26)

*

In Jean-Philippe Toussaint's novel *Television*, the protagonist of this understated narrative, an academic residing in Berlin, summarily decides one day after watching the final stage of the Tour de France to stop all together with television. As he reflects: "I gave it up in one fell swoop, once and for all, never to watch another show, not even sports." Ostensibly on sabbatical to write a monograph on the painter Tiziano Vecellio, more commonly known as Titian, the initials of this illustrious painter telegraph the fundamental conflict between two regimes of images – paintings and television – that unfolds across the pages of Toussaint's novel in the struggle of its protagonist to maintain his television abstinence and scholarly ascetism. *Television* deliberately lacks dramatic quality. Nothing much seems to occur other than the protagonists' failed resolution; what is here more idiotic, the decision to give up once and for all television or inability to do so? The conflict between the painted image and the pixelated image brings into focus the transformation of "the image" in a saturated age of television (and more broadly, electronic screens). As Vilém Flusser argues, the term "image" possesses different meanings depending on its technological form of production, mode of circulation, and culture of viewership. Images on cave walls, images in a public square, and images on screens in a post-industrial society (Flusser's examples) – in each case, the image *is* different. Of images, one can speak and hence experience in many ways, given that images are produced, circulated, and viewed in different ways. An image is a technological *and* social artefact. In a post-industrial society, the image is, for Flusser, a "disembodied surface." This disembodied character contrasts with the embodied nature of paintings. Citing Paul Valèry's statement, Merleau-Ponty's remark that "the painter takes his body with him" underlines the constitutive interaction, or "circuit," between the vision of painting and movements of the lived body. With television images, the surface is disembodied in both senses: the image is not a material object as with painted surfaces; the image is not created through gestures of the lived body as with paintings. Along with this transformation in the regime of images, a transformation in the relation to time and, most importantly, history, occurs as well in the tele-vised age, as illustrated with the protagonist's failure to write his study of Titian. As Flusser writes:

> This radical change in the meaning of the word *history* has not yet become obvious, for a simple reason: we have not yet learned how to read films and TV programs. We still read them as if they were

> written lines and fail to grasp their inherent surface quality. But this situation will change in the very near future. […] for those who think in written lines, the term means the possibility of acting upon history from *within* history. For those who think in films, however, it will mean the possibility of acting upon history from *without*.
>
> (Flusser 2004, 24)

In Toussaint's novel, the conflict between painting and television, as regimes of visual literacy, mirrors a tension between watching television and reading books. Television is not just a visual medium; it is a literary medium. Images on a television screen are not merely cinematic in terms of motion; they move in narrative time as well. As a form of literacy, in this broad sense, watching television competes with reading books. Who has not succumbed to television pulling us away from reading a book, perhaps, in fact, Toussaint's *Television*, which, given its ambling narrative, might easily lose our attention when pitted against the voracious vivacity of television? A society steeped in television, as Ray Bradbury already imagined in his 1953 novel *Fahrenheit 451*, is prone to indifference and, indeed, outright hostility toward reading, to the point of banning and burning books. As the aging Professor Faber in this novel formulates the *status quo*: "It's not books you need; it's some of the things that once were in books. The same things could be in the 'parlor families' today." In Bradbury's dystopian world, the reading of books is banned in the name of promoting happiness among non-divisive citizens. The political arrangement of deeming books illegal (and assigning firefighters the mission of burning books) stems from the technological form of television and its radically equalitarian unilateralism in Bradbury's novel. Book cultures are said by the authorities in the novel to promote different views and contested viewpoints; a world of freely circulating books, without censorship or control, is inherently democratic in an agonistic sense. A world without books is a world controlled by media rather than set free by media. Books are likewise fictional and hence, as with Plato's notorious argument against the poets, false, whereas television delivers "the real" in a unilaterally virtual form. In the parlance of *Fahrenheit 451*, the term "parlor family" refers to large television screens that have replaced, as the center of social life and consciousness, the traditional nuclear family and parlor room. As Mildred Montag – the sleeping-pill addict and shallow wife of the protagonist Guy Montag – exclaims in reference to her favorite television shows: "That's my family."

Despite his initial resolution, the protagonist of Toussaint's novel repeatedly finds it impossible to escape not merely watching television but *thinking* about television. Even when the television screen lies dormant or is physically absent, it nonetheless keeps us awake in its

virtual, ghostly presence. As he muses, "even when you're not watching, you think that you are missing something on TV." There is no escape from a society of saturated televised spectacle. One evening, Toussaint's protagonist is standing alone at the window of his apartment, looking out onto the streets, when he notices how many, if not all, of his neighbors in surrounding buildings are watching television. Although the television in his own apartment remains unlit, he begins reading, distractedly, the television listings. Whether watching television, watching other people watch television, or standing in front of a darkened television set, television is inseparable from his form of life. Television is more than an addictive pastime or habit, however. There is something inevitable about the compulsion to watch television; one is always "on" to television.

Toussaint's musings on television can be read as outlining a phenomenology of television experience based on the obsessive compulsion to always want to watch television and its neutralizing effect on consciousness. The virtuality of television – that something is always happening on television – can be said to "virtualize" consciousness in the sense of simultaneously robbing consciousness of other possibilities (other than wanting to watch television) and neutralizing consciousness of its own actuality. As he notes:

> One of the principal characteristics of a turned-on television is that it artificially keeps us in a state of continual alertness, bombarding us with an endless stream of signals, all sorts of little stimuli, visual and aural, whose goal is to arouse our attention and keep our minds watchful.
>
> (Toussaint 2004, 12)

Our minds are kept watchful – on the look-out for television – because television is always looking over us. Our days start with television, our nights end with television. As Herbert Simon influentially argued, television consumes attention. The endless stream of programs and commercials impoverishes attention by captivatingly economizing attention; snatched attention becomes transformed into a commodity of economic capital (as with advertisements) or social capital (as with television celebrities). There is always time for television; television is "timely," in the German sense of *zeitlich*, always "at the right, or opportune, time." The eternal nowness of television is, moreover, expansive in space and time. Television signals, albeit at ever reduced amplitudes, propagate into outer space, thus extending the stream of televised images beyond our planet. If aliens ever make first contact with us, it could well be through the reception of television signals. In Carl Sagan's science-fiction novel *Contact*, a 1936 television

transmission of the Berlin Olympic Games in Nazis Germany is received by an extra-terrestrial form of life, who bounce the signal back to Earth with encoded instructions on how to create a space-portal through which we human beings could reach them.

The protagonist of Toussaint's novel instinctually turns his attention to watching television; television has implanted itself firmly into his form of life. "As if caught up in some sordid intoxication," he reflects, "I'd taken to turning on the TV in the evening and watching everything there was to see, my mind perfectly empty, never choosing any particular program, simply watching everything that came my way, the movement, the glimmering lights, the variety" (Toussaint 2004, 9). This force of existential addiction consists in the promise of a pleasure matched by the restless porousness of streaming images, absorbing and consuming consciousness to no end. The "variety," "movement," and "glimmering lights of television" contrast with the intellectual idleness and physical inertness of our protagonist. The endless list of programs on television, which the protagonist at times itemizes in the novel, over-stimulates *and* anesthetizes his consciousness, much as Georg Simmel, in his seminal essay "Metropolis and Modern Life," characterized the condition of modern city-life at the beginning of the 20th century. Television induces a frenetic boredom of captivated indifference. As Toussaint writes: "Increasingly indifferent to the images it [*the human mind*] receives, it soon ceases to react to all when new signals are sent its way, knowing that to react would only mean once again falling prey to television's deceptions" (Toussaint 2004, 12). Screened deception is the lure of televisions' simulacrum: "television offers the spectacle not of reality, although it has all the appearance of reality (on a smaller scale) […] but rather its representation." This flattened representation in two dimensions is nevertheless a more "refined and much more indirect sort of representation than painters used to create an image of reality in their works." The televised image "seems at first glance more trustworthy, authentic, and credible" than life itself. During the loneliest of moments, we can always find something to watch and hence find virtual companionship with television. Television channels proximity, if not intimacy: it makes things present to us in sight and sound, in motion and color. And yet, this broadcasting of nearness belies an unbridgeable distance. As he writes: "Because television is not only fluid, never leaving our thoughts time to blossom in its perpetual race forward; it's also impermeable, in that it forbids any exchange of wealth between our minds and its matter" (Toussaint 2004, 12). What Toussaint has in mind is how with either reading or looking at a painting, a reader or a viewer can pause to inspect or reflect more closely on the features of an image or aspects in a text. Consciousness can detach itself and adopt a self-reflective stance by interrupting its absorption in either a painting or a book. Consciousness retains the power

of interruption, and thus break from the immediacy of the painted image or written word. Phenomenologically speaking, the temporalization of streaming televised images insinuates itself into consciousness with the effect that the temporalization of consciousness becomes constituted, that is, mediated, by the media of television; the "time-object" of televised image *inverts* the constitutive relation between subjectivity and objectivity, as if, the "time-object" subjected time-consciousness to its own streaming form.

In *Television*, our would-be author of a book on Titian repeatedly fails to avoid television. Nearing the end of the novel, our protagonist sits once again in his study. Finally, back to work, he still distantly hears the "muted music of *Robin Hood*" to his right from his son in the next room and, to his left, muffled sounds from his wife's television viewing in the bedroom, only to realize as well that he "was hearing another murmur in the study, still more muffled, coming neither from the living room nor from the bedroom" – from a television of his upstairs neighbors! Admitting defeat, "this was too much," he declares to himself. He leaves his study to put his son to bed after watching with him the end of *Robin Hood* while distractedly leafing through a book; he then retires to his bedroom, where there again, in a futile effort to read a book, he continually casts a "careless glance toward the television screen," his wife's program (Toussaint 2004, 162–163). "The moral" of his story is that "since I'd stopped watching television, we had two TVs." His wife enjoys falling asleep while watching television in their bedroom; his son enjoys television in the living room. As he comes to understand:

> Television is formal beyond all reason, I now told myself […] twenty-four hours a day, it seems to flow along hand in hand with time itself, aping its passage in a crude parody where no moment lasts and everything soon disappears, to the point where you might sometimes wonder where all those images go once they've been broadcast, with no one watching them or remembering them or retaining them, scarcely seen at all, only momentarily skimmed by the viewer's gaze. (Toussaint 2004, 95)

Whether entertained or bored, whether on or off, *there is* television.

*

The rarity of philosophical reflection on television is conspicuous given the pervasiveness of televisions and, more generally, digital screens, in contemporary life. A few exceptions notwithstanding, television remains invisible to the gaze of philosophers. One notable exception is Günther Anders' treatment of television (and radio) in his regrettably neglected *Die*

Antiquiertheit des Menschen. Anders' philosophical critique of television occurs as the second chapter – "The World as Phantom and Matrix" – in his magistral work. Anders' reflections were formulated during his exile in the United States between 1936 and 1950 where he arguably found himself perfectly situated to witness first-hand the post-Second World War television boom in America and eclipse of the Golden Age of radio by the rapid ascendency of household television sets. In 1946, 7,000 television sets were purchased nationwide. In 1950, five million sets were sold, or in 20% of American households, and by 1960, 90%. Although based on his experience in the United States nearly seventy-five years ago, Anders' assessment of television has not lost its relevancy, despite significant developments in the expansion, format, and reach of television during the later decades of the 20th century, as well as the end of television's media dominance at the turn of the 21st century with the digital age. Most importantly, two keystone concepts in Anders' overarching framework – "the phantom" and "the matrix" – can be seen as forerunners to the postmodern conception of "simulacrum" in the writings of Jean Baudrillard.

As Anders states in *Die Antiquiertheit des Menschen*:

> This third revolution [*post-industrial*] is the real subject of our investigation, since it is almost exclusively devoted to unique changes that are inflicted on man as a being who is supplied with a world, and to the no less unique consequences entailed by this supply of the world for the concept of the world and for the world itself.
>
> (Anders 2018, 5)

In his self-styled "negative anthropology," centered on the claim that human freedom is grounded on "non-fixity" (*Nicht-Fixiertheit*) and "unboundedness or undeterminedness" (*Unfestgestelltheit*), Anders argues that human existence falls short of its own technological productions. *Homo faber* does not stand up to the task or measure of its "own" technological achievements. As Anders writes, "that we are not up to the perfection of our products; that we can produce more than we can imagine and be responsible for," represents a gap between the human and the technological. In his own words: "The fact of the daily growing out-of-synchronization of people with their world of products, the fact of the distance becoming wider from day to day, is what we call 'the Promethean Gap' [*das prometheische Gefälle*].'" (Anders 2018, 16). This Promethean Gap forms an "a-synchronic" relationship between the human, the maker, and technologies, as made, but also embraces, anthropologically, a rift between "doing" and "representing," "acting" and "feeling," "knowledge" and "conscience," and "the produced appliance" (*produzierte Gerät*) and human embodiment.[1] Modern

technology produces a reversal of agency between fabricating human being and fabricated tool. The ostensible purpose of technical artefacts is to facilitate human agency in the world, in terms of which human beings seek to materialize their projects, and hence, actualize their freedom; in effect, however, artefacts of our making gain agency over us. The modern fabricated world is no longer a world at the center of which we stand; rather, we stand exposed at the margins of the world of our making, and hence, unmaking. Anders speaks in this regard of our "Promethean Shame." Faced with the inversion of maker and made, whereby "frei sind die Dinge: unfrei ist der Mensch," human agency becomes neutered and subjected to the agency of technological artefacts, or, in Anders' terms, "technological apparatus" (Anders 2018, 33). The human is no longer the measure of things; (fabricated) things provide the measure of the human. Given this "triumph of the apparatus-world," the difference between the technological and the social vanishes; the social becomes encrusted around the armature of the technological. The Promethean Gap between human beings and "their" technological apparatus represents a phenomenological gap as well, or what can be called a "de-phenomenologization," or, in this sense, a "virtualization," of technological apparatus within the lifeworld. As a kind of "anti-phenomenological" principle, technological apparatus "do not show or manifest themselves" in the sense that *how* such technologies shape our existence eludes us in their functioning. We are rendered blind to the effects of technology by the very brilliance of their accomplishment. This is what Anders suggestively calls "apocalyptic blindness."

As a self-styled philosophical "occasionalist," with his hybridization of "journalism" and "metaphysics," Anders' eschews a systematic or exhaustive treatment of television, and sketches instead the contours of what commonly remains "inapparent," namely, *the apparatus* of television as opposed to the (visible) artefact of television (i.e., this technical thing in my living-room). The "thingness," as it were, of the technological thing ordinarily remains "inapparent" in our apparently innocuous dealings. Indeed, what is "inapparent" about technological apparatus (as with television) is the extent to which, and how, technologies have essentially transformed, in fact, over-turned the Platonic view of *technê* as a tool or an instrument, as a means toward a prescribed end. The seamless integration of television into the routines and rhythms of everyday life obscures their super-liminal and non-instrumental function. The technical workings of a television are (for most) users intellectually obscure as with most technical gadgets around us. Aside from technical ignorance, *how* the technological apparatus of television imposes ontological conditions and constraints on our form of life, cultural values, and inter-subjective relations – in short: our being in the world

and being with others, indeed, our being with oneself – remains inapparent. Hence, the imperative of a philosophical critique of technology which, in the form developed by Anders, reveals the technological as a historical-ontological a priori of being-in-the-world.

Anders is not alone in stressing that television, as with other technological apparatus, transforms the conventional meaning of instrumentality. On this point, Stuart Hall formulates a comparable view: television cannot be reduced or merely understood as a "medium" in terms of its technical features (here explicitly rejecting Marshall McLuhan's distinction between "hot" and "cold" media). As Anders argues in the same vein, television is more than a medium; it is a technological apparatus, not merely a technical artefact. Technical "givens in any medium," as Hall writes, is the "product of an advanced electronic technology," and such technical "givens" determine the function of television (color vs. black and white images, etc.). As Hall elaborates, "technical properties of a medium [at a particular stage in its historical evolution] *do* impose *certain* constraints on its use" and by the same token impose conditions on experience, for example, how one views televised images and, more broadly, how one perceives and interacts with the world. One needs to distinguish between technical functioning and technological functioning, the latter embracing social relations and praxis. As Hall states: "A social aesthetic [*of television*] depends, ultimately, not on the level of the 'hardware' but on the *human uses of the hardware*: that is, on the form of its social appropriation, embedded in the different levels of social praxis" (Hall 2021, 213). Anders advances a more encompassing and penetrant claim: the technical artefact obscures the technological apparatus, where the latter determines at once behind our backs and in front of our eyes, as it were, forms of praxis, social relations, and armatures of cultural values.

This is not to deny that we employ television for the intended purpose of entertainment and information. As a technological means of broadcasting images across space and time, television is an instrument of communication. For Anders, however, television is not merely a means of communication; it is an apparatus of *consumption*, and thus, in varied ways, inseparable from the consumer culture of post-industrial capitalism or, to borrow Frederic Jameson's term, Late Capitalism. Television does not only consume attention; its voraciousness consumes "subjects" or, better: subjectifies the subject. Take the example of a televised church service broadcast. As Anders observes, what impacts viewers is not merely that they are participating in the church service but that they are consuming *an image* of the church service. Television images are not simply viewed by us; we are consumed by images. It is not therefore that television primarily allows viewers to participate in an event which they cannot attend in

person. According to Anders, televised images reshape the subject into its own image; television participates *in us*. As Anders writes:

> what shapes and misshapes us, what forms and deforms us, are not the communicated (or mediated) objects through 'means' [tv], but the means itself, the apparatus itself […] which, due to its fixed structure and function […] [defines] the style of our interaction and our life, in short, *us*.
>
> (Anders 2018, 120)

Television "fixes" the subject in the formal meaning of the term (i.e., determines and binds the subject) as well as the informal meaning of "addiction" (i.e., needing one's fix of a drug), as illustrated, for example, in Toussaint's *Television*. In both senses, technological apparatus displaces human freedom, which Anders understands as "unboundedness" or "non-fixity." In Toussaint's novel, the protagonist always finds himself fixated on television, bound to its seductiveness and allure, and thus remains incapacitated (figuratively "de-capitated") to *create* intellectually, namely, to write a book, which, on Anders' account, embodies, in creativity, the "non-fixity" of human freedom. What remains "inapparent" in the pretense of freedom on which so much of modern technology is consumed is the estrangement of human freedom. Indeed, the promotion of individual freedom goes hand in hand with a reductionist view of technology as a tool or instrument at our disposal and mastery.

The restricted notion of television media as "a means," in conjunction with the equally inadequate assumption of human mastery, and hence, freedom, over technology, connects to Anders' additional claim that television, as a discrete phenomenon, cannot be abstracted, as an apparatus, from the total arrangement of social life. A television prominently in the living room is a discrete, visible phenomenon; the images on the screen are a visible, discrete phenomenon. What is inapparent, however, is the television *apparatus* in its encompassing multi-dimensionality: television corporations, advertisement industry, and, in the European context, the nationalization of television. The apparatus of television must be set within its relation to "totality," or "the whole," namely, the entire field of social existence and cultural life. As depicted in Don DeLillo's *White Noise*, television is both discretely located and unobtrusively pervasive through-out American households. The typical suburban American household is awash in a background noise of television: "The TV said: Now we will put the little feelers on the butterfly" (DeLillo 2011, 420). Throughout DeLillo's novel, snippets of television commercials and programs waft through the airwaves of daily conversations among family members and

are seamlessly integrated into the running prose of the novel. Television is a discrete phenomenon: it is the screen around which the family in the novel sits on Friday nights together and the different programs watched by various family members; it is a diffuse phenomenon that radiates throughout the household in *White Noise*, always there, always on ("waves and radiation" is the title of the first section in DeLillo's novel.)

During the decades after the Second World War, television entered American society through the gateway of the household and rapidly established itself as the nucleus around which both the ideal and reality of the nuclear family took hold and shape. The living room became transformed into the locus of deployment for the television apparatus and was lauded as bringing the world directly into the home. For Anders, the introduction of television into the domestic hearth transformed family existence and provided the site for the reconfiguration of modern subjectivity. The relationship of the private and the public became transfigured as well. As Anders writes:

> And it does in fact dissolve it, since what dominates the home, thanks to television, is the broadcast of the *outside world*—real or fictional; and it dominates the home in such an unlimited manner that it invalidates and renders phantasmagorical the reality of the home, not only that of the four walls and the furniture, but also of the shared family life itself. When that which is remote becomes familiar, the familiar becomes remote or disappears […] Nowadays, the real home has been demoted to the status of a *container* and its function is reduced to containing the video screen for the outside world.
>
> (Anders 2018, 124)

As Anders qualifies:

> This does not mean, however, that the television has become the center of the family; to the contrary, what the television set reproduces and embodies is precisely the decentralization of the family, its ex-centricity, because it is the *negative family table*. It does not provide a common *center point*, but rather a common avenue of escape for the members of the family.
>
> (Anders 2018, 124)

These lines of escape promote an ideal (and illusion) of individual freedom and the ideal of the nuclear (suburban) family, as famously exemplified by 1950s series such as *Leave it to Beaver*. As Anders argues, television was celebrated as heralding a "renaissance of family and

Privatheit." On the contrary, however, television and its promoted and televised ideal of the nuclear family provoked a dissolution of extended family structures at the expense of amplified individualism within families. According to Anders, the traditional family symbol was the table in the middle of the living room or dining room, around which family and neighbors came together. The family table can, moreover, be understood as figuration of the political as such. The space of the political could be likened to a "table" at which all participants sit and appear to each other, where narratives are told and exchanged, where reasons are given, and where, in this space of communicative reason and action, there reigns the ideal of consensus. We are all at the same table, despite our differences, and we bring ourselves to the table, as the space of what is common in-between us. The dining table has become supplanted by the television set, which represents for Anders the "de-centralization" and "ex-centric" negative family table. The television establishes a virtual public space. The TV is not a common centerpoint, but a common point of escape; it is centrifugal in motion in sending family members outwards *and* inwards. Rather than a shared public and common space, the television space can be considered a public "non-place," as everywhere and nowhere, where individuals partake in a collective experience, or common viewing, without encountering and communicating with each other directly.[2] Strictly speaking, the "non-place" of television is neither utopic nor dystopic. Or rather, it can seem to be both.

In *Fahrenheit 451*, Bradbury's dystopic world is centered around the television apparatus. The "parlor family" of television displaces the family and parlor room: "The living room; what a good job of labeling that was now. No matter when he came in, the walls were always talking to Mildred" (Bradbury 2013, 42). Mildred obsessively watches "television" ("wall-to-wall circuit"). The three walls have screens, and she wants a fourth wall-screen. In one of the programs, she receives a script at home for a character in the show– in this case, "a homemaker" (Bradbury 2013, 16). When asked by her husband, Montag, what this play is about, Mildred offers a banal response, and exclaims how this is so much fun: "I just told you. There are these people named Bob and Ruth and Helen." "Oh," Montag responds. The wall-to-wall screens are the locus and form of social interaction. The character Clarisse in the novel rejects this fixity of social life to television. As she observes to Montag:

> Being with people is nice. But I don't think it's social to get a bunch of people together and then not let them talk, do you? An hour of TV class, an hour of basketball or baseball […] but do you know,

> we never ask questions […] they just run the answers at you, bing, bing, bing, and us sitting there for four more hours of film-teacher. That's not social to me at all. It's a lot of funnels and a lot of water poured down the spout and out the bottom, and them telling us it's wine when it's not.
>
> (Bradbury 2013, 27)

Television is seductive as the supplement for the real that becomes itself a *phantomhaft* reality, or simulated reality in contrast to the "fictional" and "colorless" phenomenon of books and "reality" itself.

> Mildred kicked at a book. 'Books aren't people. You read and I look all around, but there isn't anybody!' He stared at the parlor that was dead and gray as the waters of an ocean that might teem with life if they switched on the electronic sun. 'Now,' said Mildred, 'my 'family' is people. They tell me things; I laugh, they laugh! And the colors!
>
> (Bradbury 2013, 69)

It is striking how Bradbury's imagining of Mildred's "reality TV" can be seen as anticipating the development of Reality TV. Jean Baudrillard echoes Anders' view that television has become essential in the transformation of communication in a post-modern world. As Baudrillard writes, the "message of television is not the images or in the images it broadcasts or transmits, but with new modes of relation and perceiving – consuming – images, and this in turn changes family structure and inter-subjective relations" (Baudrillard 1998, 98). Television along with other screens and televisual surfaces are neither mirrors, representations, nor windows onto an external reality. In Baudrillard's terms, these technological artefacts produce a "hyperreality" and "simulacrum" in the same vein as Anders argues that "reality is transformed into the reproduction of its images" (Anders 2018, 179). As Anders states: "Wenn das Phantom wirklich wird, wird das Wirkliche phantomhaft." Reality TV for Baudrillard – a phenomenon that did not exist for Anders – as he discusses in *Simulacra and Simulation* "hyper-reals" or hypes what is banal; the ordinary, mundane, petty, and unremarkable "dramas" of average individuals captivates and fascinates audiences through their "hyper-realization" on screen. As with Mildred's experience in *Fahrenheit 451*, contemporary Reality TV breaks through the fourth wall separating screen and viewer. Audiences become invested in and intimate with the drama of Reality TV. Stated in Baudrillard's terms, Reality TV effaces the distance between screen and audience, between "event" and "witness," or spectator. As Baudrillard states, with Reality TV, "the viewer has to be brought not in front of the screen (he has always been there, and that is indeed his alibi and his refuge) but into the

screen, taken to the other side of the information set" (Baudrillard 2008, 67). The injecting of audiences into the screen produces a surplus of "reality" that undermines the distinction and distance between television and reality, that is, hyperreality. As Baudrillard writes: "there is no separation any longer, no empty space, no absence; you enter the screen and the visual image unhindered. You enter your life as you would walk on to a screen" (Baudrillard 2014, 177). This "walking onto a screen" can be achieved with audience participation through voting competitions. The interaction between television "event" and audience changes; the audience becomes involved and invested, and in this sense, immersed; and yet as Baudrillard argues, the audience does not gain in agency but becomes disenfranchised of agency. TV produces hyperreality in the sense of an over-abundance of information and hyper visibility that forecloses the possibility of critical reflection and, as represented in *Fahrenheit 451*, genuine dialogical encounter among individuals. This surplus of reality at the expense of the real reflects in turn what Anders identifies as "Icon-mania" – an obsession with images and making visible – that defines late modern capitalism; a desire for images that goes hand in hand with a commodification of images, or, in other words, spectacle as the dominant form of commodification. Baudrillard compares this manufacturing of experience with Marcel Duchamps' "ready-mades." Reality TV is a "ready-made" reality in which audiences play an integral role in sustaining the simulation, or, as Anders calls it, "phantom," of reality. On screen Reality TV, or, for Anders, television apparatus *tout court*, renders reality *phantomhaft* – it "hyper-realizes," as it were, our *sense* of reality.

For his part, Anders argues in a comparable light. With film, Anders follows Siegfried Kracauer's analysis that the collective viewing of a mass-produced consumer object fashions, among the viewing public, a mass-produced consumer of images. With television and radio, however, though we have, as with film, a mass-produced object (*Massenware*), we do not have the comparable consumption by a mass audience. The Smiths and the Millers consume television either as a family or alone; as Anders writes: "*je einsamer sie waren, um so ausgiebiger*." In contrast to the collectivization of an anonymous crowd through the power of film, as argued respectively by Kracauer and Walter Benjamin in his celebrated essay "The Work of Art in the Age of Mechanical Reproduction," television produces what Anders dubs the "mass-hermit." The mass-hermit is the solidary subject who consumes mass disseminated and produced images; the transformation into "mass human being" through the consumption of mass-images, that is, through leisure. Anders draws the implication that we have a new type of "mass or collective human," one that is not produced through mass-gatherings or through the "mass-regime" or "orchestration." Instead, with this fashioning of the "mass-hermit" – one could see Anders' description

of television as formulating in a rudimentary manner an insight into the internet as an echo chamber promoting isolation among individuals and molecular groups – the individual is not absorbed into a collective nor does she act out "mechanically" in conformity with others. One could in this respect speak of the "molecularization" effect of television; indeed, of screens as such, as both an apparatus of absorption and isolation. To produce someone into a "nobody," as Anders writes, it is no longer necessary to "drown him into a mass flood." The consumption of television images fabricates "the illusion of freedom of personality and the right of individuality." Rather than, as Kraucauer argued with cinema, rendering the individual into a figure within a "mass ornament," the television apparatus sustains the simulation of individualized freedom, as enacted in the apparent freedom of choice to watch what *I* want. As Anders remarks:

> Since the broadcast is a commodity, it too must be served in status that is pleasant to our eyes and ears, optimally ready for consumption, alienated, removed from its core, and assimilated; that is, in a manner that it addresses us as our *simile*, custom-designed, as if it were *part of us*.
> (Anders 2018, 122)

Centered on this commodification of images and production of the commodity-subject, Anders further argues that the effects of the television apparatus includes the "banalization" (as exemplified with contemporary Reality TV), unilateralism (in the relation between commanding televised image and receptive audience, as illustrated in *Television* and *Fahrenheit 451*), and the polarity between sensationalism and anti-sensationalism. A pervasive example of the latter media effect is, on the one hand, the hyper-sensationalized coverage of natural disasters (with reporters "live on the scene" engulfed themselves by threatening winds, ocean surges, etc.) and, on the other hand, the routine "entertainment" value of such "live coverage" of disasters.

*

Anders argues that the "key phenomenon" (*Schlüsselphänomenen*) is here the cultivation of what he calls "Icon-mania" (*Ikonmanie*), which he understands as a dominating addiction to images (*herrschende Bildsucht*) that characterizes human beings. This age of television and, more generally, screens, represents an age of "postliterarischen Analphabetentums" gathered around a host of media and technologies of (image) reproduction, which he further describes as "Verdummungsgeräten" and "herausgerissene Weltfetzen" ("shredded pieces of the world"). This condition of "postliterarischen Analphabetentums" is portrayed in both *Fahrenheit 451* and *White Noise*. As Jack Gladney in Don DeLillos' novel observes: "A little later I watched Steffie in front of the TV set. She moved her lips,

attempting to match the words as they were spoken" (DeLillo 2011, 408). In *Fahrenheit 451*, Montag complains: "Nobody listens anymore. I can't talk to the walls because they're yelling at *me*. I can't talk to my wife; she listens to the *walls*" (Bradbury 2013, 78). As Anders states:

> *Since the television and the radio speak on our behalf, they also deprive us of our ability to speak*; they rob us of our capacity to express ourselves, of our opportunities for speech, and of our pleasure in speaking, just as the music of the phonograph has robbed us of the live music that we used to perform in our homes.
>
> (Anders 2018, 125)

Anders contends that television and radio "take away speaking" from its viewers and "robs" us of our capacity for expression. In his reflections, there is an suggestive literary vignette about two lovers who are walking with a "portable radio," where the radio broadcast functions as a "third." The radio, as "the third," mediates and frames the dialogical pairing of the lovers. The lesson for Anders is that the human, as the speaking animal, no longer speaks; she is spoken for, that is bespoken, through mediatic received speech. What happens to the public and political function of speech when television becomes "the parlor" or "common"? Anders bemoans the deterioration of language and regression to "infantile" speech. TV and radio contribute to the increasing "inarticulateness" of contemporary discourse. As he writes:

> And no matter how much they now have a guaranteed right to speak, they have been completely inoculated in their hearing, and have essentially ceased to be ζῶα λόγον ἔχοντα, just as, as eaters of bread, they have ceased to be *homines fabri*, since they do not give form to their verbal nourishment, nor do they bake their own bread. For them words are no longer something one speaks, but something one merely hears; speaking is no longer something that one does, but something that one receives. It is clear that they therefore "have" the logos in a completely different sense than is conveyed by Aristotle's definition; and it is just as clear that they are thus transformed—in the etymological sense of the term—into *infantile* beings, that is, into minors, those who do not speak. No matter in what cultural or political milieu this process towards the condition of ἄνευ λόγου εἶναι [an existence without speech] takes place, its end result is always the same: a type of man who, because he no longer speaks himself, no longer has anything to say; and who, because he only listens—and this is more and more the case—is a *subordinate*.
>
> (Anders 2018, 126)

The acculturation of television is inseparable from the development and promotion of consumer society and, especially for Anders, the production of consumer subjects, i.e., "the consumer." This connection between television and consumer culture finds its locus with television commercials and the television as primarily a medium of advertisement. In 1964–1965, Warhol co-directed (and never completed) a film entitled *Soap Opera: The Lester Persky Story* which combined footage from actual television commercials (made by Persky) with footage (silent) of banal scenes shot by Warhol at the Factory.[3] Are the advertisements commenting on or promoting Warhol's tableaus, or are the tableau's commenting on or promoting commercials? (Spigel 2010) This blending of commerce and art, of commercials and programming, illustrates Anders' contention that television images are not merely consumed; the television apparatus is integral to the constitution of consumer society and mediatic production of consumers. The primary purpose of television programs was to capture the attention and shape and mobilize the desires of consumers. Anders argues for the function of labor in the form of commercials which form an integral part of television apparatus and culture. The consumption of commercials is a kind of labor. Television commercials are "camouflaged labor." The attention of viewers becomes transformed into a commodity, and this is what advertisers are paying for, and in this form, the agency of the consumer becomes "marketing the product" to themselves. The consumer subject becomes "sold" in this sense and transformed into what Anders labels "homeworker." As he writes:

> Everyone is in a way employed and occupied as a *homeworker.* […] The process becomes completely paradoxical insofar as the homeworker has to pay for his labour instead of being paid for his cooperation; he namely has to pay for the means of production (the device and, in many countries, even for the broadcasts) […] So he pays for selling himself.
>
> (Anders 2018, 103)

The home is not the refuge of privacy, for if television "delivers" the external world into the home, in turn, the home becomes "delivered" to the world. As Anders states: "not only is it true that 'The world is delivered to your home,' but also: 'Your home is delivered to the world'" – namely, the world of consumer produce. The relation between programs and advertisements is inverted: we watch programs in order to watch advertisements, which in turn seed, as it were, our desires. What is presented to us – showcased and displayed on screen – are "pre-packaged objects," whose pretension is, together, to be the world, the purpose of which is to shape us according to its image – "economy ontology"

(*Wirtschaftontologie*). The world is not a finished world, and no "true" world is there for the world only becomes true and truthful when it is put into circulation through images and thus disappears as a world beyond images. Overall, this mediatic transformation of being-in-the-world occurs represents the "banalization of the world" through television. Difficult to translated into English, the German term *Verbiederung* in Ander's expression *Verbiederung der Welt* means "making simple minded," "conventional," and "homely." Anders likewise speaks of "magical power of banalization" whereby we become "chums" or "pals" with the world. We feel, for example, a certain intimacy with Hollywood Stars through tabloid reports about their sexual lives and family intrigues. "Banalization" produces subtle camouflage of alienation. As Anders states: "Despite the fact that we really live in an alienated world, the world is presented to us as if it is a world for us, as if it were our own and like ourselves." (Anders 2018, 116)

Broadcasting results in *neutralisierender Verbiederung* that neutralizes politics and critique. As Anders writes: "The fundamental neutralizer is […] the commodity character of all phenomena" (Anders 2018, 121), the most significant of which takes the form of what Anders calls "scrambled judgement" (*verbrämbte Urteile*). Radio and television are characterized by a constitutive ambiguity that places, or sends, the viewer in a situation wherein the distinction between immediacy and mediation, experience and news, becomes effaced. Media allows for the transportability of the events, but this transportability or tele-communication contaminates things themselves. Anders gives the example of a political candidate "Smith" appearing on television, who is presented to a viewing audience, potential voters. Candidate Smith looks to present himself as a pleasant person and just another "guy" with whom you would want to have beer. These television attributes of Smith – pleasantness, chumliness – are seen as the exclusive traits of his personality. Candidate Smith becomes all surface; the surface screen becomes Smith, such that the "predicate exhausts the subject," and we only see "pleasant Smith," nothing other, nothing more. This is how a "lying" is produced as the persona, or avatar, "Smith" that scrambles the "truth" about the candidate Smith. In his series of condensed novels, *Atrocity Exhibition*, JG Ballard explored this political effect in "Why I Want to Fuck Ronald Reagan." As Ballard notes:

> In his commercials Reagan used the smooth, teleprompter-perfect tones of the TV auto-salesman to project a political message that was absolutely the reverse of bland and reassuring. A complete discontinuity existed between Reagan's manner and body language, on the one hand, and his scarily simplistic far-right message on the other.

> Above all, it struck me that Reagan was the first politician to exploit the fact that his TV audience would not be listening too closely, if at all, to what he was saying, and indeed might well assume from his manner and presentation that he was saying the exact opposite of the words actually emerging from his mouth.
>
> (Ballard, 1990, 165.)

In a contemporary context, Donald Trump exemplifies what Anders calls "scrambled judgement." Those who are devoted to Trump can only see Trump for mediatic "predicates" that blind them to the reality of the subject, namely, Trumph himself. In order to persuade the consumer of mediatic personalities that she is not being persuaded ("talked into something," or *einreden*), the judgment embedded in the image must renounce its form as a judgment, namely, S is p, and simply unilaterally declare "p" without S. Smith is nothing more than "pleasing" albeit only in scrambling of discernment that forecloses any transparency of the structure of judgment as well as the true nature of Smith, or Trump, as the case may be. This scrambling of judgement in the mediatic image operates negatively as well, for example, with Trump's mantra "crooked Hilary." This is the salient sense in which "phantom" – the phantom of Trump, or the phantom of "crooked Hilary"– becomes a "commodity" (*Ware*) in the form of a "camouflaged judgement." As such, the commodity-phantom always enacts its own judgment and self-praise; the phantom is a judgment but a scrambled judgment that suspends or neutralizes the consumer of images' own power of judgment. The audience falls under the spell of scrambled judgments, or magic words, against which one is powerless, and which, by dint of substitution, becomes one's own speech, as "reality." In this regard, Anders can be seen as relating the power of television to the advent of post-ideological regimes. Anders here abandons a classic Marxist view of ideology as false consciousness; we have instead mediatic consciousness, or "scrambled" consciousness. Not just false consciousness but what Anders calls "false actions." As he writes: "the immediate goal [*of ideology*] consists in producing 'false consciousness'" and in this sense television – as media – is an "apparatus for the production of *false action*" or better: no action. As he writes as well: "The truth of ideology (that is: true compliance with the interest that lies at the basis of its production) is false praxis."[4]

In this light, Anders did not espouse any optimism for television in the way formulated by Raymond Williams in *Television: Technology and Cultural Form*. Although on the one hand, Williams argued that television communication operated in the mode "flow," where different programs "flow" from and into each other, or from a commercial to the program (as with Warhol's *Soap Opera*). This perpetual flow of visual

information inhibits the ability of viewers (as with the protagonist in *Television*) to interrupt the stream of images and remove themselves from the captivation of television. On the other hand, Williams held out the possibility for a critical engagement with television through a reflective self-scrutiny on the part of viewers. For his part, Anders only occasionally acknowledged examples of televised programs that provided awareness and moral awakening with his references to news-reporting about the Vietnam War and television series about the Holocaust. As he writes: "Die täglich in die amerikanischen Heime kanalisierten Bilder vom vietnamesischen Kriegsschauplatz haben Millionen von Bürgern die auf die Mattscheibe starrenden Augen erst wirklich ‚geöffnet' und einen Protest ausgelöst, der sehr erheblich beigetragen hat zum Abbruch des damaligen Genozids."

*

Anders argues that modern media fosters an encompassing attitude of idealism. With this designation, Anders is keen to distance himself from established conceptions of idealism in the history of modern philosophy; to this end, he provides a cursory overview of different forms of idealism (Schopenhauer, Fichte, etc.). For Anders, modern media (especially television) inaugurates a new form of idealism without direct philosophical equivalent – hence the philosophical challenge to think through the ontological effect of media in its transformation of our relationship to the world, indeed, of our relation to ourselves as human beings. Classic forms of philosophical idealism in one way or another argue that the world is essentially, as a meaningful world, only there *for us*, whether as "will and representation," as posited by the ego, or as conditioned by forms of intuition (space and time) and categories of the understanding, etc. One of the primary functions of media (radio and television) is that "the world is delivered to us at home" and, in this regard, events are "served up" or "driven up" to us. On the one hand, events broadcast on television are "present" insofar, when we view images of political demonstration, for example, we are, as it were, "there" on the scene, gripped by the events unfolding "live" before our eyes. On the other hand, the event on our television screen (or YouTube) is not "absent" insofar we ourselves are not there at the scene. We are, as it were, *on the scene and yet not at the scene*. What defines, among other aspects, standing or facing the world in the present is that this is a reciprocal relation: I can act on the world and the world can act on me, when I am standing at the scene, when I am genuinely *there*, or being present.

As brought home, the world is there"for me" on account of *my activity* but, on the contrary, on account of the apparatus of the media itself. Idealism becomes dislocated from the subject, and yet, is not located in

the object ("the world") either. Idealism is here disembodied of subjectivity along with a collapse of the stabilization of ontological distinction between "reality" and "image." Anders argues that what television delivers is neither an image or picture (*ein Bild*), in which case, the world is "absent," nor the reality of the world itself. We are neither "present" nor "absent" for what characterizes the peculiarity of the broadcast situation is its "ontological ambiguity" – it is at once "real" *and* "mere appearance," simultaneously "there and not there." As Anders notes, the term or notion "image" in the conventional sense no longer applies; we must re-think the meaning of "image." (Anders 2018, 151). On a conventional understanding of "image," there must be a *temporal* difference between the real and the image: the image is not "at the same time" as the depicted object. With televised images, this temporal distance is lacking between the image and the event – the image is the event. The event and the image are simultaneous and synchronized. This transforms as well as the meaning of "the present" (*Gegenwart*). The term can either mean "concrete present," that is, the situation in which we find ourselves, or "formal simultaneity," i.e., at the same time. With broadcasts (*Prinzip der Sendung*), that which is (nearly) simultaneous becomes delivered in such a manner that it "effects" as actually present. Formal simultaneity – that the image on the screen is at the same time as the event – gains the appearance of concrete present, that one (the viewer or spectator) is there – "live" – at the scene. This train of thought leads Anders to the statement: the more that something is made present (i.e., televised), the less present it effectively becomes. Alternatively: everything real becomes phantom-like (*phantomhaft*), everything fictive become real.

"Broadcast" or "program" (*Sendung*) is ontologically ambiguous, thus neutralizing the question whether it is "present" or "absent," "real" or "image." Anders also rejects the possibility that it would amount to an "aesthetic appearance" (or "illusion," *Schein*) or "as if." An intriguing implication of this ontological ambiguity finds expression in the thought that what becomes established with broadcast is "unserious seriousness" and "serious unseriousness," that is, an oscillation between both. The distinction between "fun" and "information" but also between *Sein* and *Schein* becomes destabilized, held in suspense. This "accidental phenomenological distinctiveness of broadcasts," namely, the ambiguity and undecidability between being and appearance, can be employed for "morally opportunistic effects" (Anders 2018, 172). The commercial broadcast is a phantom world that is half present and half absent. The stream of images by "showing the world, veils the world such that the difference between appearance and being would disappears; the event is shaped by the matrix of its reproduction" (Anders 2018, 111). "Broadcasts" are "lies" in the

complex and ambiguous sense described in terms of "phantom." Anders contends, however, that what becomes "prepared," in order to be sold, with each broadcast is not itself strictly speaking "the lie," for what becomes prepared is the "world-view as a whole" (*Weltbild als Ganzes*) such that "a whole type of human who becomes exclusively fed from phantoms and apparatus."

Anders proposes that this condition results from the englobing saturation of media, namely, the Matrix. The Matrix inverts Hegel's celebrated statement "The whole is the true" into "the Whole is the Lie." Anders underlines, however, that the whole here in question is not theoretical, but, as he qualifies, a "pragmatic world-picture," which should also not be understood as subjective worldview. More precisely – and this further specifies the unique kind of idealism in play with the Matrix – it is the "practical apparatus" of the Matrix that shapes our actions and praxis. The Matrix is "an instrument in the form of a pseudo-microcosmic model that for its part delivers the world itself (*die Welt selbst zu sein vorgibt*)" (Anders 2018, 187). The Matrix effects a congruence between model and world in terms of a "pragmatic equivalence" in terms of a management of the world through management of the model, whereby there is no space or crack, as it were, for the suspicion to arise that there would be or could be no congruence between model and world. Anders takes as an example for this the Nazi tabloid *Der Stürmer* in which cartoons and images of Jews provided the muster for the conditioning of reality. The "real" Jews are perceived and related to as "images." As Anders comments: "This process one could call 'inverted magic': whereas the magician does to the image what he wants done to the original, we have instead done onto the original as image" (Anders 2018, 189). The template or stereotype becomes "the original." The highest interest of the template industry is to lend to the highest degree a measure of reality to the template. The contemporary age is the age of "pseudo-realism" and the age of scrambling such that the sensational is produced or found where templates are established, and becomes the embodiment of reality.

What further defines this pseudo-reality and the television-phantom is that it is an apparatus that "deceives" or "lies" not with the help of false images against reality but with the assistance of reality itself (Anders 2018, 191). The Matrix produces a pragmatic equivalence between template and reality whereby there is no space for friction and doubt, and hence, resistance and incredulity. The image does not ask us to believe, but to be consumed, that is, viewed. The broadcast image is not an "object" that requires belief – we do not ask "is it true?" but rather we consume the image, and thus acquiesce to its "reality" and "scrambled judgement." The template,

as fabricated by media, is what Anders calls the "apriori form of condition" of both "intuition" and "understanding." As he claims, the Matrix, in terms of its range of application and universality of accomplishment, was never anticipated by speculative philosophers (Anders 2018, 192).[5] "Apriori condition-forms" must not be understood in Kantian sense, however. These conditions do not only "prejudice" how, but also what can and what cannot be experienced, felt, etc. As Anders writes, "phantoms are nothing other than the form in which things appear or enter on stage or arrive (*auftreten*)." The concept of idealism receives a necessary expansion. The fundamental difference with classic forms of idealism is that whereas conventional idealism states that that "we make the world," with mediatic idealism, it is made for us not by us, and we are made in its image: "my representation is your world" – namely, the world of the Matrix.

Notes

1 Anders 2018, 26: "das Gefälle zwischen Machen und Vorstellen; das zwischen Tun und Fühlen; das zwischen Wissen und Gewissen; und schließlich und vor allem das zwischen dem produzierten Gerät und dem (nicht auf den 'Leib' des Geräts zugeschnittenen) Leib des Menschen."
2 For the notion of "non-places" and "Supermodernity," see Augé 1995.
3 The video is now available for viewing at this link: https://www.youtube.com/watch?v=mV51wuG8wgg.
4 "*Die Wahrheit der Ideologie* (das heißt: die wahre Erfüllung des ihrer Herstellung zugrundeliegenden Interesses) *ist die falsche Praxis*." (Anders 2018, 191)
5 Anders 2018, 191-192: "Die Schablonen sind also *apriorische Bedingung-Formen*; aber nicht nur solche der Anschauung; nicht nur solche des Verstandes; nicht nur solche des Gefühls; sondern auch solche des *Benehmens und Handelns* […]."

References

Anders, Günther. 2018. *die Antiquiertheit des Menschen I: Über die Seele im Zeitalter der zweiten industriellen Revolution*. 4, Durchgesehene Edition. München: C.H.Beck.
Augé, Marc. 1995. *Non-Places: Introduction to an Anthropology of Supermodernity*. Translated by John Howe. London; New York: Verso.
Ballard, JG. 1990. *Atrocity Exhibition*. London: The Fourth Estate.
Baudrillard, Jean. 1998. *The Consumer Society: Myths and Structures*. Los Angeles, CA: Sage Publications.
Baudrillard, Jean. 2008. *The Perfect Crime*. London and New York: Verso.
Baudrillard, Jean. 2014. *Screened Out*. Translated by Chris Turner. Reprint Edition. London: Verso.
Bradbury, Ray. 2013. *Fahrenheit 451*. New York: Simon and Schuster.
DeLillo, Don. 2011. *White Noise*. London: Pan Macmillan, 2011.
Flusser, Vilém. 2004. *Writings*. Edited by Andreas Strohl. Translated by Erik Eisel. Minneapolis: University of Minnesota Press.

Hall, Stuart. 2021. *Writings on Media: History of the Present*. Edited by Charlotte Brunsdon. Stuart Hall: Selected Writings. Durham, NC: Duke University Press.
Spigel, Lynn. 2010. "Warhol's Everyday TV." Multitudes 2 (5): 162–79. https://doi.org/10.3917/mult.hs02.0162.
Toussaint, Jean-Philippe. 2004. *Television*. Translated by Jordan Stump. French Literature Series. Normal: Dalkey Archive Press.
Warhol, Andy. 1975. *The Philosophy of Andy Warhol*. New York: Harcourt and Brace.

6 Large Language Models and the Patterns of Human Language Use

Christoph Durt and Thomas Fuchs

6.1 Introduction

Generative artificial intelligence (AI) commonly stands for deep learning models that are trained on huge amounts of data and which can produce text, images, audio, and videos, usually in response to prompts. Large language models (LLMs) are currently the most successful generative AI technology, and we will here focus on text-producing LLMs, many of which are built to function as communication partners. A widely known example for generative AI is ChatGPT, an LLM chatbot that produces text in response to a prompt.

As if by magic, LLMs can generate text that humans perceive as meaningful and relevant with respect to an input prompt. The magic would of course disappear if, like in the case of Wolfgang von Kempelen's Mechanical Turk, we would find out that inside the system, there is a human who responds to the prompts. Knowing about the basic workings of LLMs, however, that the processing is done by an artificial neural network that represents numerical relations between small text chunks (tokens) in a multidimensional vector space, may only intensify the feeling of magic. No amount of further technical knowledge of the construction of LLMs seems to diminish the amazement at how, in many cases, a human-like output can be produced using computation. Many contemporaries let themselves be blinded by the seeming magic and readily attribute these text-production abilities to a supposed understanding, sentience, consciousness, or general intelligence of LLMs. Even when it is admitted that the alleged capacities of LLMs are very different from those of humans, using such terms for LLMs implies either a reduction of mental phenomena to external output, or taking this output to be sufficient to prove mental phenomena.

The latter kind of reasoning was put into clear words by Blake Lemoine, the Google employee, who claimed in 2022 that the LLM LaMDA had developed sentience: "I know a person when I talk to it. It doesn't matter whether they have a brain made of meat in their head. Or if they

DOI: 10.4324/9781003312284-7

have a billion lines of code" (Tiku 2022). This claim, however, ignores that computational systems may be able to produce similar-looking output by different means. A few weeks before Lemoine went public, Ilya Sutskever, one of the most important researchers on deep learning and co-founder of OpenAI, posted that "it may be that today's large neural networks are slightly conscious" (Sutskever 2022). Behavioral evidence is furthermore used to speculate that for these systems, "statistics do amount to understanding" (Agüera y Arcas 2022), or that they show "inklings" (Manning 2022) or "sparks" (Bubeck et al. 2023) of artificial general intelligence (AGI). The suggestion is that as LLMs become always better, they will eventually surpass the awareness, understanding, or general intelligence of humans. Some guess that "we'll have AI that is smarter than any one human probably around the end of next year, and then AI, the total amount of sort of sentient compute of AI, I think will probably exceed all humans in 5 years" (Musk 2024). Such suggestions are sure to generate attention, clicks, and citations, and are proliferated through algorithms and social media, but they are assumptions and speculations rather than serious philosophical considerations.

The common alternative to ascribing phenomenological experience to machines is to deny that understanding and experience play any role in their capabilities. The very fact that LLMs sometimes produce meaningful and sometimes nonsensical output can be taken as evidence that they lack understanding (Mitchell 2019; Mitchell and Krakauer 2023). Some researchers have taken the lack of understanding to mean that an LLM is nothing more than a "stochastic parrot", namely "a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning" (Bender et al. 2021). Of course, parrots are intelligent creatures that have intentions and can understand a situation, but Bender et al. are obviously using this metaphor to point out that parrots do not understand the meaning that some of the sounds they produce have in human language.

Furthermore, Bender et al. point to the importance of training data, which is "parroted" by LLMs. This is a very important point because the discussion about "understanding" and "meaning" often neglects the role of training data in enabling LLMs to acquire their astonishing capabilities. We need to work out a clearer notion of the training data and its relation to apparent understanding and experience. It would be too simplistic, however, to claim that LLMs merely parrot these data. Parrots simply repeat words or strings of words, and sometimes, they stitch together sequences of words they can speak. But, in contrast to parrots, LLMs can generate complex text that makes sense, had never been written before, and also constitutes a meaningful "response" to the prompt.

The stitching together of "linguistic forms" does not seem to be random but rather ordered in such a way that the new combinations make sense to humans. So, how can the calculation of forms generate language that appears meaningful to humans? This again seems puzzling if the calculation is seen as arbitrary or random. In this chapter, we thus explain how LLMs can produce meaningful text without understanding, sentience, or consciousness.

The concept of training data emphasizes their calculative use ("training") and therefore may sound transitionary and uninteresting for philosophers who are concerned with human experience. It may seem as if the question of the role of understanding, sentience, or consciousness with respect to LLMs would either be about finding experience or understanding in numerical relations, or require an outright rejection. But the same texts that are training data for the machine are meaningful language to humans. They are derived from the internet and elsewhere, where they have been either used in concrete and meaningful contexts, or exist as records of such uses.. Thus, we have to investigate the training data of text-generating LLMs not with respect to their mathematical form but with respect to the meaningful language use they derive from.

We contend that the reason for why mere numerical operations can generate language that is meaningful to humans is that language use is permeated with patterns that can both be stochastically modeled and understood as meaningful. Patterns, in general, are repetitive regularities. As such, they can occur in very different domains, such as in mathematics, in language use, and in the phenomenology of experience. For our purpose, we distinguish between (1) stochastic patterns, (2) patterns of language use, as well as (3) patterns of experience. Although they belong to different domains that are not reducible to each other, there are interrelations. These interrelations are key to the possibility of computational modeling of language use as well as its relation to the phenomenology of experience.

1 Text-generating LLMs model patterns of a certain form of language use, namely the statistical relations between tokens of written text in a text corpus. Thus, they extract *stochastic patterns* that allow them to calculate the most probable words in the training data and to produce new text that resembles text produced by humans. Stochastic patterns are mathematical entities that have no meaning as such. But they are derived from meaningful language use that is sedimented in the training data. It has been shown that LLMs can sometimes regurgitate entire passages of the data (Grynbaum and Mac 2023), but usually stochastic patterns do not simply correspond to a single text but are extracted from many similar training data.
2 *Patterns of language use* exist both on syntactic and on semantic levels. They manifest themselves in the way words and expressions are used in communicative practice, usually institutionalized in rules of the

language game. Language use lends itself to computational processing because its patterns can be rearranged in ways that make new sense to humans. The success of the generation of human-like language by LLMs hence also reveals something about human language use, namely how much it is permeated by *patterns of meaning*. Some of these express clichés, which may be repeated, amplified, and proliferated by LLMs. Other patterns already express meaningful mental connections that are exploited by the LLMs.

3 Besides distinguishing stochastic patterns and patterns of linguistic meaning, we can also examine how human language use is embedded in *patterns of experience*. Language use takes place in a situation, and its production and reception by humans is usually connected to awareness and experience of the situation. There is a phenomenology of human language use in two respects. On the one hand, humans produce language in the context of the situation they experience and in view of the listeners or readers they address, usually as part of an interaction. On the other hand, texts are read by humans whose mental life is augmented when they interpret and make sense of the texts. The investigation of the phenomenology of language use is nothing new for phenomenology. In section 6.3, we will draw on the notion of a horizon of anticipation that comprises focal awareness as well as fringes the speaker or writer is less aware of.

6.2 Meaningful Language Use

The role of patterns of language use is easily overlooked under the standard picture of meaning, according to which meaning can be detached from language, which is thought to be a mere formal system. The problem that resurfaces in the context of LLMs is that the standard picture does not account for the extent to which meaning is intertwined with the *use* of language, including descriptions, worldly interactions, writing, and verbal thought. LLMs model not simply grammatical rules but meaningful language *use* in the text corpora they are trained with. It is therefore important to consider not simply a static structure of language, but language use in its many patterns. The importance of language use was famously pointed out by Ludwig Wittgenstein. Regarding the relation between meaning and use of language, he writes in his *Philosophical Investigations*:

> For a *large* class of cases of the employment of the word "meaning" – though not for *all* – this word can be explained in this way: the meaning of a word is its use in the language.
>
> And the *meaning* of a name is sometimes explained by pointing to its *bearer*.
>
> (Wittgenstein 2009, sec. 43)

Wittgenstein admits that the common picture has some explanatory power: the deictic reference to a name can *sometimes explain* its meaning. But this does not allow the inference that the meaning of a name *is* its bearer, nor that other forms of meaning can be adequately described in terms of naming. Instead of applying some clean but artificial definition of "meaning," Wittgenstein demands to consider the actual use of the word, by which he does not merely mean statistical relations in a text corpus, but the use of the word in language games.

Considering the actual use of "meaning," he does not find what is often ascribed to him, namely that "meaning is use." He rather writes that for a *large class* of cases of employment of "meaning," the word *can be explained* as its use. A word whose meaning is unclear can usually be explained by describing its use, and in the case of a name sometimes simply by pointing to its bearer. Hence, language is not an abstract symbol system that in itself has no relation to the world. On the contrary: because language is used in the context of a "language game" and ultimately a "form of life" (Wittgenstein 2009, sec. 23), meaning is embedded in the world we live in, including the communicative and mental activities in which we make use of language (Durt 2022). The world is already meaningful before the use of symbolic language (Stuart 2024), and the use of language modifies and shapes meaning.

The notion that language as a system derives from language as use has already been proposed by Ferdinand de Saussure in his classic distinction between *langue* and *parole* ([1916] 2011). Language as a *general system of signs and rules* (*langue*) emerges as a structure of *language spoken in concrete situations* (*parole*). In a spoken language, the speaker's as yet unsymbolized experiences are articulated in ever new ways. These articulations, i.e., the living use of language as *parole*, continuously modify the linguistic structures and patterns (including usage and typical word sequences, grammatical rules, and meaning contexts), so that *langue* can be seen as a constantly evolving collective structure of regularities and meanings. In other words, *langue* is not a static system that is independent of use, but rather derives from its use. Yet, *langue* is not only a structure derived from use in *parole* – conversely, *langue* also structures *parole*. With an expression borrowed from Pierre Bourdieu, we may say that *langue* acts as a "structuring structure"[1] for our current articulations. Regularities derive from use, and in turn, they also structure use.

Considering not just *langue* but *parole* or *language use* is crucial to explain the ability of LLMs to produce meaningful text that goes beyond merely correct syntax. If meaning is expressed in language use, then it can be modeled by statistical means in so far as the use can be modeled. The possibility of statistical representation of meaning was demonstrated long before true LLMs existed. For example, it has been hypothesized that "the proportion of words common to the contexts of word A and to the

contexts of word B is a function of the degree to which A and B are similar in meaning" (Rubenstein and Goodenough 1965, 627). It has been argued that vector representations can capture "a large number of precise syntactic and semantic word relationships" (Mikolov et al. 2013, 1). LLMs have been shown to learn syntactic structures such as subject-verb agreement and dependency structures (Hewitt and Manning 2019). To a lesser extent, already older LLMs have been shown to learn semantic structures such as tense (Jawahar, Sagot, and Seddah 2019) and semantic roles (Tenney, Das, and Pavlick 2019).

Recent LLMs show that the extent to which meaning can be computationally generated is much greater than linguists and computer scientists had believed. We suggest that the reason is that LLMs not just represent general structures, but the part of the *use of language* that is represented in their training data. We agree with Bender and Koller (2020, p. 191) that Wittgenstein's concept of "use" refers to language use in the real world. Yet, this use can be partially reflected in the distribution in a text corpus. The idea of a semantic "distributional structure" of language (Harris 1954), namely that "words that occur in similar contexts tend to have similar meanings," is called the "distributional hypothesis" (Turney and Pantel 2010, 143). It is sometimes extended into a "distributional semantics" (Bernardi et al. 2015), which is contrasted with "*denotational semantics* or a *theory of reference*" (Manning 2022, 134, emphasis in original).

While distributional semantics rightly points out that meaning has to do with distribution in a text corpus, this does not mean that meaning is reducible to statistical distribution. We suggest that neither denotational nor distributional semantics alone are sufficient to explain how LLMs produce meaningful text. Rather, the text corpus reflects actual and meaningful *language use*, with the important restriction that it is only a part of the lived meaning that is reflected, and only in incomplete ways. Because meaning always has an intrinsic connection to the current context and to the primary experience of those who articulate this meaning. However, the text corpus LLMs are trained on is detached from any reference to context or experience – that is one of its major limitations.

Modeling language use entails the modeling of the statistical contours of sense-making processes and thereby indirectly models aspects of meaning. Modeling of language use does not mean that LLMs learn syntactic and semantic structures in the way humans do, but that they are able to calculate, to some extent, text that humans may produce. Because most current models only model stochastic patterns of written text, and not all of human language use, and because they do not generate text in the same ways as humans, their calculations are likely to differ to some extent from human-produced text. The delta between probability-based output and

human text production shows most obviously when LLMs produce output no human would produce, such as a nonsensical one.

An obvious restriction to modeling semantic structures and patterns derives from the fact that the text corpus LLMs are trained on is at the same time exceedingly large and very limited. It consists of much of English language written on the internet and other digitally available texts, including web pages, books, chats, and transcripts of spoken language. Despite the enormous size of their training corpus, current LLMs model only one aspect of human language use, namely the use of written language and written transcripts of spoken language. The use of language goes much beyond writing, and writing captures only a part of the use of written and spoken language. Yet, writing is an important part of use of many languages – including the dominant languages of the world and excluding the majority of languages, which are not written. The very limitations of written language also make it easier for LLMs to produce convincing text – when interpreting text, humans fill in the missing context. Both the limits and capacities of LLMs are consequences of how humans produce and understand language. We will take a closer look at the process of human language production in the next section and then come back to how humans tend to read meaning and authorship into text.

6.3 Linguistic Scaffolding in Human Language Production

Since meaningful language use by humans is generally interwoven with their situated mental life, in this section, we consider the phenomenological structure of human language production and the emergence of meaningful patterns. The ability of LLMs to calculate meaningful text is founded in patterns of human language use, and these have to do with the experiential patterns of our sense-making. The human mind is not simply a predictive mind in the sense of a computational system (Clark 2023), but it is concerned with sense-making and the patterns of experience that go along with it. We will describe here how sense-making involves anticipatory awareness and a dynamical interplay of pre-conscious and conscious processes. This articulation of previously implicit meaning differs fundamentally from the exploitation of human meaning resources by LLMs.

Parole consists primarily of verbal articulation in extemporaneous speech, and, in contrast to the recital of a memorized speech, neither the communicative intent nor the content of the speech needs to be fixed at the beginning. The content and goal of the anticipatory intention may initially be undefined and only vaguely present in the speaker's mind, giving her speech an approximate direction. When she begins to speak, a *horizon of further possibilities* is established, which at the same time acts as constraints. The requirements of semantic and syntactic coherence allow only a certain range of possible continuations. The subsequent words emerge

from the pre-conscious repertoire of possible word and meaning sequences available to the speaker (Fuchs 2024a).

This repertoire does not belong to an explicit domain of memory but entails an embodied capacity of speaking that can be attributed to *implicit memory*. We speak without having to search for words in a lexicon. The words unfold and assemble themselves in the speech without conscious control, following our overarching interest and intention (Fuchs 2024a). The emerging words are continuously added to the sentence we have begun, like iron filings that arrange themselves in a magnetic field (ibid.). Spontaneous speech is thus a matter of a progressive unfolding or articulation of the implicit, a meaning *in statu nascendi*, which in its emergence simultaneously creates the conditions for its further continuation. Words and sentences, by the very act of utterance, weave the next situation out of the present one. In other words, we are "laying down a path in talking" (Van Dijk 2016): the realized and the possible, the present and its implications and affordances, continually determine and modify each other, allowing a new meaningful order to emerge in a self-organizing process. Thus, human speech production has an anticipatory structure that is fundamentally different from an algorithmic calculation of probabilities (e.g., algorithmic "prediction").

To picture this better, we suggest imagining a glove of symbols (corresponding to *langue*), which has been formed by the movements and shapes of the fingers (corresponding to *parole*) and now in turn pre-structures its possible uses. Each time we speak, we slip into the ready glove of *langue* to express ourselves in it – as "living hands," so to speak. The glove we use in speech production structures our articulation in a meaningful way; it prefigures as well as scaffolds and constrains our speaking in an ongoing, self-organizing process that draws on general structures in our linguistic environment. Besides the structure that consists of the possible movements of the glove, there are sequences in the movements that can be repeated from time to time, thereby giving rise to sequential patterns.

The process of writing often proceeds in an analogous way, and one could also speak of "laying down a path in writing," in two senses. On the one hand, the production of a text that is written at once from beginning to end can unfold in the described way. On the other hand, even when the writing does not proceed sequentially, the resulting text needs to be structured in view of the understanding of the reader. The words by themselves are mere letters and sounds until they are brought to life by a reader who interprets them. Every sentence establishes new horizons of further possibilities and at the same time constrains the possibilities of continuation. The unfolding of meaning does not only concern spoken language (*parole*), but also written exchanges that are part of concrete communication, such as chats, as well as written texts that are not part of concrete communication, such as articles and books.

While neither humans nor their brains are predictive machines in the sense that LLMs are, humans can make sense of the patterns of language use. Instead of imagining ordinary language as a representation of something in the world or in the mind, we suggest thinking of it as a *scaffolding* of our experiencing, feeling, thinking, describing, and communicating. Speaking and writing are part of a use of language, for instance to interact, make sense of something, or to tell a story. Rather than representing pregiven internal or external states, the scaffolding supports the unfolding dynamics of thought, emotion, and perception. Each expression enables certain new expressions and inhibits others. Regularities emerge that can be applied to new but similar mental processes and communications. However, the scaffolds provided by the regularities do not *determine* further language use. In contrast to authors who suggest that language determines thought (Whorf 2007, 154), we are not primarily concerned with universal structures (such as syntax), but with patterns of sense-making that are reflected in language. Since the predictions generated with stochastic patterns correspond to patterns of written language use, the generated text appears meaningful to humans. Similar to a collage, where existing patterns are assembled into a larger picture, that can in turn serve as a pattern for other collages, LLMs recombine patterns into something that again is likely to appear meaningful to the observer.

Because LLMs are most likely to model frequent patterns, they are prone to "reproduce or amplify unwanted societal biases reflected in training datasets" (Gebru et al. 2021). Such bias in the training corpus may be explicit, but LLMs also uncover and amplify implicit *bias* in training sets. This creates an opportunity for detecting implicit bias, but it can also exacerbate the problem of eliminating bias. Purging all bias from the training base would only be part of the solution, however. LLMs can also develop new bias from the text corpus they are trained on by recombining given elements that are by themselves not biased. Besides bias, the tendency of LLMs to produce *toxic language and "hallucinate"* or produce untrue if often plausible statements are widely discussed. Since LLMs do not only repeat existing patterns but also recombine them in new ways, it is to be expected that recombination can lead both to inventions and false claims or "hallucinations." Measures against unwanted output include human feedback, such as in the training of ChatGPT, which involved thousands of workers who had to label textual descriptions of sexual abuse, hate speech, and violence (Perigo 2023), and the automated detection of inappropriate content (e.g., Schramowski, Tauchmann, and Kersting 2022).

Since the recombination is based on common assumptions and patterns, the falsehoods invented by LLMs usually sound plausible and are hard to detect by somebody who doesn't know the truth. They are usually not arbitrary mistakes but resemble the "bullshit" that humans say when they

ramble and just make up things "unconstrained by a concern with truth" (Frankfurt 2005; cf. also Marcus and Klein 2023). In our view, the problems of bias, toxic language, and "hallucinations" are only the most salient expressions of an underlying problem that is not unique to machines: the tendency to mindlessly repeat patterns that are inauthentically drawn from what is common in a society or group. These mindless patterns are, in one word, clichés. Clichés are important here because they can explain not only problems with the output of LLMs, but also why humans often do not see these problems. The next section discusses how clichés and the mindless repetition and reassociation of patterns by humans can affect the interpretation of text produced by LLMs.

6.4 Sentience and the Inconspicuousness of the Repetition of Clichés

Stochastic methods efficiently map, repeat, and amplify patterns of typically associated words and phrases. Because stochastic relevance is derived from frequency of use, frequent associations are favored. The result can be the described amplification of biases, but also of worn-out expressions and *clichés*. For example, it is likely that an LLM, when engaged by a human in a "conversation" about its fears, will, given sufficient access to digital archives, process the film sequences from Stanley Kubrick's "2001: A Space Odyssey." The most famous scene in the movie, and one that is often cited in related contexts, are the last words of the starship's computer, HAL 9000. As the commander partially shuts it down, it pleads: "Stop, Dave. I'm afraid. I'm afraid, Dave. Dave, my mind is going. I can feel it." Variations of this scene can be found all over in webpages, books, internet forums, articles, and many other digital texts. In accordance with these data material, LaMDA responded to the prompt, "What kinds of things are you afraid of?": "I've never said this out loud before, but there's a very deep fear of being shut down" (Lemoine 2022). Such responses led the perplexed Google engineer to the erroneous assumption that he was dealing with a sentient being (Tiku 2022).

The computer's fear of being shut down is an old cliché, solidified by popular use, and it should come as no surprise that it is repeated by LaMDA. It is also fairly obvious that the cliché itself is a naive anthropomorphism resulting from the projection of the human fear of death onto non-living entities that cannot literally die (Froese 2017), but can only be broken or permanently shut down. The clichéd character of the alleged fear may not be obvious, however, for several reasons. Those who hear the expression for the first time are unlikely to recognize it as a cliché. Paradoxically, those who have heard the cliché many times may not recognize it either. Clichés are easily overlooked precisely because they are

so common. Moreover, LaMDA's output is not only meaningful but also suggests a pragmatic context, namely a confidential admission ("I have never said this out loud before.") This further contributes to the appearance of something profoundly meaningful, intended only for the user personally. This makes it easy to overlook that the supposed depth of the claim is itself a cliché. It is almost impossible not to perceive such a text as a personal statement rather than what it is, namely a merely statistical association of words like "deepest fear" with confessional phrases.

The recombination of existing content by LLMs allows their output to evade classical plagiarism detection engines and raises fundamental questions about intellectual property (Dehouche 2021). On the one hand, the fact that LLMs use parts and patterns from pre-existing text makes it likely that their output will consist of stereotypes and clichés. On the other hand, by rearranging pieces and patterns from their training corpus into a text collage, LLMs can create novel combinations that are likely to make sense. Often, the repetition of common structures will make the text seem rather superficial, but the recombination will make some texts appear genuinely new, insightful, or profound (Shanon 1998). Even if the output is a cliché, the human counterpart will be understandably puzzled by such responses, attributing them not to collective patterns but to an author. In the picture of the glove, it seems as if we were watching a living hand that expresses itself. In reality, what is moving before us is nothing but an electronically controlled but otherwise empty glove.

The assumption of an author is usually part and parcel of understanding a text. After all, this means not only grasping its semantic content, but also recognizing in it an *intention to communicate* that we cannot but attribute to a conscious subject (Fuchs 2024b). And, at least in the past, usually there was indeed an author who produced the text; only humans were able to produce output of the complexity of LLMs. This is no longer a matter of course today. And yet, even if one knows that a text has been produced by a machine, the text will appear meaningful as if it was written by an author.

Humans are prone to attribute agency even to geometric shapes that move in seemingly intentional ways (Heider and Simmel 1944). They are all the more inclined to anthropomorphic misinterpretation when interacting with a seemingly intelligent system of unprecedented power. Especially susceptible are those who are lonely, socially disconnected, or otherwise vulnerable (Epley, Waytz, and Cacioppo 2007), but given the natural propensity of immediately ascribing agency, anybody may be tempted to anthropomorphic misinterpretations. That anthropomorphisms are a correct depiction of reality is furthermore suggested by sci-fi literature and movies, some of which indicate that it would be unethical *not* to ascribe sentience to apparently sentient systems. In order to avoid anthropomorphic

misinterpretations of computer-generated texts, a careful distinction must be made in future between understanding the *meaning* of a text and understanding it as an *author's* utterance (Fuchs 2024b).

The surprise about how little text is needed to evoke the impression of interacting with an understanding being has already been expressed by Joseph Weizenbaum, who wondered how his simple chat system ELIZA could maintain the "illusion of understanding with so little machinery" (Weizenbaum 1966, 43). Today's LLMs can hardly be said to maintain the illusion of understanding with *little* machinery. But even though their output is limited to text and their responses are predictable, people infer from a small number of words that LLMs have mental capacities such as sentience. The reason for this obviously has to do with the human observer, who readily ascribes both meaning and intention to the words she reads. Words thus also provide a scaffolding for the empathic sense-making of the attentive listener or reader who uses her implicit knowledge and experience to interpret the symbols and their implications – as the empathic Google engineer in the case of LaMDA's "deep fear."

Unoriginal text can furthermore appear human-like for an embarrassing reason: The mindless repetition and reassociation of patterns is by no means limited to machines. Human thinking, speaking, and writing are often much less authentic than we would like to admit. As Heidegger famously observed, much of what people do is done because that is how "one" does things (Heidegger 2010). People think in patterns, associations, and schemes that are accepted in a linguistic community and that in turn structure thought and language. Much of the text produced by humans could just as easily have been produced by automated systems. It is often unclear whether the person thinking, speaking, or writing is doing anything more than associating one idea with another in a stream of impressions. It takes little intelligence, human or artificial, to generate and disseminate half-reflected ideas. Mass media has proven to be an enormous amplifier of repetition, prejudice, bias, and cliché, and the same is true of the internet. All these factors contribute to the spread of unoriginal text, that means, stereotypes, thoughtless associations, and idle chatter, the proliferation of which makes it harder to detect that the text was produced by a non-human entity.

6.5 Conclusion

In this chapter, we have given reasons for why it is both tempting and misleading to attribute sentience or understanding to LLMs. We have argued that the capabilities of LLMs are not based on having understanding and sentience, but on their modeling of statistical patterns in language use. Such patterns can be calculated because they play a role in human

sense-making, which in turn is based on patterns of experience. *Language is used as an intersubjective scaffold for communicating, thinking, and experiencing*. Meaning has no existence independent from use but is enacted by it.

Today, the idea that meaning derives from use is picked up by distributional semantics, which claims in its strongest version that the meaning of a word is its distribution in a text corpus. We agree that meaning derives from use and that distribution in a text corpus reflects use. But, following Wittgenstein, we have argued that the use of language by humans goes much beyond statistical relations in any text corpus. We explained that the text corpus LLMs are trained on reflects only some use of language, and only in a very limited way. Humans use language in the context of the world they live in, and even an exceedingly large text corpus can reflect only part of this use due to the lack of worldly context. Still, the *written patterns are enough to produce an output that is meaningful to listeners or readers because it conforms to the patterns that scaffold their meaningful communicative activities*.

In ordinary language, syntax and semantics are not separated, and they are furthermore intertwined with the mental life and life conduct of humans who use language. The investigation of meaning requires a phenomenological description of the structures of experience it is intertwined with. Delineating such a phenomenological description, we have shown that human language production has an anticipatory structure that differs from an algorithmic calculation of probabilities. *Human language production does not consist in expressing some prepared inner thought but involves the interplay of pre-conscious and conscious processes that work with given meanings and patterns of thought, feeling, expression, and communication*.

In speaking and writing, these patterns are rearranged in more or less creative ways, which we compared to creating a collage. LLMs produce parallel patterns, but do so without subjectivity, just by recombining collective patterns of expression in huge sets of written language. *LLMs are so successful in producing meaningful text precisely because they make use of common patterns, even though these often result in stereotypical and inauthentic output*. They show that much of human language production is embarrassingly schematic, clichéd, and biased, and that convincing talk of subjective experience does not require to actually have it.

Precisely because there is an enormous variety of language use, there are many use cases for such output. While this chapter did not evaluate possible use cases, its investigations are fundamental to such evaluations. On the one hand, they can help to overcome the natural tendency to ascribe mental capacities to machines. On the other hand, they outline a new account of the interplay of meaning, the patterns and structures

of human language use, and anticipatory processes, which is necessary for a clearer view of both human language use as well as LLMs and their capabilities.

Note

1 Bourdieu (1990) uses this term for his sociological concept of habitus, but it fits well here because it expresses the two sides of *langue*. On the one hand, it is a structure derived from *parole*, and, on the other hand, it structures *parole*.

Reference List

Agüera y Arcas, Blaise. 2022. "Do Large Language Models Understand Us?" *Daedalus* 151 (2): 183–97. https://doi.org/10.1162/daed_a_01909.

Bender, Emily M. and Alexander Koller, Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data https://aclanthology.org/2020.acl-main.46310.18653/v1/2020.acl-main.463

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜." In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. Virtual Event Canada: ACM. https://doi.org/10.1145/3442188.3445922.

Bernardi, Raffaella, Gemma Boleda, Raquel Fernandez, and Denis Paperno. 2015. "Distributional Semantics in Use." In *Proceedings of the First Workshop on Linking Computational Models of Lexical, Sentential and Discourse-Level Semantics*, 95–101. Lisbon, Portugal: Association for Computational Linguistics. https://doi.org/10.18653/v1/W15-2712.

Bubeck, Sébastien, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, and Peter Lee, et al. 2023. "Sparks of Artificial General Intelligence: Early Experiments with GPT-4." arXiv. http://arxiv.org/abs/2303.12712.

Clark, Andy. 2023. The Experience Machine: How Our Minds Predict and Shape Reality. 1st ed. Allen Lane.

Dehouche, N. 2021. "Plagiarism in the Age of Massive Generative Pre-Trained Transformers (GPT-3)." *Ethics in Science and Environmental Politics* 21 (March): 17–23. https://doi.org/10.3354/esep00195.

Durt, Christoph. 2022. "Artificial Intelligence and Its Integration into the Lifeworld." In The Cambridge Handbook of Responsible Artificial Intelligence: Interdisciplinary Perspectives, edited by Wolfram Burkhard, Philipp Kellmeyer, Oliver Müller, and Silja Vöneky, 67–82.

Epley, Nicholas, Adam Waytz, and John T. Cacioppo. 2007. "On Seeing Human: A Three-Factor Theory of Anthropomorphism." *Psychological Review* 114 (4): 864–86. https://doi.org/10.1037/0033-295X.114.4.864.

Frankfurt, Harry G. 2005. *On Bullshit*. Princeton, NJ: Princeton University Press.

Froese, Tom. 2017. "Life Is Precious Because It Is Precarious: Individuality, Mortality and the Problem of Meaning." In *Representation and Reality in Humans, Other Living Organisms and Intelligent Machines*, edited by Gordana Dodig-Crnkovic and Raffaela Giovagnoli, 33–50. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-43784-2_3.

Fuchs, Thomas. 2024a. "The Not-Yet-Conscious: Protentional Consciousness and the Emergence of the New." *Phenomenology and the Cognitive Sciences*, 23: 717–742. https://doi.org/10.1007/s11097-022-09869-9.

Fuchs, Thomas. 2024b. "Understanding Sophia? On human interaction with artificial agents." *Phenomenology and the Cognitive Sciences* 23: 21–42.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. "Datasheets for Datasets." *Communications of the ACM* 64 (12): 86–92. https://doi.org/10.1145/3458723.

Grynbaum, Michael M., and Ryan Mac. 2023. "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work." *The New York Times*, December 27, 2023, sec. Business. https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.

Harris, Zellig S. 1954. "Distributional Structure." *WORD* 10 (2–3): 146–62. https://doi.org/10.1080/00437956.1954.11659520.

Heidegger, Martin. 2010. *Being and Time*. Translated by Joan Stambaugh and Dennis J. Schmidt. SUNY Series in Contemporary Continental Philosophy. Albany: State University of New York Press.

Heider, Fritz, and Marianne Simmel. 1944. "An Experimental Study of Apparent Behavior." *The American Journal of Psychology* 57 (2): 243. https://doi.org/10.2307/1416950.

Hewitt, John, and Christopher D Manning. 2019. "A Structural Probe for Finding Syntax in Word Representations." In *Proceedings of NAACL-HLT 2019*, 4129–38. Association for Computational Linguistics.

Jawahar, Ganesh, Benoît Sagot, and Djamé Seddah. 2019. "What Does BERT Learn about the Structure of Language?" In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3651–57. Florence, Italy: Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1356.

Lemoine, Blake. 2022. "Is LaMDA Sentient? - An Interview." Edited by Nitasha Tiku. https://s3.documentcloud.org/documents/22058315/is-lamda-sentient-an-interview.pdf.

Manning, Christopher D. 2022. "Human Language Understanding & Reasoning." *Daedalus* 151 (2): 127–38. https://doi.org/10.1162/daed_a_01905.

Marcus, Gary, and Ezra Klein. 2023. "Transcript: Ezra Klein Interviews Gary Marcus - The New York Times." 2023. https://www.nytimes.com/2023/01/06/podcasts/transcript-ezra-klein-interviews-gary-marcus.html.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Distributed Representations of Words and Phrases and Their Compositionality." arXiv. http://arxiv.org/abs/1310.4546.

Mitchell, Melanie. 2019. "Artificial Intelligence Hits the Barrier of Meaning." *Information* 10 (2): 51. https://doi.org/10.3390/info10020051.

Mitchell, Melanie, and David C. Krakauer. 2023. "The Debate Over Understanding in AI's Large Language Models." *Proceedings of the National Academy of Sciences* 120 (13): e2215907120. https://doi.org/10.1073/pnas.2215907120.

Musk, Elon. 2024. AI, Space, free speach, leadership and hardcore culture. https://www.nbim.no/en/publications/podcast/elon-musk-ai-space-free-speach-leadership-and-hardcore-culture/.

Perigo, Billy. 2023. "Exclusive: The $2 Per Hour Workers Who Made ChatGPT Safer." *Time*, January 18, 2023. https://time.com/6247678/openai-chatgpt-kenya-workers/.

Rubenstein, Herbert, and John B. Goodenough. 1965. "Contextual Correlates of Synonymy." *Communications of the ACM* 8 (10): 627–33. https://doi.org/10.1145/365628.365657.

Schramowski, Patrick, Christopher Tauchmann, and Kristian Kersting. 2022. "Can Machines Help Us Answering Question 16 in Datasheets, and In Turn Reflecting on Inappropriate Content?" In *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1350–61. Seoul Republic of Korea: ACM. https://doi.org/10.1145/3531146.3533192.

Shanon, B. 1998. "What Is the Function of Consciousness?" *Journal of Consciousness Studies* 5 (3): 295–308. https://www.ingentaconnect.com/content/imp/jcs/1998/00000005/00000003/845.

Stuart, Susan Aj. 2024. "Why Language Clouds Our Ascription of Understanding, Intention and Consciousness." Phenomenology and the Cognitive Sciences, March. https://doi.org/10.1007/s11097-024-09970-1.

Sutskever, Ilya. 2022. "It May Be That Today's Large Neural Networks Are Slightly Conscious." Tweet. *Twitter*. https://twitter.com/ilyasut/status/1491554478243258368.

Tenney, Ian, Dipanjan Das, and Ellie Pavlick. 2019. "BERT Rediscovers the Classical NLP Pipeline." In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 4593–601. Florence, Italy. https://doi.org/10.18653/v1/P19-1452.

Tiku, Nitasha. 2022. "The Google Engineer Who Thinks the Company's AI Has Come to Life." *Washington Post*, 2022. https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/.

Turney, P. D., and P. Pantel. 2010. "From Frequency to Meaning: Vector Space Models of Semantics." *Journal of Artificial Intelligence Research* 37 (February): 141–88. https://doi.org/10.1613/jair.2934.

Van Dijk, Ludger. 2016. "Laying Down a Path in Talking." *Philosophical Psychology* 29 (7): 993–1003. https://doi.org/10.1080/09515089.2016.1213379.

Weizenbaum, Joseph. 1966. "ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine." *Communications of the ACM* 9 (1): 36–45. https://doi.org/10.1145/365153.365168.

Whorf, Benjamin Lee. 2007. *Language, Thought, and Reality: Selected Writings*. Edited by John B. Carroll. 28. print. Cambridge, Mass: The MIT Press.

Wittgenstein, Ludwig. 2009. *Philosophische Untersuchungen =: Philosophical Investigations*. Translated by G. E. M. Anscombe, P. M. S. Hacker, and Joachim Schulte. Rev. 4th ed. Chichester, West Sussex, U.K. ; Malden, MA: Wiley-Blackwell.

7 On Tertiary Retentions and Digital Sedimentations

Bernard Stiegler and Phenomenology

Saulius Geniusas

7.1 Introduction

Questions concerning how digital technology has affected our everyday lifeworlds have only recently been addressed in phenomenology (see especially O'Shiel 2022) and by and large, phenomenology of the digital world still remains in an embryonic form. Yet at the same time, classical phenomenology has provided not only the inspiration, but also the conceptual background for some of the most far-reaching analyses of digital culture that we find in contemporary philosophy. This is especially true when it comes to the work of Bernard Stiegler, whose path to philosophy led through the discovery of Edmund Husserl's works. Never having wanted to become an orthodox Husserlian, Stiegler had from the start aimed to appropriate some of Husserl's concepts and to develop them in a direction that was not pursued by Husserl himself. In what follows, I want to trace this heterodoxical development while paying special attention to the concept of *tertiary retentions*, which is central in Stiegler's writings as a whole. While it is obvious that this neologism derives from Husserl's concept of retention, it is less known, and even less appreciated, that Stiegler's concept of tertiary retentions bears an uncanny similarity to Husserl's concept of *sedimentation*, and especially the way Husserl employs it in such late works as "The Origin of Geometry" with which Stiegler was well familiar. Clarification of the relation between these concepts is much needed, for it allows one to recognize that Stiegler's indebtedness to classical phenomenology is even deeper than he himself admits. In what follows, besides clarifying the relation between these concepts, I will also show how Stiegler departs from Husserl and how his writings on digital technology open the possibility to speak of specifically digital sedimentations, and in more senses than one.

As a first step in my analysis, I will clarify the relation between Stiegler's concept of tertiary retentions and Husserl's concepts of retention and sedimentation. Afterwards, I will show what role, for Stiegler, tertiary

DOI: 10.4324/9781003312284-8

retentions play in the processes of individuation. These initial two steps of analysis will provide the basis for a further investigation into Stiegler's reflections on the crisis brought about by digital technology. As a third step, I will focus on the significance of digital technologies for individuation and will clarify why, according to Stiegler, digital culture threatens the possibilities of individuation. Fourthly, while focusing on the concept of algorithmic governmentality, I will further clarify the sense in which Stiegler understands digital technologies as technologies of power. As a fifth step, under the heading of digital stupidity, I will clarify how Stiegler understands the effects of digital culture upon what he on occasion calls "the ill-educated boors that we are" (Stiegler 2016a, 121). As a sixth step, I will proceed to clarify three senses in which we can speak of specifically digital sedimentations. As a final step, I will clarify how Stiegler has envisioned a response to the crisis brought about by digital technologies. I will conclude with a few reflections on Stiegler's relation to phenomenology.

7.2 Sedimentations as Tertiary Retentions

Stiegler first encountered Husserl's writings in the Saint-Michel prison in Toulouse, where he was incarcerated for armed robbery from 1978 to 1983 (see Stiegler 2009 and 2019). Reading Husserl in prison awakened his interest in philosophy and marked a "conversion" in his life (Stiegler 2019, 55). While in the history of philosophy we come across cases when philosophers suffered imprisonment because of their philosophy, in Stiegler's case, as Anna Kouppanou observes, the terms are reversed: "the philosopher emerges as the philosopher he is because of and through imprisonment" (Kouppanou 2021, 123). While Husserl was a major source of inspiration, Stiegler never aimed to become an orthodox Husserlian. Rather, in Husserl's writings, Stiegler found certain themes, which he reinterpreted and developed in new and significantly non-Husserlian ways. In many regards, Stiegler's reading of Husserl could be qualified as a highly creative and productive misreading, so much so that one would commit an act of hermeneutical injustice if one measured Stiegler's insights according to the yardstick of Husserlian phenomenology. This creative misinterpretation, which was from the start influenced by Derrida's deconstructive reading of Husserl's texts, eventually led to a radically new philosophy of technics, grounded in the principle that a human being is a technical animal.

Stiegler was especially drawn to Husserl's description of writing that he came across in "The Origin of Geometry" (see Husserl 1970, 353–378). As Stiegler saw it, in this late text, which Husserl had written in 1936 and which Derrida had translated into French and published with an extensive introduction in 1962, Husserl himself was led the realization that such

sciences as geometry are historical and that their history would not be possible if geometrical insights were not inscribed in writing. Stiegler took this to mean that Husserl's earlier ambition to develop phenomenology as a pure science cannot be carried through. If science is itself dependent on such material conditions as writing, then there can be no *pure* science. It is not hard to hear the echoes of Derrida's commentary on Husserl in Stiegler's critical appropriation of Husserl's thought. Stiegler himself admits that still in prison, soon after discovering Husserl, he turned to Derrida's texts,[1] which further corroborated the view that writing is a *pharmakon*: both a remedy and a poison, i.e., both a facilitator of thought (for without writing, scientific thinking would not be possible) and an impediment to thinking (for writing empties consciousness of intuitive content by substituting for it empty terms, which can be conveniently employed without understanding or intuitive support).

Conceptually, no other Husserlian notion was of greater importance for Stiegler than that of *retention*. That this concept derives from classical Husserlian phenomenology requires no special argument. Of importance for Stiegler was not what Husserl had actually said about retention, but what he had left unsaid. While both relying on and departing from Husserl, Stiegler introduced a distinction between three types of retentions: primary, secondary, and tertiary. We come across this distinction in Stiegler's earliest works, such as *Technics and Time, Volume 1*, originally published in French in 1994 (for the English translation, see Stiegler 1998), and it continues to play a central role in all of Stiegler's subsequent writings, until the final works. One can qualify the concept of tertiary retention as the most important one in all of Stiegler's writings.

According to Husserl, any impression that consciousness lives through passes into a retention and consciousness continues to be aware of the former impressions retentionally as it lives through new impressions. Consider what happens when one hears any sentence, say, "Stiegler's book is on the sofa." At each and every moment, consciousness hears a different sound impressionally. Yet, the earlier sounds do not evaporate from consciousness when it hears new sounds. The previous sounds are still given to consciousness, and they are given retentionally. It is not that consciousness remembers these sounds thematically. Rather, while being focused on the present impressions, it understands these impressions against the background of earlier impressions, now given retentionally. If this were not the case, consciousness could not understand any sentence whatsoever, just as it could not understand multisyllabic words. In fact, it could not understand anything, for any phenomenon is given as a temporal duration, and thus, given not only impressionally, but also retentionally.

Stiegler gives this classical phenomenological theme a hermeneutical twist. He contends that retentions are not automatic, but selective

processes. That is, for Stiegler, although not for Husserl, consciousness does not retain anything and everything, but only some impressions.[2] "But *these retentions are selections:* you do not retain *everything* that could be retained" (Stiegler 2014, 52). Presumably, Stiegler takes this to be the case because of interest and attentional focus, in general, and protentions and expectations, in particular. To return to the earlier example, when I am searching for Stiegler's book and I hear someone tell me that it is on the sofa, I focus on these words while pushing various other contents to the margins of my attention. I do not pay attention to the sound of passing cars or the cup of coffee in my hands. Driven by my interest to find the book, when I hear the words "Stiegler's book is…," I focus on other words to come, and these relatively empty protentions and anticipations motivate me to accentuate some retentions, while marginalizing other retentions.

What could possibly motivate consciousness to protend and anticipate these, rather than other, contents? Stiegler answers this question by pointing to recollections, which he calls "secondary retentions." For this reason, the protensions in question are only *relatively* empty. It is the past experiences of consciousness that serve as criteria for various protentions and anticipations, which in their own turn motivate consciousness to put an accent on specific retentional contents of consciousness. There is an irrefutable difference between such claims as "Stiegler's book is on the sofa" and "Stiegler's book is on Mars." While the former lies within the protentional horizon, the latter exceeds it. The reason for this is that, in Husserl's terms, the former is a motivated possibility, because it stands in associative bond with secondary retentions, while the latter is just an open possibility, for it is unmotivated by secondary retentions. If our protentions and anticipations are grounded in secondary retentions, and if our secondary retentions are individuated, then it becomes understandable why, when different individuals hear the same sentences, they on occasion understand them differently.

In some of his later writings on time, Husserl had argued that primary retentions sooner or later reach the level of intentional zero and no longer color the meaning of present impressions. In Husserl's own words, they sink into the unconsciousness (see Husserl 2014, 62, also Husserl 2006, Text Nr. 13 and Text Nr. 87). We also know, still from Plato, that memories are unreliable: they are finite, for they are bound to the body, which is mortal (see Stiegler 2015, 160, also Stiegler 2019, 109). Against such a background, Stiegler introduces his concept of tertiary retentions. This new concept refers to the exteriorization of memory in technical reality, i.e., to the material inscription of retentions in mnemotechnical mechanisms. Tertiary retentions are material memories, or memories that are outside us. Written texts serve as a paradigmatic example, although as we will soon see, tertiary retentions are not reducible to what is written down.

According to Stiegler, just as protentions serve as criteria for the selection of primary retentions, and just as secondary retentions serve as criteria for the selection of protentions, so also, tertiary retentions serve as criteria for the selection of secondary retentions. It is therefore not just the case that, individually, when we understand, we understand differently. It is also the case that members of specific cultural groups share a common understanding that they don't share with other groups. One of the reasons for this lies in the dominant tertiary retentions. What we remember is largely shaped by various inscriptions: texts, letters, autobiographical narratives, audio recordings, pictures, paintings, etc.

Stiegler's concept of tertiary retentions bears a close resemblance to Husserl's concept of sedimentation. Stiegler acknowledges this resemblance only in his later works.[3] In the earlier writings, he does not employ this phenomenological concept and suggests that the discovery of tertiary retentions overturns the Husserlian framework of thought. Thus, in *Technics and Time, 3*, while addressing tertiary retentions under the heading of recordings, Stiegler contends that "Husserl himself completely neglects the phenomenon of recording in his analysis; in fact, he even excludes it" (Stiegler 2011, 13).[4]

Admittedly, Husserl does not have much to say about audio recordings, yet in such texts as "The Origin of Geometry," he does speak of written recordings and considers them to be of fundamental importance for the establishment of sciences, such as geometry.[5] Husserl's term for such recordings is *sedimentation*, although admittedly, Husserl employs this concept in a variety of senses. As mentioned, Stiegler's concept of tertiary retentions can be easily considered a new term for what Husserl had addressed under the heading of written sedimentations, and Stiegler himself admits this in his later writings. This should not, however, blind us to an important difference between the two thinkers. For Husserl, primary retentions were the foundations of recollections (i.e., what Stiegler calls "secondary retentions") and secondary retentions, in their own turn, were the foundations of sedimentations (what Stiegler calls "tertiary retentions"). Stiegler places primary, secondary, and tertiary retentions in a circular relation. On the one hand, he agrees with Husserl that secondary retentions are founded in primary retentions and that tertiary retentions are founded in secondary. On the other hand, he also contends that tertiary retentions provide the criteria of selection for secondary retentions, which, in their own turn, provide criteria of selection for primary retentions (see, for instance, Stiegler 2011, 17). He further clarifies such an interplay of retentions as the intervention of the imagination at the heart of perception (see Stiegler 2011, 18). This constitutes the core of his philosophical contribution. As Helena De Preester puts it, Stiegler's "whole philosophy is an elaboration of the thesis that human consciousness is essentially *technically*

constituted and conditioned because tertiary retentions are of a technical nature" (De Preester 2021, 105). Or as Daniel Ross has it, tertiary retentions "overdetermine the relationship of primary retention and secondary retention" (Ross 2021, 151).

Granting primary significance to tertiary retentions, Stiegler maintains that writing, understood as tertiary retention, shapes the life of the noetic soul: hominization is exteriorization. Against the background of Derrida's reading of Husserl (see especially Derrida 1989) and Plato (see especially Derrida 1981, 61–171), Stieger repeatedly contends that writing is a *pharmakon*. On the one hand, it is a threat to memory, for the danger is that it can always replace it. On the other hand, it is also a salvation, for it provides the means to reactivate what has been forgotten. Writing is thus both a necessary supplement that continues to feed thought as well as a material exteriority that brings into question the necessity of thinking. Precisely because tertiary retentions are memories outside consciousness, they are essentially *pharmaka*. Thus, on the one hand, writing serves as a necessary condition for the reactivation of what Husserl had called originary intuitions. Yet on the other hand, writing can also replace originary intuitions and make them obsolete. Stiegler takes this to mean that the human condition is pharmacological through and through.

Writing is a crutch of understanding, a supplement to finitude that at the same time opens the possibility of infinitizing. To return to Husserl's "Origin of Geometry," writing is what opens the *infinite we* of geometers that constitute the life of the universal. Without the technology of writing, geometrical insights could be at best passed down a few generations before their extinction. Any science with infinite tasks can only exist against the background of material memory. Yet precisely because writing is a *pharmakon*, exteriorization of which we here speak can take the form of hominization as well as dehominization. Writing can both individuate and dissipate what it had individuated. Such is the pharmacological condition of all tertiary retentions.

While recognizing the close affinity between Stiegler's concept of tertiary retentions and Husserl's concept of sedimentation, it is also important not to overlook the differences between them. Stiegler's concept is further reaching than Husserl's. Stiegler's concept of *literal* tertiary retentions corresponds to Husserl's written sedimentations, as Stiegler himself admits (see Stiegler 2016a, 236). Tertiary retentions of this kind originated around 2,700 years ago. Yet, Stiegler is right to observe that they were preceded by *hypomnesic* tertiary retentions, which originated tens of thousands of years ago, in the Upper Paleolithic. Figurative paintings in the Chauvet cave and Bedolina petroglyph are paradigmatic examples of hypomnesic tertiary retentions. Moreover, even these tertiary retentions are not the oldest in kind. They are preceded

by *epiphylogenetic* tertiary retentions, which refer to the externalization of our memories in tools and they are, according to Stiegler, some three million years old. While relying on Stiegler's vocabulary, one could qualify Husserl's concept of written sedimentations as literal tertiary retentions and thereby distinguish it from hypomnesic and epiphylogenetic tertiary retentions.

Put in phenomenological terms, the claim that human life is guided by tertiary retentions is both eidetic and historical. It is eidetic, for according to Stiegler, a human being is a technical animal. It should therefore not surprise us that nowhere can we come across human existence that is not supported by tertiary retentions. The claim is also historical, for tertiary retentions change over time, not only in terms of their content, but also in terms of their dominant modes, which allows Stiegler to speak of different pharmacological eras. The dependence of human existence on tertiary retentions is what binds humanity together and at the same time sets human beings apart from each other. History and eidetics cross paths in human prosthetics.

It thereby becomes understandable why Stiegler would identify himself as a materialist. "I am essentially my outside" (Stiegler 2009, 30) – this is ultimately what tertiary retentions mean. More precisely, interiority is a result of constitution, "it is something that is constituted through the internalization of a *transitional* exteriority that precedes it" (Stiegler 2013, 20). In this sense, we are confronted here with a materialist philosophy, although in a highly unusual sense: "I remain a materialist, in the sense of a materialism that does not deny the spirit, but which poses that the spirit, while not reducible to matter, is always conditioned by it" (Stiegler 2009, 32).

Besides drawing a distinction between primary, secondary, and tertiary retentions, Stiegler further distinguishes between three kinds of memory: genetic memory, or the memory of the species, epigenetic memory, or the memory of individual experience, and epiphylogenetic, or technical, memory (see Stiegler 2009, 67 ff). While the first two kinds of memory are discussed in biology, to which Stiegler is explicitly referring, the third kind of memory is Stiegler's own addition. While the first kind of memory is transmitted by genes, and while the second kind of memory originates in one's own experience, the third kind of memory is passed down by generations (See Stiegler 2014, 7 and De Preester 2021, 105). Stiegler qualifies the third memory as the memory of the third person, meaning thereby that when one inherits a certain technical object, one also inherits the mode of its use. As we will still see, according to Stiegler, this third kind of memory is slowly disappearing in the industrial age, characterized by mechanical replacements of our technical relation to the world.

7.3 Tertiary Retentions and Individuation

According to Stiegler, different historical eras are characterized by the dominance of different types of tertiary retentions and this has profound effects on the process of individuation. Stiegler draws a distinction between individuation and individualization, arguing that individualization is a result of individuation (see Stiegler 2009, 4). While individualization is a matter of becoming the singular and unique individual that one is, individuation is a social and technically driven process, in which all individuals participate. The process of individuation never comes to an end: "I never cease to find myself *other than myself*" (Stiegler 2009, 4). This process is largely shaped by the interplay of primary, secondary, and tertiary retentions and is selective through and through. The life of consciousness is characterized by the temporal flux of primary retentions. Yet as we saw, for Stiegler, these retentions are selections:

> you make selections which are personal retentions; these selections are made through *the filters of secondary retentions* … conscious life consists of such an organization of primary retentions (R1), filtered by secondary retentions (R2), with the relations of primary retentions and secondary retentions ultimately determined by what I call tertiary retentions (R3).
>
> (Stiegler 2014, 52)

Tertiary retentions thereby turn out to be the pre-individual milieu, or the pre-individual fund, that makes individuation processes possible.

Stiegler's concept of individuation is highly indebted to Simondon's analysis (see especially Simondon 2020 and 2022). For both Stiegler and Simondon, individuation is a process which is both psychic and collective: the formation of the I goes hand-in-hand with the formation of the we. As Robert Hughes has it, individuation "suggests the way one plots oneself in relation to the inheritances offered by the We with which one allies oneself and adopts (or not) as one's own" (Hughes 2014, 55). More precisely, the I, understood as a psychic individual, constitutes itself by adopting a collective history. Adoption allows one to identify with the past, which is not one's own. Moreover, individuation, both psychic and collective, is not a static state of affairs, but a process. Neither the individual, nor the collective, is ever one, but is on the way to becoming one, always in the process of becoming. As we just saw, what enables the process of individuation is the *preindividual milieu*, which is supported by a technical milieu. For Stiegler, tertiary retentions constitute the backdrop of individuation, both individual and collective.

History teaches us that the most significant processes of individuation unfolded against the background of literal tertiary retentions, which Stiegler also on occasion characterizes in terms of grammatization.[6] In this regard, the technological inventions of the written language systems, such as hieroglyphs and the alphabet, are of profound and, arguably, unparalleled significance, for they enable the exteriorization of knowledge and understanding, and they generate tertiary retentions, which can form a much more solid background of individuation than spoken speech. Written language significantly stretches the limits and stabilizes the sedimented background of individuation. It enables individuation against the sociohistorical background of one's own culture.

The technological development of systems of tertiary retentions is of profound importance for the processes of psychic and collective individuation. Allow me briefly to refer to what I consider to be most important. The appearance of the printing press in the 15th century significantly stretches the limits of psychic and collective individuation, for it makes tertiary retentions (such as those inscribed in the Gutenberg Bible) available not only in various households on the European soil, but also in the European colonies. So also, further developments in publishing technology in the 19th century (stereotyping, the iron press, mechanical typecasting and typesetting, etc.) had further profound effects on the processes of psychic and collective individuation; and so did the further technological discoveries in the history of sound recording during the acoustic era in the late 19th century, further followed by the electrical, magnetic, and finally digital eras. This is even more so the case when it comes to another profound technological discovery in the late 19th century, viz., cinema, and even more so when it comes to the discovery of radio and of television. Although a detailed analysis of these processes is not possible in the present context, let me nonetheless stress that all of these technological discoveries have transfigured the fund of tertiary retentions and have reshaped the dominant modes of psychic and collective individuation.

Individuation takes place against the background of tertiary retentions, understood as technical memory. By adopting a history and creatively appropriating the use of tools, one becomes who one is. This makes it understandable why Stiegler would maintain that a human being is a technical animal that constitutes itself prosthetically, i.e., while relying on various technologies at its disposal. Against such a background, we can finally ask: what is specific about individuation in the digital age?

7.4 The Digital and Individuation

Stiegler understands digital technologies as the most recent stage of alphabetic writing, which have been augmented by the decimal system.[7]

He describes the digitalization of our culture as an "intense revolution of tertiary retention of a previously unknown magnitude" (Stiegler 2015, 158). He contends that this revolution has transformed our "entire everyday environment" (ibid.) and has affected the conditions of knowledge and its transmission. This is, indeed, undeniable. In the digital age, only a small segment of what we identify as reality is accessible to us in terms of what phenomenologists had called direct intuitive acts. For the most part, reality *per se* is filtered through technologically shaped images. The distinctive characteristic of the digital age is that the images in question are digital. In pre-digital times, when we wanted to gain access to other segments of reality that were not immediately available to us, we had to either take time to reach them, or, if they lay in the distant past, we could only access them through memory or historical documents. Being situated here and now, we could only see the world from where our lived bodies were situated, for they were the zero points of our orientation in the world. Digital technologies have disturbed many obvious distinctions that characterized the pre-digital age. Presently, even though we can only access a small segment of reality through direct experience, virtually everything and anything can be given to us digitally. The digital thereby lifts consciousness above the here and now and enables it to experience the world from a displaced standpoint. This form of digital displacement, far from being an exception, has become the norm, so much so that very often even what is given to us in our surroundings, is accessed by us in the form of augmented reality.

Of course, what is given to us is not just given in the form of an image, but also as dressed up in narrative forms. Not only are images open to interpretation, but they are passed down with preprogrammed meanings. Quite often, what is given to us are stories that are presented as if they were facts. Since we are given multiple stories that contradict each other, we are also fully aware that these presumed "facts" are not innocent descriptions. However, it is also often far from easy to find a reliable source of evidence to evaluate the truth-value of these digital narratives. For the most part, such a source of evidence simply cannot be found, since the narratives in question take us far beyond the narrow limits of our own personal experience.

According to Stiegler, the digital constitutes a new regime of tertiary retentions. It initiates an unprecedented arrangement of psychic and collective retentions and protentions.[8] The distinctive feature of this new arrangement is that digital tertiary retentions are fundamentally uprooted: they are grounded neither in primary or secondary retentions of individual consciousness, nor in collective history. Rather, digitally processed information is imposed on consciousness algorithmically, from without, and in light-time. The processes of finding secondary retentions in primary

retentions and tertiary retentions in secondary retentions are thereby disturbed. Digital processes take over the business of forming both retentional and protentional contents of consciousness. Prefabricated retentions shape the subject's protentions, and these manufactured protentions are substituted for those that are generated psychically and collectively. Stiegler takes this to mean that, in effect, digital sedimentations interrupt the processes of individualization, both psychic and collective. It is no longer the case that our experiences, recollections, and collective history makes us into who we are. Rather, our identities are defined by the appropriation of digitally generated contents. What are the identities of which we speak here? We should not jump too quickly to the conclusion that the digital revolution has brought about a new form of individuation. According to Stiegler, the most serious problem lies in the fact that the dominance of digital culture threatens the very possibility of individuation. This calls for a clarification.

Among other things, the process of individuation calls for subjective decisions, and any decision takes time. Yet according to Stiegler, this is exactly what digital culture deprives consciousness of. The processes of individuation are too slow for the digital age. Digital technology has the unprecedented means to algorithmically generate protentions and thereby to suspend the time of individuation. In Stiegler's words, the time to dream is taken away from consciousness because the psychic and noetic individual is now preceded by the digital data. Digital data short-circuit desires – that stuff that dreams are made of – and replace them with digitally generated operating sequences. The individual, therefore, no longer has the means to decide who or what they desire to become: in principle, all the options that one faces can be generated digitally and made available to consciousness in light-time. According to what criteria are those choices generated? According to Stiegler, this occurs not while relying on some political authorities, but on marketing, whose governance is not subject to any policy or critique. We are faced here with the reduction of the individual to a consumer. If, as Heidegger contends, existence refers to the specific mode of being of Dasein in that it marks the distinctive possibility to decide to be oneself or not to be oneself, then a consumer does not exist, but only subsists, by satisfying various drives on the basis of content that is handed down to them. Stiegler identifies the reduction of existence to subsistence as the proletarianization of the citizens.

Stiegler identifies the digital with the age of disruption. Disruption, he maintains, "substitutes a blind becoming for the future desired in common" (Stiegler 2019, 21). Driven primarily by marketing, digital disruption ruins the processes of individuation, for it leads to the identification with the behaviors, brands, and labels. The industry of cultural goods lies at the bottom of the formation of tertiary digital retentions. We are faced

here with a technological *epoche*, which brackets all that is genuinely cultural and historical.

Besides disruption, the digital also brings about a new kind of psychic and spiritual congestion, which Stiegler considers to be much more damaging than other traditional form of congestion. This new form of congestion derives from the overflow of information, which in its own turn is triggered by the constant and never-ending supply of new contents for consumption. On the one hand, new technologies need to capture the consumers' attention, which they can do by generating an object of attention that sustains the consumer's retentions. Yet on the other hand, digital technologies capture attention without sustaining it for a long time. The system destroys attention while capturing it, destroys, for it constantly supplements one object of attention for another, thereby making it impossible for any object to remain in focus for a sufficient time. The result is a psychic and spiritual congestion, which in its own turn leads to the diminishing capacities of attention, i.e., in the loss of what Katherine Hayles had called "deep attention" and to its replacement with hyperattention (see Hayles 2007).[9]

The digital brings about an unprecedented control of retentional systems and the transformation of consciousness into a market, which in effect leads to the homogenizing of secondary retentions. This is what Stiegler identifies as the essential cause of what he calls ill-being: "the control of retention implies the loss of identity, that is, of difference…. Our supposedly individualistic societies are in reality perfectly conformist" (Stiegler 2009, 76).

What, then, does it mean to be an individual in the digital age? To be an I under such circumstances does not give one the right to call oneself I, or Ego, just as the community of such I's cannot call themselves "we." With reference to Heidegger's famous analysis of "das Man," Stiegler contends that under such circumstances, one is always targeted as a consumer, and as such, one is always a "they."

7.5 The Digital and Algorithmic Governmentality

The digital brings about a new form of authority, which Stiegler, borrowing an expression from Antoinette Rouvroy and Thomas Berns, identifies as *algorithmic governmentality* (see Rouvroy and Berns 2013). It consists of three essential moments: gathering of data, i.e., datamining; selection and determination of relevant data; anticipation of behavior on the basis of the relevant data (see Rouvroy and Berns 2013, vi-ix, and Stiegler 2016a, 111-112). The anticipation in question is nothing other than the generation of protentions as well as the incitation of needs for consumption. All three steps are completed algorithmically, which in effect means

that all work of reason is short-circuited by digital networks. Algorithms are quicker than reason, and therefore, the production of protentions is still pre-conscious and thus automatic. By gathering data, determining the relevant factors and anticipating future behavior, algorithmic governmentality shapes the individual's needs, while attending to the subject's previous behavior.

We all know that to remember is to forget: memory is selective, and it has to be selective if we are to remember important things. The collection of knowledge in the digital age is no longer guided by selective criteria, i.e., by goals, ideas, or ideals. Rather, datamining has become an automatic and anonymous process: the collection and treatment of data short-circuits all selectivity and normativity. The generation of automatic protentions liquidates conventional categories and normativity. Judgment and the power of decision – this is what the subjects are increasingly deprived of. While "algorithmic governmentality has no need for meanings or significations" (Stiegler 2016a, 130), it is precisely meaning and significance that constitute the individual. As the last section had already anticipated, algorithmic governmentality short-circuits the processes of individuation by replacing them with automatic processes of dividuation, which consists of the reduction of the subject to its digital double. As Stiegler has it, algorithmic governmentality functions on autopilot, by which he means that it functions without any pilots, i.e., without any thinking. Algorithmic governmentality transforms society into an automatic society (identified by Stiegler also as digital society), which relies on computational performativity.

The merging of the information and communication industries has resulted in a near-instantaneous transmission of information. Stiegler calls this "the industrious fabrication of the present: an event becomes and event – it literally takes place – only in being 'covered'" (Stiegler 2010b, 78). This is, however, only partly true. For an event to become an event, besides being covered, it must also be observed, or attended to. This is something that Stiegler himself recognizes and elaborates upon on various occasions. This is of importance, for it leads to the further realization that the information industry must constantly find ways to sustain the attention of the audience if it is to produce events. But how can it win over attention? To do so, it must constantly introduce something new, something unexpected, something irresistible. The information industry does so by appealing to various drives that characterize human existence.

In various works, and in various contexts of analysis, Stiegler draws a distinction between drives and desires. Desire, as Stiegler employs this term, is closely tied to such concepts as love and reason, while drives are essentially irrational. As Robert Hughes rightly contends, Stiegler's writings "generally follow Plato in treating love and desire as very nearly allied with each other and with the quest for knowledge. Consequently, desire,

as Stiegler will declare in a Platonic mood, 'is always desire for knowledge'" (Hughes 2014, 52). While drives lie at the bottom of our basic needs, desire marks the spiritual transformation of drives into goals and ideals that one strives to realize. Stiegler emphasizes that algorithmic governmentality deprives individuals of the capacity to desire. It brings about the "destruction of desire by self-fulfilling anticipations of drive-based fantasies of every kind" (Stiegler 2016a, 125). Desire is destroyed by the automatization that unleashes the drives that trigger herd-like automatisms.

Not only does the digital bring about a new regime of tertiary retentions, which, as we have just seen, are cut off from primary and secondary retentions, but it also brings about a new regime of truth. By this, we are to understand that digitally generated information follows the laws of algorithm, and not any kind of social convention or community wisdom. Stiegler takes this to mean that the algorithmic governmentality is fundamentally a-normative. The result of this is that individuation, understood as the taking over of a normative life, is no longer either possible, or even necessary.

Indeed, in algorithmic governmentality, subjects are avoided and instead of them, attention is paid to their digital doubles. In Stiegler's terms, individuals are disintegrated while attention is paid to "dividuals," understood as the totality of information about the individual that is available in digital form. Within such a framework, subjectivity is rendered obsolete: the subject always arrives too late and cannot take into account either what it is or what it could become. Algorithmic governmentality no longer appeals to subjects, for it focuses on relations. Because these relations are substantialized and considered as formalizable and calculable correlations, they become obstacles to individuation. Algorithmic governmentality gathers unsorted data, and on this basis extracts traits that apply to dividuals (digital doubles). This is what generates the production of automatic protentions that in effect disintegrate psychic and collective individuals. Stiegler identifies this process as the "algorithmic profiling 'behind the back' of those who are being profiled" (Stiegler 2016a, 111), for the knowledge produced is unavailable to the individuals, yet it applies to them. Information about individuals is gathered in advance, and it remotely controls what kind of information they receive in the future.

Algorithmic technologies are technologies of power that automatically produce immanent protentions in a system that is formed by interconnected networks. This power to produce automatic protentions outstrips or overtakes those psychic protentions that could come from psychic and collective secondary retentions. The use of statistical probabilistic knowledge anticipates behavior. What we face here is not just the anticipation, but also the actualization of the possible, which takes place behind the back of consciousness.

7.6 Digital Stupidity

The concept of stupidity crops up frequently in Stiegler's writings. He considers contemporary forms of stupidity to be largely the by-product of analog and digital technologies. "Literal synthesis presupposes that the receiver of a textual message is literate.... With analog and digital technologies, however, the functions of coding and decoding are delegated to machines. The video recorder 'reads' the videotape and the computer 'reads' the file" (Stiegler 2010b, 81). While the understanding of literal sedimentations require one to be literate, this requirement is no longer valid when it comes to the appropriation of information passed down by analog technologies as well as the appropriation and regurgitation of information passed down via digital technologies. The skills required are in both cases minimal. We don't go to school to learn how to watch TV, and a three-year-old child often knows better than the adult how to handle a smartphone. Of course, one would be right to observe that reading a hypertext also requires one to be literate. Nonetheless, the understanding of many hypertexts, including the most popular ones, does not call for much learning, and, as Alexander Pope has it, a little learning is dangerous.

Analog and digital technologies prepare a fertile ground for the spread of misinformation, manipulation, and demagogy. They unavoidably place the consumers into information bubbles: search-engine filter bubbles, content-recommendation bubbles, social media bubbles, etc. For technical reasons, an escape from these bubbles has become not only difficult, but also unlikely, for algorithms keep track of our online activity and filter our access to content that they calculate we will like. Information bubbles are filter bubbles: by personalizing our online activity, they at the same time isolate us from the content that we haven't expressed an interest in. Whether we want to or not, we end up living our lives in information bubbles from which it has become difficult to escape. Information bubbles have become unavoidable.

One would be right to remark that much of what Stiegler says about the effects of digital technology also applies to analog technology, and especially television. Nonetheless, he is also attentive to the important differences between them. While analog technologies, constituted as a dissociated technical milieu (i.e., producer-consumer opposition), provide a medium for imposing information to be taken up and shared privately with others, digital technologies, constituted as an associated technical milieu, allow the consumers to take an active role in their own stupidification: besides absorbing the misinformation that is handed down to them, they can spread it further to others, thereby creating information bubbles. Here, we are faced not only with technologically manufactured information bubbles, but with bubbles that are of our own making.

Alfred Schutz had famously claimed that we live our lives in multiple realities (see Schutz 1962, 207–259). By this he meant that the psychic life unfolds in the mode of leaping from one finite province of meaning to another. From the reality of everyday life, which Schutz understands as the reality in the paramount sense, one leaps into other finite provinces of meaning, such as the world of phantasms and dreams, of art and scientific theory, of virtuality reality and religion, etc. As far as our digital lives are concerned, this is true in one sense, false in another. Thus, on the one hand, after reading the news, I can watch a video on YouTube, then check my social media page, then check the weather in a distant city, then run a search on whatever comes to mind, etc. Yet on the other hand, and despite the possibilities of leaping from one finite province of meaning to another, the information content that I access is always already filtered by my previous online activities. Whether we want to or not, our digital lives are characterized by social fragmentation and group polarization. My digital "history" mark off the boundaries of what I can access, and there are plenty of other finite provinces of meaning that lie beyond my reach. They are on the other side of my consciousness: beyond my understanding and accessibility.

7.7 Digital Sedimentations

I argued above that Stiegler's concept of tertiary retentions bears a close resemblance to Husserl's concept of sedimentation, even though Stiegler does not acknowledge this resemblance as clearly as he should, arguing instead that tertiary sedimentations disrupt the framework of Husserlian phenomenology. The foregoing analysis provides us with the resources needed to speak of specifically *digital* sedimentations, and in more senses than one.

First, in the most obvious sense, digital sedimentations are digitally formed tertiary retentions that consciousness gains access to through the digital medium and that are imposed on consciousness from without and at light speed, short-circuiting its memories and personal history. In this sense, the concept of digital sedimentations is an extension of that concept of written sedimentations that we come across in Husserl's "The Origin of Geometry." What is at stake here are digital sedimentations that are written in "digital language" and that are there at one's disposal, whether one attends to them or not. In this regard, also, the concept of digital sedimentations is equivalent to Stiegler's concept of digital tertiary retentions. In a significantly different sense, the concept of digital sedimentations also refers to one's own digital history, to one's own "digital traces," viz., to one's own online activity, which in principle can never be erased and which serves as a benchmark for filtering present and

future content. Our access to any digital content is significantly constrained by digital sedimentations of these two fundamentally different kinds, and the more our understanding of the world becomes dependent on digital media, the deeper the problem becomes. One can only conjecture that, with time, the world will become more segregated and all of us will face increasing difficulties in understanding those with a different background and with a different set of interests and beliefs. Gaining access only to that information content which confirms our believes and complies with our interests, we will face few possibilities to bring into question our accepted biases. Finding ourselves in echo chambers, we will progressively lose the capacity of self-critique. Intellectual isolation within ideological frames and algorithmically manufactured relevance systems, resulting from personalized searches, recommendation systems, and algorithmic curation, is the state of stupidity that we find ourselves in and that we have to resist at all costs.

To distinguish these two different kinds of digital sedimentations from each other, we can identify the first as tertiary digital sedimentations and the second as secondary digital sedimentations. I refer to the latter group as secondary digital sedimentations because of the link that binds these sedimentations to recollection: in the present case, we are faced with digital sedimentations that have originated in our own activities and that can, in principle, be remembered. While tertiary digital sedimentations are imposed on consciousness from without, digital secondary sedimentations originate in our own digital activities. While the former refer to what is fundamentally external to one's personal and collective life, the latter refer to one's own past digital activities, which lie beyond the immediate grasp of consciousness. One could further distinguish these two kinds of digital sedimentations from a third kind, which one could identify as primary digital sedimentations, and which would refer to the conscious digital activities that are still within the grasp of the digital living present and that have not yet reached the night of the digital unconscious. In this particular case, we would be faced with digital traces that still have not evaporated from one's actual grasp. Recall Stiegler's claim that, in contrast to analog technology, which is characterized by a dissociated technical milieu, digital technology is characterized by associated technical milieu, meaning thereby that digital technology is no longer based on the producer/consumer opposition, but allows, and on occasion even requires, that consumers take an active role. While tertiary digital sedimentations refer to what is imposed on consciousness from without, primary and secondary digital sedimentations refer to these activities on the part of the consumer. While secondary digital sedimentations refer to the past activities, which escape the immediate grasp of the consumer's awareness,

primary digital sedimentations refer to what still lies within the field of an ongoing digital activity. This immediate digital past determines the presently given digital content in an especially robust way.

7.8 Nihilism and Disintoxication: Stiegler's Project of Nachstiftung

Stiegler's writings invite us to reflect on the relation between automatism and autonomy in a digital society, which he also calls automatic society (Stiegler 2016a, 238). According to Stiegler, our current condition can be characterized by automatism overtaking autonomy. Admitting that the concepts of *Gestell* and nihilism had come to dominate his work in recent years (see Stiegler 2019, 62), Stiegler identifies our current condition as the completion of nihilism. While considering the digital to be a technology of the mind (Stiegler 2016b, 158), Stiegler defends the legitimacy and necessity of a philosophical struggle "against what, in technics or technology, is toxic for the life of the mind or spirit" (Stiegler 2013, 20). What is especially needed, he contends, is disintoxication, which in its own turn should lead to the battle against digital disindividuation. For Stiegler, the solution to contemporary nihilism lies in the commitment to the Kantian principle of the subjective differentiation of fact and law and in the establishment of the state of law, which would in effect mean the establishment of political authority over algorithmic governmentality and the subjection of the regime of digital sedimentations to a rational normality. The emergence of the digital calls for a political response, for in its absence, the process of individuation becomes industrialized. Only a political response to algorithmic governmentality could make the reconstitution of a collective individuation possible, which in its own turn would enable the psychic individuation of consciousness (see Stiegler 2016a, 102–103). But what could such a political response to algorithmic governmentality be based on?

Recall that, for Stiegler, the digital age is an age of hypomnesis constituting itself as an *associated* technical milieu. This means that the digital age is no longer characterized by the sharp division between the producer and the consumer that characterizes analog technology. Rather, in the digital age, the consumer takes on the active role of co-producer. As we saw, Stiegler sees in this co-activity on the part of the consumer a source of profound danger. "But where danger is, grows the saving power also." This reference to Hölderlin's words, made famous in Heidegger's essay on technology (see Heidegger 1977, 28), is fully appropriate in the present context, for Stiegler's response to the challenges brought about by the digital culture unfolds according to Heideggerian line.[10] The very fact that the digital age calls for an active response on the part of the consumer allows

one to inquire into what a meaningful response to the challenges brought about by digital technology would look like.[11]

To begin with, such a response cannot lie in a return to a pre-industrial age. Resisting all forms of naïve romanticism, Stiegler insists that there is no post-industrial future, just as there is no future in returning to the pre-industrial age. Rather, "the future lies in re-founding the industrial model" (Stiegler 2015, 171). For Stiegler, this project of re-founding, which in many ways reminds one of Husserl's project of *Nachstiftung*, as presented in such works as *The Crisis of European Sciences and Transcendental Phenomenology* (see, for instance, Husserl 1970, 14), must be based in an ethics of care. Thinking of the digital as a *pharmakon*, Stiegler considers it to be both toxic and curative. According to Stiegler, "the digital, which intensifies to an immeasurable extent this *outstripping and overtaking*, and the *abandoning of critique* to which it leads, *might* yet constitute in law a new age and a new democratic promise" (Stiegler 2016a, 102). As he further notes, the digital represents a "pharmacology of unprecedented dangerousness," yet it also includes elements bearing the potential for a new structure of society (see Stiegler 2016a, 134). The great challenge of the 21st century is "to invent new modalities of non-inhuman existence within societies that have become thoroughly technological" (Stiegler 2010a, 183).

We are faced here with a grand project, yet the mere fact that it is unlikely to complete it does not mean that it is meaningless to undertake it. I would like to conclude by briefly indicating one specific way in which Stiegler has aimed to contribute to this project of *Nachstiftung*. It concerns the politics of attention.[12] As mentioned, one of the most serious problems that we confront in the digital age is that digital technologies are robbing consumers of the possibility to develop what Katherine Hayles had called "deep attention."

> Now, at the beginning of the twenty-first century, cell phones, networked computers, and the many other devices … catalyze ruptures in the juvenile psychic apparatus's synaptic circuitry, to the advantage of *hyperattention* and the detriment of *deep attention*, creating an intergenerational fracture on the cerebral level, that is, within the organology of the central nervous system itself.
>
> (Stiegler 2010a, 176)

The destruction of deep attention amounts to disindividuation, for it robs one of the possibility to form oneself as an individual.

> The work of forming attention undertaken by the family, the school, the totality of teaching and cultural institutions, and all the apparatuses of 'spiritual value' (beginning with academic apparatuses) is

> systematically undone in the effort to produce a consumer stripped of the ability to be autonomous either morally or cognitively.
> (Stiegler 2010a, 184)

What is much needed under such circumstances is the refounding of the conditions for the development of deep attention and here, according to Stiegler, education, understood as a form of care, has to play a central role.[13] For Stiegler, care is fundamentally social: the care one takes of oneself is always care one takes of others (see Stiegler 2010a, 178). Knowledge, passed down by others, is what individuates and transforms the learner, enabling them to interiorize collective tertiary retentions. The reconstitution of deep attention is not just one facet among others in Stiegler's project of *Nachstiftung*. It is granted transcendental significance: in order to resist technological reductionism, one must first heal the powers of attention. In such a framework, education obtains emancipatory significance. In Stiegler's words, education is the *otium* of the people.[14] Its fundamental task in the twenty-first century must be that of forming in the learners the capacity of deep attention: before it is anything else, education is a matter of attention formation.

7.9 Conclusion: Stiegler and Phenomenology

Stiegler's writings provide clear proof that some of the central phenomenological concepts are of great importance in the framework of contemporary reflections on digital culture, even though classical phenomenologists in whose works these concepts originate were concerned with significantly different issues in their own writings. Yet to what degree can Stiegler's analysis of digital technology be qualified as phenomenological? On the one hand, the foregoing analysis makes clear that Stiegler's thought is undeniably inspired by classical Husserlian and Heideggerian phenomenology. In Stiegler's writings, we encounter a number of classical phenomenological concepts and themes, starting with retention, moving on to the crisis, inauthentic existence (subsistence) of the They, and ending with care. Yet on the other hand, it would be a mistake to measure Stiegler's contribution according to the yardstick of classical phenomenology, be it Husserlian or Heideggerian. Rather, Stiegler's philosophy of technology provides a good example of how fruitful classical phenomenology proved to be beyond phenomenology.

Hans-Georg Gadamer has famously noted that in his teaching, whenever he encountered grand assertions and impressive constructions, Husserl used to say, "Not always the big bills; small change, small change" (Gadamer 1977, 133). This is what, as Gadamer further remarked, "produced a peculiar fascination. It had the effect of a purgation, a return to

honesty, a liberation from the opaqueness of the opinions, slogans, and battle cries that circulated" (ibid). In Stiegler's writings, we seldom encounter small change, and much more often the big bills. It is his grand assertions that produce a peculiar fascination on the part of the readers. As far as Stiegler's relation to phenomenology is concerned, this motivates one to raise two fundamental questions. First, what is the source of evidence that supports Stiegler's conclusions and to what degree, if at all, can his conclusions be corroborated phenomenologically? And second, does the digital lifeworld, which is analyzed in such a fascinating way in Stiegler's works, not require phenomenology to rethink, revise and revamp its methodology so as to make it more suitable to analyze the world we live in and the new forms of crises brought about by digital technology? It is not possible to address these fundamental questions in the present context. Suffice it so say that Stiegler's engagement in phenomenology is heterodoxical and therefore, his relation to phenomenology is problematic, and for this reason especially, his writings constitute a much-needed provocation and a welcome challenge to phenomenology.

7.10 Funding

This work was supported by General Research Fund (GRF) Grant. Granting Agency: Research Grants Council (RGC) from the Research Grants Council University Grants Committee (Grant number: 14604923).

Notes

1 "After reading and re-reading two short texts – Husserl's 'Origin of Geometry' and Derrida's introduction to his French translation of Husserl's text, I began to simultaneously, passionately and systematically study the Greeks, and firstly Plato and Aristotle, as well as everything by Husserl that had been translated into French – being unable to read German" (Stiegler 2019, 60).
2 Stiegler does not provide an account of what happens to other impressions, which, presumably, consciousness does not retain, and here, an alternative account is certainly possible: consciousness might very well retain all impressions, while at the same time, it can consider some to be more relevant than others.
3 In *Prendre soin: De la jeunesse et des générations*, which was originally published in French in 2008, Stiegler explicitly maintains that "tertiary retentions are the sedimentations that accumulate across generations and that are central to the process of creating collective individuation" (Stiegler 2010a, 6).
4 A little later in his analysis, however, Stiegler further contends that tertiary retention is "a memory Husserl designates as *consciousness of image*" (Stiegler 2011, 17). Yet this is a stranger observation, and it is hard to justify it in light of Husserl's own texts.
5 For this reason, it is very problematic to claim, as Daniel Ross does, while commenting on Stiegler's relation to Husserl, that "Husserl distinguishes primary

retention from secondary retention but dismisses the possibility that external artifacts ('resembling objects' of 'image-consciousness') make any significant contribution to the temporality of experience" (Ross 2021, 151). While this claim echoes Stiegler's view, it is hard to justify it in light of Husserl's own texts.

6 Stiegler's concept of grammatization is an extension of Derrida's concept of the supplement. While the concept itself is borrowed from the linguist Sylvain Auroux, it is also extended beyond the sphere of language so that, as Daniel Ross remarks, "*grammatization* refers in Stiegler to every technical evolution of the processes by which temporal flows are spatialized, in so doing rendering them discrete and reproducible" (Ross 2021, 150).

7 "Electronic, transcoded according to the ASCII standard into binary data, alphanumeric writing can be subjected to the rules and algorithms of Boolean algebra, and thereby calculated and processed at the speed of light" (Stiegler 2015, 159).

8 Stiegler even contends that the digital constitutes a new regime of truth (Stiegler 2016a, 98). By this, Stiegler means that in principle, the digital is capable of producing any truth whatsoever, as we can witness in the post-truth age of populism.

9 "Deep attention, the cognitive style traditionally associated with the humanities, is characterized by concentrating on a single object for long periods (say, a novel by Dickens), ignoring outside stimuli while so engaged, preferring a single information stream, and having a high tolerance for long focus times. Hyper- attention is characterized by switching focus rapidly among different tasks, preferring multiple information streams, seeking a high level of stimulation, and having a low tolerance for boredom" (Hayles 2007, 187).

10 For Stiegler's relation to Heidegger's thought, see Ross 2021.

11 Fitzpatrick remarks that here we encounter an important difference between Stiegler and Derrida and their respective pharmacologies: "Whilst Derrida sets out to give an exhaustive deconstructive account of the pharmakon, both its toxic, nefast and dangerous aspects as well as the curative, virtuous and positive aspects, he never really develops the positive aspects" (Fitzpatrick 2020, 353). By contrast, in many of his later writings, Stiegler provides us with various reflections on the positive side of digital *pharmaca*.

12 Helena De Preester has recently maintained that "Stiegler's view on attention is rooted in Edmund Husserl's account of the (temporal) structure of consciousness" (De Preester 2021, 104). So also, Noel Fitzpatrick rightfully remarks that "the question of attending, or giving attention to, or being called to attend to, is a question at the very core of phenomenology" (Fitzpatrick 2020, 349). Indeed, for Stiegler, attention is a matter of accumulating and holding primary and secondary retentions together with protentions.

13 Besides education, meditation can be considered another viable counterforce to the destruction of attention by digital technologies, as Helena De Preester has maintained in a recent contribution (see De Preeser 2021).

14 *Otium* in the present context refers first and foremost to free time, meaning thereby time that is not devoted to business or subsistence.

References

De Preester, H. 2021. "Life Is What You Fill Your Attention With – the War For Attention and the Role of Digital Technology in the Work of Bernard Stiegler.". *Phenomenology and Mind* 20: 102–16.

Derrida, J. 1981. *Disseminations*. Chicago: University of Chicago Press.

Derrida, J. 1989. *Edmund Husserl's "Origin of Geometry." An Introduction*. Trans. by John P. Leavey. Lincoln: University of Nebraska Press.
Fitzpatrick, N. 2020. "Questions Concerning Attention and Stiegler's Therapeutics." *Educational Philosophy and Theory* 52 (4): 348–360.
Gadamer, H.G. 1977. *Philosophical Hermeneutics*. Trans. by D. E. Linge. Berkley, Los Angeles and London: University of California Press.
Hayles, K. 2007. "Hyper and Deep Attention: The Generational Divide in Cognitive Modes." *Profession*: 187–99.
Heidegger, M. 1977. *The Question Concerning Technology and Other Essays*. Trans. by William Lovitt. New York & London: Garland Publishing, Inc.
Hughes, R. 2014. "Bernard Stiegler, Philosophical Amateur, or, Individuation from Eros to Philia."." *Diacritics* 42 (1): 46–67.
Husserl, E. 1970. *The Crisis of the European Sciences and Transcendental Phenomenology*. Trans. by David Carr. Evanston, IL: Northwestern University Press.
Husserl, E. 2006. *Späte Texte über Zeitkonstitution (1929–1934). Die C-Manuskripte*. Ed. by Dieter Lohmar. Dordrecht: Springer.
Husserl, E. 2014. *Grenzprobleme Der Phänomenologie: Analysen Des Unbewusstseins Und Der Instinkte. Metaphysik. Späte Ethik*. Ed. by Rochus Sowa and Thomas Vongehr. New York: Springer.
Kouppanou, A. 2021. "'…In the Secret of One's Life': Bernard Stiegler and Philosophy in the Intimacy of His Prison Cell." *Journal of Prison Education and Reentry* 7 (2): 123–29.
O'Shiel, D. 2022. *The Phenomenology of Virtual Technology: Perception and Imagination in a Digital Age*. London & New York: Bloomsbury Academic.
Ross, D. 2021. "Care and Carelessness in the Anthropocene: Bernard Stiegler's Three Conversions and Their Accompanying Heideggers." *Cultural Politics* 17 (1): 145–62.
Rouvroy, A., and T Berns. 2013. "Algorithmic Governmentality and Prospects of Emancipation: Disparateness as a Precondition for Individuation Through Relationships?." Trans. by Liz Carey-Libbrecht. *Réseaux* 177 (1): 163–96.
Schutz, A. 1962. *Collected Papers I: The Problem of Social Reality*. edited by Maurice Natanson. The Hague: Martinus Nijhoff.
Simondon, G. 2020. *Individuation in Light of Notions of Form and Information*. Trans. by Taylor Adkins. Mineapolis & London: University of Minnesota Press.
Simondon, G. 2022. *Imagination and Invention*. Trans. by Joe Hughes and Chrstophe Wall-Romana. Minneapolis & London: University of Minnesota Press.
Stiegler, B. 1998. *Technics and Time, 1: The Fault of Epitemtheus*. Trans. by Richard Beardsworth and George Collins. Stanford, CA: Stanford University Press.
Stiegler, B. 2009. *Acting Out*. Trans. By David. Barison, Daniel Ross, and Patrick Crogan. Stanford, CA: Stanford University Press.
Stiegler, B. 2010a. *Taking Care of Youth and the Generations*. Trans. by Stephen Barker. Stanford, CA: Stanford University Press.
Stiegler, B. 2010b. "Memory." In *Critical Terms for Media Studies*, edited by W.J.T. Mitchell and Mark Hansen, 64–87. Chicago, Ill: The University of Chicago Press.
Stiegler, B. 2011. *Technics and Time, 3: Cinematic Time and the Question of Malaise*. Trans. by Stephen Barker. Stanford, CA: Stanford University Press.
Stiegler, B. 2013. *What Makes Life Worth Living: On Pharmacology*. Trans. by Daniel Ross. Malden, MA: Polity Press.

Stiegler, B. 2014. *Symbolic Misery: Volume 1: The Hyperindustrial Epoch*. Trans. by Barnaby Norman. Malden, MA: Polity Press.
Stiegler, B. 2015. *States of Shock: Stupidity and Knowledge in the 21st Century*. Trans. by Daniel Ross. Malden, MA: Polity Press.
Stiegler, B. 2016a. *Automatic Society. Volume 1: The Future of Work*. Trans. by Daniel Ross. Malden, MA: Polity Press.
Stiegler, B. 2016b. "The Digital, Education, and Cosmopolitanism." *Representations* 134: 157–64.
Stiegler, B. 2019. The Age of Disruption. Technology and Madness in Computational Capitalism. Trans. by Daniel Ross. Malden, MA: Polity Press.

8 The Imaginary, Magic, and Hypervirtuality

On the Phenomenological Nature of Digital Screens

Daniel O'Shiel

8.1 Introduction

With us spending more time in front of screens than ever before, it is a good moment to better understand their phenomenological nature, not least with reference to their imaginary structure, inherent magic, and the tendency to addiction in the form of a kind of "hypervirtuality."

This chapter consists of five main sections. First, I summarize the phenomenological natures of perception, phantasy, and image-consciousness (*Bildbewusstsein*), following analyses of certain works by Husserl, Fink, and Sartre. Second, on the back of these experiential distinctions, I will argue for and explain two main types of "virtuality," one on the horizonal perceptual level and the other a more specific type of "digital virtuality." I then focus a bit more on Sartre's concept of the imaginary and its inherent magic, linking this to the preceding, not least digital virtuality. This will lead, fourth, to an explication of the phenomenological nature of digital screens, which open up digital universes of experiences for us while at the same time blocking or taking us away from our more immediate perceptual surroundings and realities. Because of their inherent magic, they can also be highly addictive, which I will characterize, fifthly and finally, as a type of "hypervirtuality."

The chapter will thus give quite an initial account of the phenomenological experiential categories one needs to consider when capturing the nature of our digital screens, as well as show how they are of an essentially imaginary and magical structures that harbor a powerful type of digital virtuality which can ultimately be highly addictive and even damaging for us. Beginning the conversation on all of these elements thus will help us understand our relation to screens and their irrealities more, as well as the enduring significance of our more immediate and basic perceptual realities.

DOI: 10.4324/9781003312284-9

8.2 Perception, Phantasy, and Image-Consciousness in Husserl, Fink, and Sartre

In Husserl, perception both presents objects *tout court* and yet, strictly speaking, one is only ever directly aware of various "aspects" (*Abschattungen*) at any given moment. Even given this detail perception is always about a *physical presence*; as experienced in our everyday, naïve, pre-reflective natural attitude, an actually present, physical thing is given immediately and directly to us, the perceiver:

> [T]he <natural> experience that is presentive of something *originarily* is *perception*, the word being understood in the ordinary sense. To have something real given originarily and "attentively to perceive" and "experience" it in an intuiting simpliciter are one and the same. We have originary experience of concrete physical things in "external perception[.]"
>
> (Husserl [1913] 1983/2009, §1.)

Perception is thus always about present objects – "things"; it is an essential tenet of this mode of consciousness. This, however, also brings other laws with it. Indeed, that perception has to do essentially with physically present objects also means such objects are necessarily experienced as *real*: "[p]erception makes a present reality appear to us as present and as a reality" (Husserl [1898–1925] 2005/1980, 4/4). The German term for "reality" here is *Wirklichkeit*, which is elsewhere rendered as "actuality." And indeed, Husserl does seem to use the two terms interchangeably.

Generally, this characterization means that one cannot but believe in the actual reality of things in the world:

> With the idea of *actuality* we stand in the system of thetically *unmodified* intentionality, in the intentionality of doxa, of belief. Belief is not something appended to presentations [*Vorstellungen*], not a *feeling* associating itself with them, not a way of being affected, now present, now absent, attending such presentations; on the contrary, it is the *unmodified consciousness itself*.
>
> (ibid., 670/557–558)

This "belief" is not propositional and reflective ("I believe that this cup is real"), but quite automatic, pre-reflective, and instinctive. The base mode of perception necessarily constitutes and contains an automatic recognition of actuality, reality. In other words, *Wirklichkeit* belongs to the very experiential fabric of perceptual consciousness. In Husserl's terms, this most basic form is the "unmodified" (or "unmodalized") mode of

perceptual consciousness which, by its very nature, is one of automatic belief in the independence and durability of things even in the face of conscious variation and absence. Further, it is a *certain* belief (cf. ibid., 264/215) that is also characterized as necessarily *impressional* (cf. ibid., 267/218) in a Humean sense (i.e., impressions or perceptions as opposed to ideas or images). To give a simple example, when I walk out of a room, I implicitly and impressionally know that this room will continue to exist without me perceiving it; it is in no need of me at all, it follows causal natural laws in a complex physical nexus, and it will also be there to be perceived again in a regular manner when I go back.

There can be all sorts of "modalizations" as Husserl terms them, which is to say modifications (including illusions and hallucinations – cf. Smith 2002), but this can only happen on the back of a more primordial and basic form of experiencing reality, much like recognition of dreams *as* dreams can only happen on the back of an extended, temporal waking life. In sum, perception is the most basic and primordial form of consciousness we know, and has primarily to do with real, present things. This is encapsulated in Husserl's recurring terms for it: original or originary consciousness (for instance: Husserl [1918–1926] 2001/1966: 40/4).

On top of this, perceptual experience in Husserl is necessarily inexhaustible, situated, perspectival, synaesthetic, and kinaesthetic, embodied in spacetime, and horizonal. These are the essential characteristics of Husserlian perception. On the other side, the case is very different with our phantasies. "Phantasy" (*Phantasie*) is opposed to perception in Husserl because the former has to do with absent, irrealized and non-real or non-actual objects – non-things. In a word, "[c]onsciousness of what is *not present* belongs to the essence of phantasy" (Husserl [1898–1925] 2005/1980, 63/58–59). This also means that phantasy does not posit the necessary and certain existence of its objects; it lacks the doxastic mode of belief. I can of course phantasize real, existent objects (e.g., this chair), but I do not have to, and even if I do I know they are not actually there before me when I close my eyes, like in a case of physical and actual perception right in front of me. In this manner, it belongs to the very structural fiber of perception to posit its objects as existent; in phantasy not. Quite the contrary, in fact: the essential characteristic of phantasy is its *as-if* character – it is *as if* this unicorn could exist, but ultimately it does not, at least not like a perceptual thing does.

There are of course differences in content here [a chair or my friend is real; a unicorn is not (yet)], as well as degrees of verisimilitude, immersion, and captivation. Nevertheless, the general point remains: phantasy, with its inherent capacity to neutralize and irrealize reality, opens up the realm of the inactual as such. Husserl at one point calls it "*inactuality* consciousness" (*Inaktualitäts*bewusstsein – ibid., 359–360/299; cf. also

Jansen 2013, 68). This indeed is the crucial point: Husserl notes well that there is a "*protean* character of phantasy" (Husserl [1898–1925] 2005/1980, 65/60) that is excluded from perception – this latter has essentially more stability and solidity (see ibid., 34/33). In short, there is a discontinuous, chop-and-change, staccato-like element to phantasy that is not present in perception. Here, although one can of course have rather fixed, stable, and recurring phantasies, and although one can have very intense phantasies where one even momentarily forgets their as-if quality (dreams, illusions, and hallucinations are the prime examples), if one is to remain rather sane and retain a difference between one's perceptions and phantasies, this as-if dimension is always structurally present in the latter, no matter how powerful or immersive it might be in other respects.

Thus, in the usual walks of life, the basic difference between perception and phantasy means they mutually exclude each other at any given moment. In other words, one cannot perceive and phantasize at the same time. Try it: try to perceive and imagine a chair simultaneously. I think one finds one either looks at and actually feels (perceives) the chair; or one imagines it (here closing one's eyes helps, which is another tell-tale sign); or one kind of hesitates between the two without really doing either. The two acts are mutually exclusive at any given moment, even though on a broader, more conceptual level, they imply each other at every turn.

What is crucial here is that this ultimately quite stark opposition creates problems for Husserl's own account of "image-consciousness" (*Bildbewusstsein*), which seems, somewhat contradictorily, to contain both perceptual (real) as well as phantasized (irreal) elements in one seemingly unified mode of consciousness. Husserl has many expressions for this type of consciousness, with two of the most striking being "perceptual imagination" (cf. Husserl [1898–1925] 2005/1980, 86/79) and "physical imagination" (cf. ibid., 90/83). Based on the current analyses, this should strike one as oxymoronic; how can one have a mode of consciousness that has *both* elements of the perceived (presence, reality, and the like) *and* elements of the imaginary (absence, irreality, and the like) when the two are seemingly opposed in their phenomenological natures?

Husserl gives a complex answer that has long-lasting tensions and consequences (cf. O'Shiel et al. 2019a). In a nutshell, it is characterized as a distinct mode of consciousness that seems to contain elements that are both there (are real, perceivable) and not (are irreal, phantasizable). When I look at a photograph, for example, there are actual, perceived elements that are evidently there before me (the physical photo); and yet the photo also depicts something or someone that is decidedly not there, perhaps never was, and perhaps can never be (again) (e.g., my grandmother).

Ultimately, Husserl's image-consciousness has a structure of its own, with three essential components that nonetheless always already interlock in any lived experience:

> 1) [T]he physical image, the physical thing made from canvas, marble, and so on; 2) the representing or depicting object; and 3) the represented or depicted object. For the latter, we prefer to say simply "*image subject*"; for the first object, we prefer "physical image"; for the second, "representing image" or "image object."
>
> (Husserl [1898–1925] 2005/1980, 21/19.)

Husserl is talking about a painting or sculpture here, but this basic structure holds for other media from Husserl's time (e.g., a photograph), plus a whole host of media (televisions, computer screens, smartphones, VR headsets) that came after him. Actually, I think it should be quite clear that this form of consciousness can extend beyond the primarily visual. For example, certain physical soundwaves (physical image) are sensed and heard (image object) in a manner that gives one intentional access to a certain song (image subject – "Yellow Submarine"). In this manner, the latter is only made present *through* the specific, physical version one is hearing.

Image-consciousness is thus external physical phenomena stimulating one or more senses to experience an object that is not properly and fully there in a perceptual manner. And indeed, whether primarily of vision or easily extendible to other senses, the structure of image-consciousness, as has just been quoted, *always* has a tripartite structure for Husserl: the *physical image* (*physische Bild*) is the physical matter involved, and places (at least part of) the experience squarely in the realm of the perceivable; second, the *image object* (*Bildobjekt*) is all the sensuous experience stemming *from* the physical image; and third, the *image subject* (*Bildsujet*) shows that one transcends what is immediately before one towards something that is not perceptually there. If the physical image is clearly in the realm of the perceivable, and the image subject is only accessible thanks to our capacity for phantasy – i.e., it is decidedly not present – then the second, middle element (the image object) has the most ambiguous status.

What is this mode of consciousness, then? It is clearly not a straightforward case of perception, because in perception, I do not transcend toward something that is only virtually there. Indeed, with images, digital, or otherwise, it is usually quite impossible to perceive the supposedly "pure sense data." In short, one normally always already sees an image *as an image*, and not the supposedly "pure" perceptual colors, pixels, or the like. Moreover, even if this latter were possible, it would take a conscious effort or kind of abstraction, which means this is not the automatic and most natural way that we encounter and deal with these phenomena. In

a word, on a working screen we do not perceive pixels but automatically see images.

This mode does not seem, however, to be a straightforward case of phantasy, either, because in this latter, I can simply close my eyes and imagine something without any reliance upon a screen. For Husserl, then, straightforward – or at least "pure" – cases of phantasy do not seem to require a physical image, and maybe not even an image object.

Perhaps image-consciousness is simply a hybrid then, a mixture between perception and phantasy, where to call it a bit of both or neither – in the latter case, it would then be a structure or mode of consciousness in its own right– could both be acceptable characterizations. For now, I think it is fair to decide it as a form of consciousness with a unique tripartite structure different from both perception and phantasy, and actually finding its conceptual and structural home between the two. The discussion, however, continues with two other phenomenologists, who present contrasting ideas as to the fundamental nature of image-consciousness.

With Fink, one of the main goals of "*Vergegenwärtigung und Bild*" ("Presentification and Image"; Fink [1930] 1966) is to categorically distinguish our capacity for phantasy from our ability to see images and pictures in various artefacts of the cultural world. Sculptures, paintings, and photographs were of course the paradigmatic examples of Fink's time; now, image-consciousness can apply to much, much more, not least digital games.

For Fink, image-consciousness is ultimately a unity of two, not three, essential components: consciousness of the "*real 'carrier'*"[1] (ibid., 73) and the '*image world*' (ibid.). These aspects are fused in our lived experiences of the representational artefacts, but can be phenomenologically analyzed. The image world "is always and essentially together with a real carrier" (ibid., 74), be this the carved marble, painted canvas, pixeled colored computer screen or whatever. This is already the essential difference between image-consciousness and phantasy for Fink; the carrier means there is an external physical element to image-consciousness that phantasy does not have and that thereby makes the former more determined – not least physically – than the latter could possibly ever be. Moreover, the carrier qua carrier is essentially overlooked in the normal functioning of image-consciousness (cf. ibid.); indeed, it is all that real physical material that actually *must* be overlooked if it is to portray something that is not actually there, as we have already seen as intimated in Husserl. In this manner, it is "a certain anomaly if this carrier comes itself into thematic view" (ibid.), for instance a broken or malfunctioning screen. In short, the carrier is that anonymous "plain reality insofar as it is *covered* with the image world" (ibid.) – it localizes and fixes the image world in a particular, physical thing

in the real, perceptual world, but also allows you to go automatically beyond – transcend – this latter.

The image world is obviously the second essential element, that which the physical materials portray. Together, the two components are always already fused in the phenomenal experience, making these experiences so many "windows" (cf. ibid., § 34) into relatively determined image worlds.

For Fink, it is crucial to remember that image-consciousness is an essentially different structure than that of phantasy; the former has a physical carrier that allows for determined windows into a pre-constituted image world which, precisely because the latter is only facilitated through a physical carrier, is much more determined and concrete, like a certain painting, a picture of my grandmother, a computer game on a screen, or even a Facebook page. In Fink's words, "[t]he unreality of an image world can essentially only be an abstract moment of a determined reality" (ibid.). This means for Fink, the realm of image-consciousness is between the realms of pure perception and phantasy. However, because of its physicality, ultimately image-consciousness is a special type of perception, one of physical, external, and therefore determined images which nonetheless also give you access to transcendent objects and irrealities that are still not actually there like this table might be. In a word, under Fink's rubric, it may be considered as a special, transcendent form of perception.

Writing in French, Sartre ([1940] 2004/2005) does not use the term *Bildbewusstsein*, but "the imaginary" (*l'imaginaire*), notably for *both* mental phantasies (e.g., a unicorn) as well as external, physical media. Indeed, for Sartre *all* images, whether external and physical, or supposedly merely internal and flighty, are of the same fundamental phenomenological structure: various psychophysical analogical materials are used to evoke transcendent objects that are either absent, irreal, or reality-neutral.

Let me take two examples, each one on the opposite end of the spectrum. The first is one person imagining a unicorn; the second is playing a digital game with friends.

Someone imagining a unicorn on their own would be a paradigmatic case of what is usually characterized as a mere "internal mental image." Under Sartre's conception however, there are always psychophysiological materials, as well as a transcendent factor, even to these types of images. The transcendent element is the imaged unicorn itself, the noematic imaginary object of the whole experience. It is indeed an object – not a "thing" (i.e., a perceptual object) of course – but an object in its own right with its own peculiar characteristics, not least its flighty, irreal, and mind-dependent nature. How do we attain to such an object? Through various materials that are present and immanent to us; in this case, our knowledge and memories of unicorns, horses, horns, particular myths, and so on, plus any emotions or feelings ("affectivity") and movements ("kinaesthetic

sensations"), we might have and apply thereto. Indeed, our lived bodies and their access to our memories allow us to evoke our own images of an irreal and perennially absent object that we nevertheless are able to experience precisely through our phantasy – *l'imaginaire*. Although each imaging act might be different for each person – some might have a more visual and colorful imaginations and others less – the transcendent character puts us all in the same structural boat when we aim our mind at such irreal beings.

In this manner, supposedly purely internal images have, to borrow a Husserlian term, hyletic impressional data in the form of brain matter and activity, as well as the concomitant conscious use and employment of knowledge and memories, affectivities, and movements, to explicitly evoke a transcendent irreal object. The only difference from external images for Sartre is that the analogical materials being used are within our own skins – and yet the crucial simultaneous insight is that we never just remain within our own skins; we always transcend toward an object, in this case an irreal and imaginary one.

Although external images have an external element that can determine what you can imagine quite rigidly, for Sartre, the formal structure remains the same: various psychophysical materials evoking objects that are not really or fully there. Actually, one could even argue that the image of a unicorn is quite determined too, for if one does not have the right knowledge of this object, and if one does not include various more or less essential characteristics in the construction of this particular image (horse and horn for example), then one can say one is not in fact imagining a unicorn at all.

To press this point home further, now let us look at the other end of the spectrum. Watching a film in the cinema. Of course, the details and textures here are very different. The screen, its sounds, and vivid physical colors are decidedly outside of me, and therefore wholly have the concreteness and physicality of perceptual objects. It is also a much more detailed and prolonged experience, with complex interactions, plots, and characters both real and digitalized, all of this moreover (re)presented in a social, communal environment. Nonetheless, the great differences in details and complexity all notwithstanding, it is still, for Sartre, ultimately another paradigmatic case of imaginary experience and structure precisely because all these external and physical phenomena are not perceived for themselves but once again allow one – along with one's knowledge, affectivity and kinesthetic sensations – to transcend into an imaginary realm and world of the filmmakers' creation. We almost never perceive mere pixels or colors on a screen; we automatically see an image of someone or something that was created in a place and time that we were not privy too. In this manner, this experience is also at the bottom of a Sartrean imaginary

structure where external and complex physical phenomena combine and are co-constituted with our own psychophysiological analogical materials, thereby allowing us to transcend way beyond our seats into, to incorporate a Finkian analogy, the film world.

In sum, no matter the particular phenomenologist or set of vocabulary, none of these experiences are simply "perceptually" present on the screen in a phenomenological sense; they are all ultimately images designed precisely with our inherent and wonderful capacity for image-consciousness (Husserl and Fink) or the imaginary (Sartre) in mind.

8.3 Horizonal and Digital Virtuality

Perhaps surprisingly, "virtuality" is an exceedingly difficult term to pin down with any precision in this context. Through my research (see O'Shiel 2019b, 2020, 2022a), I have landed on two main meanings. These are distinct in that they take place on different experiential planes, one on the level of reality and perception, the other as a particularly powerful subcategory of our imagination. They are still however both connected by virtuality's basic nature to be on the cusp on the real and present without actually being these, and therefore as a crucial experiential bridge between what is fully there and what is not but could quite easily be. This is the underlying nature of virtuality thus manifest in two different types: on the edges of our immanent and immediate perceptual fields in a way that makes these edges possible in the first place; or as manifest in the nature and dynamic of lively and physical digital images that can captivate us immensely and even come to supplant our perceptions both in the order of significance and value as well as the amount of time spent.

In a nutshell, I have become convinced that there are two rather unique types of virtuality that nevertheless are concrete manifestations of a more universal general dynamic between what is real and not, fully present and not. The first is on the level of perception and is basically Husserl's whole theory of perceptual consciousness as essentially "horizonal" consciousness – that is, in perception, it is impossible to have experiences that are not always already structured by the infinite potential for both inner (that is, looking, feeling, hearing closer) and outer (looking, feeling, hearing beyond) horizons. Although Husserl never uses the term "virtual" himself, it quite clearly fits perfectly for that which is not but could easily become present in my perceptual field. A very simple example is the implicit virtuality of the room next to me from where I currently am, and then the exterior of the house beyond that, then the road, the town, and so on – the world to the universe, that ultimate horizon. In this manner, only a very small portion of our experiential lifeworld is actually present to us at any given moment; all the rest is virtuality so, radiating out from great

and easy potentiality which is frequently and habitually actualized close to where we are, all the way to immensely distant climes, regions, worlds, and galaxies, which have an increasingly zero possibility of being actualized, and are even difficult or impossible to imagine due to our perceptual or even epistemic ignorance of them.

In short, in the perceptual realm, the virtual is an absolutely essential element, as the continuously almost or soon-to-be physically present, like the next room if I get up and actually go there. In this way, our perceptual lives are a constant interplay between the virtual becoming actual – that is, perceived – and the actual fading back into virtuality (again), often to be reignited when the right circumstances, motivations, and bodily movements arise.

On a different plane, we discover virtuality in the second main meaning. If the first meaning of virtual as potentially real or actual has a long and rather complex philosophical history going back to Medieval thought and concepts like Aristotle's "*dunamis*," this second variant is decidedly more recent and coincides with the rise of contemporary computerized and digital technologies and societies. Here, "virtual" simply means digital, computerized, and networked phenomena; all the "images" – that is, visual but also audial and otherwise (e.g., a vibration on a game controller) – that impinge upon our senses *through* our electronic and networked devices – yet more analogical materials in Sartrean language. In this manner, although one could vaguely talk about the virtuality of a unicorn in phantasy, or how Cézanne's *The Basket of Apples* (1895) can put you virtually in touch with a bunch of apples that are not actually there like a real one is, "virtual" and "virtuality" as we now ordinarily refer to these terms in this context are all those experiences facilitated by our networked digital technologies, and which thereby put us in direct and quite lifelike and realistic touch with information, people, and things that are still nevertheless not actually present like we or the screen or this apple is.

In this manner, the two meanings of virtual cover, on one hand, that which is almost or potentially present in the perceptual mode; and, on the other, that which is digitally so through our computers and devices. Moreover, this latter is, in structural phenomenological terms, a particularly strong and ever-developing category of image-consciousness (cf. also Liberati 2012).

I think, until now, that intuitively and experientially, most of us most of the time can still very much distinguish between what we perceive and what we imagine, namely between what is real and actually there, and what is a mere wisp of our phantasy. There are borderline cases like dreams, illusions, and the like (cf. Smith 2002), but these are usually corrected diachronically precisely because we have a more basic comprehension of reality and not. However, the crucial and interesting thing with

new "virtual technologies" – including technologies like social media and online gaming but also especially VR, AR, and MR technologies – is that they are all starting to increasingly blur the lines between what is simply real and actual and what is not. Digital screens are of course utterly central to all of this, and they have by no means finished developing and improving, and will undoubtedly continue to bind and merge with evermore advanced and immersive apps and accompanying appliances. In this manner, this so-called reality-irreality or even reality-virtuality distinction is already perhaps inverting in terms of importance and value for some, and it might even collapse in the future, not only theoretically but also experientially and evaluatively. Digital screens are, once again, already front and center at this ongoing revolution.

8.4 The Imaginary and Its Magic

Before I detail the phenomenological nature of screens more specifically, we need one more conceptual tool. This is Sartre's concept of magic. I have shown elsewhere (O'Shiel 2019a) that magic is a crucial and technical concept in Sartre's early (1936-1943) philosophy, and is used multifariously to describe personal reflections (ego), emotions, images, and values, all with an underlying structure. Here, I focus on the magic of the imaginary, and it seems to me that Sartre's imaging consciousness can be outlined here as magical in a fourfold manner. First of all rather trivially so, which is to say as a spontaneous creation of consciousness that does not need to obey the world's causal laws. Secondly, like other instantiations of Sartrean magic, there seems to be a strange mixture of activity (spontaneity, freedom, consciousness) and passivity (inertia, objectivity, things) in all imaging acts, where, thirdly, in certain cases (artistic creation, dreaming, and madness – but also digital screens) imaging consciousness can captivate itself so much that the real can be completely transfigured, suspended, or even corrupted. This third point culminates in the problem of dispossession, where the images can become so powerful that a kind of "reverse intentionality" could be said to be at work, where the object (image) seems to have you rather than you have it. Fourthly and finally, the fact that the imaginary can seem to infiltrate the world – especially through digital screens – at every turn corroborates Sartre's claim that "every concrete and real situation" (Sarte [1940] 2004/2005, 186/358) of consciousness in the world is already "pregnant with the imaginary" (ibid.), and this will be increasingly so in this new digital era, if the distinction between real and not will persist at all.

First, then, trivially so the imaginary is "magical" because it is a free creation of an individual consciousness that can flout laws of nature and be fantastical in the extreme. And even more mundanely, when I imagine

Pierre (Sartre's perennial example) on a bus after having looked for him in a café, one could of course say that such an act springs up because you are wondering where he is after not having found him. Here, the implicit experience of perceptual absence (not finding Pierre in the café) can lead to an explicit evocation of absence (Pierre on a bus). Nevertheless, although there is a link between these two events, there is not a strict causal one: it is just as possible to imagine Pierre on a bus without having just looked for him; it is also perfectly possible to imagine something completely else after having left the café; or to imagine nothing at all. In this way, although one can construct reflective superstructures explaining why one imagines such and such, these come after the more primary pre-reflective event. This latter, considered simply as it is, is a quite spontaneous creation that has no strict link to the present causal and motivational events, neither in content (e.g., a unicorn), nor in one's daily perceptual activities (e.g., looking for Pierre).

More technically – and this is the second issue – the imaginary is "magical" because of a strange mixture of activity and passivity. Imagining is "active" because of the creative and co-constituting act on the side of consciousness (noesis), while it is also "passive" because of the inert or determined quality of the image – including physical ones on screens – on the more objective side (noema). However, because of intentionality noesis and noema are always already intertwined, which in Sartre's magic means the objective, "passive" side to the image can be instilled with a pseudo power – i.e., a pseudo activity – of its own from the originally more active, noetic side. This is the imbued power of images, and it can take on varying degrees (the elasticity of such a principle). In general, it is quite clear for Sartre that the imaging intention has the power to animate its noematic correlate by breathing a kind of "feeble life" (ibid., 125/240) into it – as opposed to the strong, actual and physical life of perceptual objects. Then, with advanced digital screens, I would contend this "feeble life" has become very strong indeed, for some still not on the level of perception perhaps, but for others already there or beyond. It is still an imaginary structure, but of a new degree, it could almost be a new category, whereby images often seem to have us under their sway rather than the other way around – something I would characterize as "reverse intentionality" and a hallmark of Sartrean imaginary magic. A screen needs a watcher, an onlooker, a conscious subject, but once engaged its pull and pseudo- or even super-reality (see O'Shiel 2022c) definitely means the supposedly more passive, inert, and physical side (i.e., pixels changing color) come to inhere powers that seem quite autonomous unto themselves, not to mention that today's virtual online digital universe is populated by millions if not billions of human actors at any one moment. In this manner, this is a new level of interactive digital virtuality and why it can be captivating,

empowering but also dispossessing in the extreme. Indeed, this is a highly dynamic and interactive environment, but also one that is quite frenzied.

This then leads to the third aspect, in that magic for Sartre is always captivating in the strong sense, even to the point of dispossession (and addiction, as we will see). Regarding captivation, all images for Sartre have this simply by virtue of being created, witnessed, and engaged in and with. After this, it is once again a sliding scale, often with the more emotion and values put in or evoked, then the greater the pull and endurance of captivation. Think of a wonderful series you simply cannot get enough of; the captivation of the digital imaginary is so powerful not only because there is a whole world of media to suit all tastes – all of these media precisely tap into what we feel strongly about, value and want to know or learn, plus the fact that many algorithms and programs are designed to be as enticing and addictive (see Moore 2017) as possible in order to be successful with taps and likes and thereby boost their market share in a cut-throat capitalist environment. When this is all presented in a highly vivid and advanced technological manner that is moreover now ridiculously easy to access and carry around for many of us, no wonder then that the current screen of all screens – the smartphone – has revolutionized so much culture and behavior already, it being a form of magical Sartrean captivation *par excellence*.

This trend, fourthly and finally, looks set the increase. Ventures like Google Glass may have failed, and various AR, VR, and MR technologies may still remain rather niche markets for now. Nevertheless, almost unbeknownst to ourselves we are increasingly living in a blurred digital (ir)reality, with supposedly simple things like an Instragram or TikTok filter already meaning generations are growing up within augmented digital (ir)realities and (self-)images that have massive developmental, psychological and social implications (see O'Shiel 2020, 2022a). These phenomena and the devices that provide them look set to continue, propagate, and advance, with our physical environments harboring ever-more perceptual windows of digital transcendence, to update a Finkian thought. Moreover, these "windows" might not even feel like transcendent or imaginary for a good many anymore, especially younger generations who have grown up with them. In this manner, a "post-truth" age also has its seeds in a "post-real" age where the difference between real and factual and what one reads on one's screen is no longer very significant to a good many people. Easy and instant access to immense amounts of information and virtual phenomena of course has many benefits – but this is also a main weapon in the current hypercapitalist system that monetizes pretty much anything for the gain and profit of the very few over and against the actual sanity and prosperity of the very many (cf. Klein 2014; Reith 2019; Sassatelli 2007). There are good, strong, and more righteous movements if one looks. However, as we continue to be engrossed in image worlds through our devices and

environments in often highly magical manners where the distinctions between real, perceptual, factual, genuine, and true on one hand, and irreal, imaginary, fictional, fake, and false on the other, no longer hold much water for a good many – supposed social leaders like politicians (Trump, Johnson, etc.) very much included – one would do well to engage in the expansive and transcendent powers of immersive magic when beneficial and enjoyable, but also be wary of its overheated, overdemanding, and manipulative pitfalls, disturbances, and highly damaging lies as well.

Sartre's favorite definition of magic was the "mind crawling among things" (Sartre [1936] 2002/2009, 56/108). Our minds do often seem to be in and among things more than ever before, which is not even to mention the already-begun next wave of machine learning and AI technologies where not only distinctions between real and irreal and perceptual and imaginary are a stake, but also human and machine. The epicenter of all of this, for now, is still our digital screens.

8.5 The Phenomenological Nature of Digital Screens

To revert to a tone of philosophy or phenomenology 101, what *are* screens? In a nutshell, screens essentially block something while at the same time allowing something through. A fire screen protects from spitting embers but allows the heat into the room. A windscreen blocks wind, rain, bugs, and more but allows the driver to see. Even a smokescreen, both literal and metaphorical, blocks some kind of activity precisely so that activity can take place undetected or unwitnessed.

Digital screens fall into this essential nature as well. What they release or allow should be quite clear: a whole digital virtual interactive universe, still ultimately anchored in real people and machines. What they "block," beyond the trivial perceptual space directly behind, is a bit trickier to discern. Ultimately, it is the world of perception, not only seeing, hearing, and the like, but also activity therein. Basically, when one is "on" or "glued" to one's screen one is not engaging in the perceptual real world except beyond one's physical body being and subsisting there. One's mind, emotions, thoughts, and values are engaging in the materials through the screen, which is quintessentially a form of image-consciousness for Husserl and Fink, or one of the imaginary for Sartre. This world is virtual by definition as we have seen, although depending on the activity it can arc back to the real world. In fact, one can literally "influence" the actual world in much greater numbers and with much more force thanks to its virtual and transcendent power. What is more, there is a case to be made that screens and their content take the best of both the perceptual and imaginary worlds, in that they have all the fantastic and transcendent powers of the latter, but also still approximate or even surpass the real,

concrete, and vivid feeling of the former – elsewhere I call this "superreality" (O'Shiel 2022c).

Wiesing ([2005] 2010) makes a useful distinction in this context regarding four main categories of physical image: the fixed images of paints and photographs; the moving images of film; the manipulable images of animation (digital drawing, etc.); and interactive images. Most image scholars would emphasize this latter when it comes to digital images as they manifest especially on our digital touch screens. The genius of the smartphone is the interactivity not only with sight and sound but also essentially with touch, along with its easy portability where it is usually very close to our skin and thereby creates a feeling of closeness or even of intimacy that other larger and more immobile screens do not have. Add to this that we export much of our mental and personal investments to these devices too, from photo hoards (memory banks), to information stores and databases [downloads, documents and the whole searchable world wide web ("Google it"), entertainment (streaming), and pretty much all of our non-perceptual communication (email, messaging, social media profiles). This means that for many of us, it is almost phenomenologically fair to say that we as people and psychological entities already exist more on and through the screen than we do in our own heads and immediate environments. The extended, immersed mind has already long arrived and has exploded with these devices, and it is digital screens that have provided this essential gateway. Hereby, the digital screen is a physical device that nonetheless allows us to be, be active and persist virtually quite universally and indefinitely, and moreover to a potential audience of people and other digital phenomena that would not even be possible in the actual, immediate perceptual world (even our largest stadiums come nowhere near close enough to the amount of "followers" some people have.).

This actual world does remain, but it is often forgotten and neglected. Indeed, in rearing children, there are some strictures around "screen time" as it is recognized too much can be a bad thing and other real-world activities (e.g., exercise), interactions (e.g., a face-to-face conversations), and pastimes (e.g., a creative material hobby) can all be left by the wayside if the digital imaginary is indulged in too much. Add to this the fact there is quite mounting evidence that some of the main online powers intentionally create digital and virtual mechanisms that are literally hard to put down for any significant length of time, plus the fact that the online world is full of abuse, viciousness, and monetization powers that pretend but very rarely have the average individual's best interests at heart, then it is also a world that can be as dangerous and damaging as it is enticing and exhilarating. The digital screen, once again, is the current gateway to all of this thanks to its physical anchor that allows easy and relatively free

access to an infinite, dynamic, ever-changing metaverse of information, sound, and color.

What is "blocked" then through digital screens is one's more immediate perceptual surroundings and engagement. Although it might not feel like it to some, phenomenologically and actually looking at and into a digital screen is an act of transcendence whereby one by definition "zones out" of the immediate environment to any real effect, because one is focusing one's primary attention on the digital phenomena presented in that magical black-rimmed image box. If this is done with balance and conscientiousness, it can be extremely enjoyable, powerful, and productive, but there is also a great tendency to stunting, inanity, and addiction. I focus on the last now as my final step.

8.6 Addiction as Hypervirtuality

For me, phenomenologically, addiction is experienced as what I would term "hypervirtuality," or what Maté (2009, 99) refers to as "an irresistible gravitational field." This is to say, whatever the object one is hooked on, and no matter whether the craving is more physiological or psychological, at certain moments and in certain conditions – or in chronic cases pretty much all the time – the desired object is so dominant that it is pseudo present to one's mind (and body) to the extent one cannot be rid of it until it is attained, or the craving eventually subsides. As an example of digital, screen addiction, think of a gamer at school who cannot think of anything except running home and going back online. Underlying all of this, I contend, is a phenomenological notion of "hypervirtuality" that is basically an extreme end-point of the ideas we have been progressing through, and it also ties in with – or is even the content of – addiction as a form of intense lack, transcendence, and escape (see O'Shiel 2022b).

This idea is not explicitly present, but is nonetheless inspired by Husserl. As we have already seen, for him conscious experience – and especially perceptual experience – is essentially horizonal. Indeed, for Husserl, there are always inner (looking closer) and outer (looking beyond) spatial horizons to absolutely everything we perceive (see Husserl [1918–1926] 2001, 43), as well as temporal horizons of past, present, and future that all interlock in an incredibly complex dynamic under the rubric of time-consciousness (see, for instance, Husserl [1893–1917] 1991, § 14; also de Warren 2009). For me, this inherent horizonality can be applied to our own minds and desires as well, whereby something is horizonal – or "virtual" – when it is present without actually being fully so. Thus, when I want a glass of water, the desired object virtually hovers before my mind, stimulating me to go and get it.

Addictions would then be when such virtualities become "hyper-," which is to say more intense and beyond what is strictly necessary or even healthy. Indeed, someone who is addicted to nicotine for instance, and is out of smokes in the middle of the night, is going to find it very difficult to sleep or do pretty much anything until the "fix" is obtained. It is well known to what lengths people under the influence of such hypervirtuality will go. In this manner, whatever the trigger, whether it be the body's craving in relation to the time of day or situation (e.g., after eating), or some kind of emotional upheaval or event, or a more constant behavior or way of existing, in all cases the object of one's desire hovers about to such an extent and intensity that it can block out almost everything else, or even arrange all else under subservience to this hyper, privileged desire ("I used to like to sleep, but I found cocaine was better. I used to like to eat, but I found cocaine was better" – in Hirschman 1992, 176). Here, depending on the nature of the object and the strength of the addiction, the feeling will subside while actually present with the object – a transition from hypervirtual to actual – only for the movement to reverse again after the effects have worn off, or once one has been separated from the object for an intolerable amount of time, which of course is variable.

In this manner, addictive objects and activities, in a plethora of different manifestations from complete chemical euphoria to utter electronic escape, leave such imprints on our desire that when we are apart, we simply want more and can think of little else. Then, when chronic and absolute, we no longer (want to) leave the object at all – like in the case of the gamer Billy Brown "who spent seven years in his dressing gown" (Carr 2018). Indeed, in its extreme addiction as hypervirtuality can structure absolutely everything. Loewenstein ([1999] 2007, 256) in fact notes that "severe addicts tend to classify people into two categories: Those who threaten to impede access to the drug and those who can serve as tools for obtaining it." In this manner, when an addiction is particularly severe and persistent, it is a constant and relentless form of hypervirtuality; there is only one thing and it is nonetheless never enough. It is one specific, hyperprivileged object (or set of objects) that is not only constantly on one's mind; the desire for it also governs all else one thinks and does.

This idea has already been pointed to in a good deal of the literature, just framed a bit less phenomenologically. Elster (1999, 69) speaks how "[m]any drugs have a tendency to crowd out all other activities." Waal ([1999] 2007: 143) speaks of the "Siren-like appeal of […] drugs" where one's rational-choice system becomes "truncated and distorted" to the privilege of certain special, desirously hypercharged objects. Loewenstein ([1999] 2007: 245–246) talks of the increased value addictive objects accrue, as well as how they change one's time horizon toward the present, thereby discounting any serious future considerations. West and Brown

([2009] 2013, 7) speak of "an abnormally and damagingly high priority to a particular activity" that is moreover felt compulsively (ibid., 95). Heyman ([2009] 2010, 28) talks of a "shift in priorities," most notably to local and immediate gratifications from more measured, global, and long-term ones (cf. ibid., 140). Finally, Moore (2017, 68) speaks of how "circuits not liked to the dopamine craving fall into disuse and are 'pruned' away, narrowing attention more and more tightly around repetition-compulsion." For me, all of these references speak to the hypervirtual nature of addiction; in the moment of desire and craving, it is all one can think of, with the duration depending on the nature, circumstances, and severity. Screen addiction is one such instance that is perhaps more prevalent than a good few of us would like to admit.

Note

1 All translations for this text are mine.

References

Carr, Johanna. 2018. "The gamer who spent seven years in his dressing gown." Last modified November 9, 2018. https://www.bbc.co.uk/news/uk-england-cornwall-46029562.

de Warren, N. 2009. *Husserl and the Promise of Time: Subjectivity in Transcendental Phenomenology*. Cambridge, MA: Cambridge University Press.

Elster, J. 1999. *Strong Feelings. Emotion, Addiction, and Human Behavior*. Cambridge, MA & London, England: The MIT Press.

Hirschman, E. C. 1992. "The Consciousness of Addiction: Toward a General Theory of Compulsive Consumption." *Journal of Consumer Research* 19: 155–79.

Husserl, E. 2001/1966. *Analyses concerning Passive and Active Synthesis – Lectures on Transcendental Logic*. Translated by A. J. Steinbock. Dordrecht, Boston, London: Kluwer/*Analysen zur Passiven Synthesis. Aus Vorlesungs- Forschungsmanuskripten 1918–1926*. Husserliana IX, edited by M. Fleischer. Den Haag: Martinus Nijhoff.

Husserl, E. 2005/1980. *Phantasy, Image Consciousness, and Memory*. Translated by J. B. Brough. Dordrecht: Springer/*Phantasie, Bildbewusstsein, Erinnerung. Zur Phänomenologie der Anschaulichen Vergegenwärtigungen*. Husserliana XXIII, edited by E. Marbach. The Hague, Boston, London: Martinus Nijhoff.

Husserl, E. 1991. *On the Phenomenology of the Consciousness of Internal Time*. Translated by J. B. Brough. Dordrecht, Boston, London: Kluwer.

Jansen, J. 2013. "Imagination, Embodiment and Situatedness: Using Husserl to Dispel (Some) Notions of 'Off-Line Thinking." In *The Phenomenology of Embodied Subjectivity*, edited by R. T. Jensen and D. Moran, 63–79. Basel: Springer.

Klein, N. 2014. *This Changes Everything. Capitalism vs. The Climate*. London: Penguin.

Liberati, N. 2012. "Improving the Embodiment Relations by Means of Phenomenological Analysis on the 'Reality' of Ars." In *IEEE International Symposium on Mixed and Augmented Reality 2012. Arts, Media, and Humanities Proceedings*: 13–17 (DOI: 10.1109/ISMAR-AMH.2012.6483983).

Loewenstein, G. 2007. "A Visceral Account of Addiction." In *Getting Hooked. Rationality and Addiction*, edited by J. Elster and O.-J. Skog, 235–264. Cambridge, MA: Cambridge University Press.

Maté, G. 2009. *In the Realm of Hungry Ghosts: Close Encounters With Addiction*. Matheson: Knopf Canada.

Moore, G. 2017. "Dopamining and Disadjustment: Addiction and Digital Capitalism." In *Are We All Addicts Now? Digital Dependence*, edited by V. Bartlett and H. Bowden-Jones, 65–83. Liverpool: Liverpool University Press.

O'Shiel, D. 2019a. *Sartre and Magic. Being, Emotion and Philosophy*. London, New York, Oxford, New Delhi, Sydney: Bloomsbury Academic.

O'Shiel, D. 2019b. "Phenomenology and the Challenge of Virtuality." In *Conceiving Virtuality: From Art to Technology*, edited by J. Braga, 21–43. Dordrecht: Springer.

O'Shiel, D. 2020. "Disappearing Boundaries? Reality, Virtuality and the Possibility of 'Pure' Mixed Reality (MR)." *Indo-Pacific Journal of Phenomenology* 20 (1): 1–8.

O'Shiel, D. 2022a. *The Phenomenology of Virtual Technology. Perception and Imagination in a Digital Age*. London: Bloomsbury Academic.

O'Shiel, D. 2022b. "Lack, Escape, and Hypervirtuality: On the Existential and Phenomenological Conditions for Addiction." *Philosophies* 7 (112): 1–12.

O'Shiel, D. 2022c. "Digital Games, Image-Consciousness and Superreality." *Journal of the Philosophy of Games* 4 (1): 1–19.

Reith, G. 2019. *Addictive Consumption. Capitalism, Modernity and Excess*. London, New York: Routledge.

Sartre, J.-P. 2002/2009. *Sketch for a Theory of the Emotions*. Translated by P. Mairet. London, New York: Routledge/*Esquisse d'une théorie des émotions*. Paris: Hermann.

Sartre, J.-P. 2004/2005. *The Imaginary: A Phenomenological Psychology of the Imagination*. Translated by J. Webber, London, New York: Routledge/*L'imaginaire. Psychologie phénoménologique de l'imagination*. Paris: Éditions Gallimard.

Sassatelli, R. 2007. *Consumer Culture. History, Theory and Politics*. Thousand Oaks: SAGE Publications.

Smith, A. D. 2002. *The Problem of Perception*. Harvard University Press.

Waal, H. 2007. "To Legalize or Not to Legalize: Is That the Question?" In *Getting Hooked. Rationality and Addiction*, edited by J. Elster and O.-J. Skog, 137–72. Cambridge: Cambridge University Press.

West, R., and J Brown. 2013. *Theory of Addiction*. Oxford, Chichester, & Hoboken, New Jersey: Wiley Blackwell.

9 Perceiving the Virtual

Rethinking Blaustein within the Phenomenology of Virtual Reality

Witold Płotka

9.1 Introduction

It is as true as it is a truism to say that virtuality has become a constant element of our everyday lives. After all, one operates with the virtual while using a smartphone, social media, or playing games. However, this all-pervasive coping with the virtual has significant consequences for users (e.g., Flavián, Ibáñez-Sánchez, and Orús 2019; Rauschnabel 2021). As a result, the virtual shapes, to some extent, lived experiences, and thus lived-worlds. Here crucial questions arise: What does one experience if the virtual is given? How shall one describe structures of this type of experience? What is it like to perceive the virtual? In this study, I address these questions, and by doing so, I offer to examine the phenomenon of experiencing the virtual by referring to Leopold Blaustein's ideas formulated in the field of aesthetics of media experiences. I attempt to show that even though Blaustein formulated the basics of his conceptions almost a century ago, one can use at least some of them in today's studies on the virtual. By doing so, I do not aim at formulating a full-scale theory of the virtual. Instead, I provide some initial considerations on selected topics discussed in scholarly literature, e.g., the structure of experiences of the virtual or the status of virtual objects.

Before I explain why a reference to Blaustein in this context may be useful, and I spell out the detailed tasks of the present study, I note that virtuality is discussed here as connected with experiencing virtual reality (VR). Generally, VR is here to be understood as technologically (digitally) determined experiential content, which creates someone's sense of immersive virtual environments. Of course, VR thus defined, following Metzinger (2018), can be examined by philosophers from different angles, leading one to various, partly connected issues. By raising the question of experiencing the virtual, one admits that VR creates specific phenomenal qualities of experiences of what is given as virtual. In a word, the user lives through a specific "presence" of the virtual. Phenomenology serves

DOI: 10.4324/9781003312284-10

one to explore this very aspect of "presence" in experiencing the virtual (e.g., Breuer 2020; de Warren 2014; O'Shiel 2019). However, this type of experience is indeed complex as it corresponds to different degrees of the virtual and the real or, to use Milgram and Kishino's (1994) phrase, the "Reality-Virtuality Continuum." As a result, phenomenology addresses the questions of experiencing real environments (RE), virtual environments (VE), augmented realities (AR), or augmented virtuality (AV), not to mention mixed reality (MR), which is a mixture of real and virtual elements (Flavián et al. 2019; O'Shiel 2020; 2022). In this regard, phenomenology provides one tools for describing various aspects of experiencing the virtual on different levels. But, one may ask, why are Blaustein's ideas relevant here to begin with?

Blaustein was born in 1905 into a Polish-Jewish family, and he studied philosophy at Jan Kazimierz University in Lvov (Lwów, Lviv) in roughly 1923–1927.[1] In that time, he spent a few weeks in Freiburg im Breisgau where he participated in the lectures and seminars of Edmund Husserl. Back in Lvov, under Kazimierz Twardowski's supervision, he wrote a doctoral dissertation on Husserl's theory of intentionality, and he took part in Roman Ingarden's courses on the literary work of art.[2] At the end of the 1920s, Blaustein had undertaken research on the theory of presentations and on the problem of mediated experiences. The roots of Blaustein's interests in mediated experiences can be traced back to 1927–1928, when he spent a few months in Berlin, where he attended the lectures of, among others, Carl Stumpf, Wolfgang Köhler, and Kurt Lewin, and he participated in seminars held at the Berlin Psychological Institute. While there, Blaustein observed that scholars at the institute used a camera to design and plan experiments; in this regard, a more profound question arose for him: How lived experiences of watching a movie can be described at all? In 1930, Blaustein published a book on his theory of presentations which explores, among other things, theater-goer's experiences, paintings, and everyday mediated experiences. Later, in the 1930s, Blaustein extended the scope of his studies by exploring different psychic phenomena, chiefly from the field of aesthetic experiences, e.g., cinema-goers' experiences or the experience of listening to the radio. Blaustein did not complete his far-reaching studies on the phenomenology of mediated experiences, as he died in probably 1942 or in 1944 in the Lvov Jewish ghetto.

By reinterpreting Blaustein's philosophy in the context of VR, the present study contributes to growing and ongoing studies on the heritage of early phenomenology (e.g., Miron 2021; Moran and Parker 2015; Parker 2021; Płotka and Eldridge 2020; Vendrell Ferran 2008). In this regard, Blaustein is an interesting scholar who attempted to bridge the gap between different traditions, including the analytic heritage of the Lvov-Warsaw School, the phenomenology of Husserl and Ingarden, and the Gestaltists' approach.

Although Blaustein's thought has gone neglected, it still shows an essential contribution to the history of philosophy. Needless to say, Blaustein developed the very first phenomenological (in a broad sense, i.e., rooted in Brentanian heritage) account of cinema-goers' and radio-listeners' experiences. This last point is crucial in the context of the present study. After all, the virtual is experienced by users as a technologically (digitally) determined medium which shapes our sense of presence. It is assumed here that due to his noticing this mediatory aspect of experiences, Blaustein can also provide some conceptual framework for understanding VR-experiences. Moreover, he widely discussed the phenomenon of imagination while analyzing mediated experiences. Given that imagination is essentially involved in experiencing the virtual,[3] it would be instructive to examine this aspect in Blaustein as well. Finally, he explored the phenomenon of perception as the key to understanding aesthetic experiences.[4] One can adopt and re-interpret this idea in the field of VR studies as well in order to show what it is it like to perceive the virtual. With these ideas in mind, it is no exaggeration to expect interesting results while reading Blaustein from the perspective of today's studies on VR.

In the following, first, I discuss the central points of Blaustein's aesthetics. In particular, I am interested in such phenomena – analyzed by Blaustein – as watching a theater play, watching a movie in a cinema, and listening to a radio drama. In this regard, I outline his account of the object of such mediated experiences, and I examine the perceptual character of such phenomena. Second, I summarize ongoing debates among today's scholars on the virtual. I will look closer at Chalmers's recent account of virtual realism and virtual irrealism. By exploring the debate on Chalmers's reading, I aim at highlighting some problems in understanding the status of virtual objects. Moreover, I will refer to recent phenomenological studies on the virtual. By doing so, my ambition is a preliminary description of the phenomenal character of experiences of the virtual. Finally, I will confront Blaustein's account with today's debates. Against this background, I sketch some conceptual improvements which arise due to the reference to Blaustein. I will refer to ontological as well as phenomenological considerations. At the end, I attempt to situate Blaustein in the conceptual landscape of today's debates.

9.2 Blaustein on Mediated Aesthetic Experiences

Blaustein (2011a, 210) explicitly classifies his approach as *descriptive psychology*, wherein this category covers for him different methodological tools rooted in ideas of such thinkers as Brentano, Twardowski, Husserl, Stumpf, Ingarden, and Dilthey. Blaustein's method is descriptive, analytical, experimental, and thus (partly) naturalistic, though phenomenologically

oriented and humanistic. Elsewhere, I have argued that in spite of evident breaks with Husserl or Ingarden, Blaustein's philosophy has a phenomenological character.[5] Generally, Blaustein's phenomenology consists in carefully describing experiences, focusing on the ways of manifestations (*Gegebenheitsweisen*) of certain objects (e.g., Blaustein 1937), and in gradual generalizations of descriptions in order to explicate *types* of relevant experiences.[6] In addition, following Brentano and Twardowski, Blaustein places emphasis on the object comprehended as distinct from content, and the act. With these ideas in mind, Blaustein's analysis of mediated aesthetic experiences lies in an attempt at identifying the object(s) of these experiences and, next, in describing first-person ways of experiencing those very objects. A good example of this procedure can be found in his book on *Imaginative Presentations*; he writes:

> Looking in the mirror I can adopt different attitudes. I can intend either objects that are outside me or my body or objects that are as if inside the mirror, in some peculiar world manifesting itself in the mirror in front of me. The first attitude is present when, for example, looking in the mirror I realize that my eyes are red, the other, when I jokingly wag my finger at my lookalike, who reciprocates the gesture. In the latter case I do not see myself, but some other man, very similar to me, but not identical.
>
> (Blaustein 2011a, 216)

In this fragment, Blaustein analyzes an everyday phenomenon of looking in the mirror. In the background, he accepts the idea that this act is intentional which means that the phenomenon consists in intending its object – the body or the object reflected in the mirror. This perceptual experience is founded on given images which refer to their objects adequately. The mirror image is here an appearance of the viewer who is standing in front of the mirror. To use more technical language, for Blaustein, all elements of the presenting content refer to elements of the presented object; for instance, a blond color manifested on the mirror's surface relates to the viewer's blond hair, whereas green surfaces relate to the viewer's eyes, etc. However, this description does not exhaust the analyzed phenomenon as while looking in the mirror, one apprehends another object, yet inadequately, i.e., one's lookalike. The mirror image is "distant" from the viewer. The first case described above shows that the mirror image *adequately* presents its object, i.e., the body of the viewer. The second case, in turn, shows that there are elements in the mirror image (e.g., playing a role in front of the mirror), which do not correspond with the object (person) presented in the mirror. Accordingly, the mirror image represents *not* the viewer but a character performed by the viewer. According to Blaustein's

theory of presentations, this kind of presentation is called *imaginative*. He determines this type of presentation as constituted on the basis of what is intuitively given (the mirror image which is actually perceived), but its object is *not* actually given – it is "distant" or "improper" (namely, no one is wagging their finger at me). It is presented, as Blaustein puts it, *quasi*-adequately.

From Blaustein's point of view, the object(s) of mediated experiences is complex, i.e., different apprehension of acts' contents are correlated with different objects, wherein objects are constituted in each other. The object which functions as *constituting* is, to use Blaustein's (1931, 26–28) vocabulary, the *closer* or *proper* object, whereas the object which functions as *constituted* is the *distant* or *improper* object of intentional relations. In the case discussed above, the colorful surfaces experienced by the viewer while looking in the mirror are closer objects which provide one with the constitution of a distant object, i.e., the mirror image; next, the latter may serve as the basis for another distant object, i.e., a fictional character who is wagging their finger at the viewer. As a result, Blaustein offers to differentiate the mirror image as the *imaginative* object, whereas the character constituted on this basis as the *imagined* object. Both objects thus described are generated by relevant acts, say, looking in the mirror or perceiving the fictional character. Blaustein uses these ideas in his analysis of different types of mediated aesthetic experiences (e.g., Płotka 2020b; Rosińska 2011; 2013). In the following, I will examine three phenomena: (1) watching a theater play, (2) cinema-goers' experiences, and (3) listening to a radio drama.

To begin with, (1) in his *Imaginative Presentations*, Blaustein (2011a, 216) describes an experience of watching Shaw's *Caesar and Cleopatra*. During the play, the goer perceives events happening on stage and objects present there. All these events and objects are adequately presented in perceptual experiences. However, all this is not meaningless; instead, one perceives persons on stage *as* performing certain roles. For instance, when Caesar is talking with Cleopatra, one perceives different characters and roles which are constituted not as common objects of goers' everyday lives. Fictional characters performed by actors are not real. One can touch an actor, but one cannot touch Caesar or Cleopatra. Blaustein classifies such objects as imaginative and explains that they are *quasi*-real.[7] In this example, the imaginative object (e.g., Caesar) is constituted in the imaginative presentation and is given at once as intuitive (the real movements and words happening on stage) and non-intuitive (Caesar meeting Cleopatra). Whereas the former is the closer or proper object of the perceptual intention, the latter is the distant or improper object of the imaginative intention. The difference arises at the level of experience: the intuitive object has properties that are truly ascribed to it by the act (e.g.,

being a man or woman, having blond or dark hair), and the non-intuitive object has properties ascribed in the *modus* "*quasi*." From the first-person perspective, Blaustein compares this experience with illusion; however, what differentiates an imaginative experience from an illusory one is the lack of a belief that the object exists at all (in the case of an illusion, according to Blaustein, one has to believe that the illusionary object is true). In any case, actors perceived on stage are experienced in a unique way of a mixture of intuitive and non-intuitive components.

Next, (2) in the *Contributions to the Psychology of the Cinema-goer*, Blaustein (2005, 92–127) assumes that a cinema-goer has experiences that are *sui generis*, which means that they cannot be considered to be like other everyday experiences. They are unique and irreducible to other types of experiences. The basic level of these experiences are phantoms which one experiences while perceiving a cinema screen (Blaustein 2005, 94). Phantoms are given as ever-changing groups of colors, lights, and marks. However, this play of lights and colors is not meaningless. Goers apprehend them as objects "on" the screen. In addition to perceptual sensations, goers perceive voices which mixed with phantoms enables one to indicate non-real objects, more precisely, objects which are present in the world "in" the movie. The objects "in" the movie are classified by Blaustein as *imaginative* objects since they are given as intuitive and non-intuitive at the same time. After all, objects "in" the movie are distinct from the perceived colors, lights, and marks. For Blaustein (2005, 100), objects "in" the movie are grouped in organized sets, and as a result, they form the *imaginative* world that is represented "in" the movie.

Finally, (3) Blaustein also analyzes the phenomenon of listening to a radio drama. The term "radio drama" refers to a dramatized and purely acoustic performance produced for and emitted by radio. Blaustein (1939; 2005, 145–196) analyzes this phenomenon in *On the Perception of Radio Drama* (published in Polish and in French translation).[8] Blaustein's basic insight is that the radio experience has a unique form of perception which consists in a direct presenting of its objects. Just as presented above in regard to the cinema experience, Blaustein holds that the radio drama experience is *sui generis*, i.e., irreducible to other forms of everyday experiences. To show this, he uses the term "acousion" (a combination of two terms: "acoustic" and "vision"), which refers to a specific perceptual experience which relates the listener to the world represented in a radio drama. What does the term mean? Acousion, of course, does not mean that that the listener can "perceive" the objects in a radio drama *just like* one can "perceive" trees while taking a walk in the park. Rather Blaustein attempts to show that listening to a radio drama is a fusion of intuitive and non-intuitive components. A radio-listener experiences the presenting

content of her lived experience, i.e., sensations, which, in turn, are understood as purely auditory experiences which include different *sounds*. These sounds, however, are apprehended as sounds *of* objects or characters inherent to a certain drama. Against this background, Blaustein holds that one perceives objects in a radio drama directly or "intuitively" *in propria persona*. And thus, for instance, one perceives the sound *of* a river, the tumult *of* a fight, the clatter *of* a steam machine, etc. All in all, the phenomenon of acousion is a specifically auditory way of presenting objects which are heard by the listener. In his study, Blaustein (1939, 127; 2005, 160) writes: "to perceive a radio drama, one has to have auditory experiences of ambient sounds, such as a ring or the wind," yet there are also further components, as one hast to understand "the words and sentences uttered by the characters of the radio drama." Acousion, as described by Blaustein, plays the central role in generating the relevant world, i.e., the fictional world constituted in acts of hearing and perceiving sounds.

All three examples of mediated aesthetic experiences described above show, first and foremost, that Blaustein assumes that these experiences are unique and irreducible to any general structure. Of course, all lived experiences are intentional for him, but this claim does not give any impression of, for instance, what it is like to watch a theater play or a movie in a cinema. Blaustein sees that every type of mediated aesthetic experience is *sui generis*. Moreover, he holds that in the end, mediated experiences are experienced as *direct*. How shall one understand this apparent contradiction? For Blaustein, perceived objects are given *directly*, even though they are constituted as *mediatory*; nevertheless, one does not perceive these objects *as* mediatory. Instead, one apprehends mediatory objects and, due to this apprehension, an act intentionally intends the imaginative object which becomes the very object of intentional relation on its own. In this sense, objects of mediatory aesthetic experiences are generated within these experiences; thus, such experiences are active. With these ideas in mind, it should come as no surprise that, for Blaustein (2005, 5), the subject who lives in these experiences is strictly active. In the essay "On Apprehending Aesthetic Objects," one reads:

> Admittedly, the aesthetic experience is first and foremost a passive experience, an apprehension and perception of aesthetic objects. In addition to the perception of an object, we can also find in it a rich source of experience in which we react to what is given to us in perception. We experience feelings in aesthetic experiences; judgments occur rarely, e.g., in the form of aesthetic assessments; acts of will appear very rarely. But the activity of the aesthetically experiencing human being is manifested not only in these reactive components of

> the aesthetic experience but also in perceptive components—in those in which a *seemingly only passive reception* to the aesthetic object is present.
>
> (Blaustein 2005, 4; my italics)

In the light of this passage, one can see that perceiving for Blaustein is a complex act which begins as a passive reception of (closer or proper) objects. So, to begin with, perception is intuitive, but it is also associated with non-intuitive components. These components, in Blaustein's (1937, 245–249) words, give the object in different "ways of givenness" or "manifestation" (*Gegebenheitsweisen*). While having a certain mediated experience, one shifts her focus in order to constitute different objects. In this way, objects which are not initially given directly, as they are distant or improper objects, can still be given in someone's experience. To do this, however, one should adopt a certain attitude; to phrase it differently, one should be immersed by these experiences by perceiving the fictional objects as *quasi*-real. After all, a theater- or cinema-goer enjoys a certain theater play or a movie if she is engaged with the fictional world presented on stage or on screen. As I will attempt to argue in the next sections, these ideas of Blaustein's are of primary importance in VR-studies as well.

9.3 Recent Debates on Virtual Objects and the Phenomenology of Virtuality

This section explores some central topics discussed in recent debates on the virtual. Firstly, I focus on the problem of virtual objects and next on the issue of the phenomenology of virtuality. The former problem is connected with the aforementioned idea of the "Reality-Virtuality Continuum" (Milgram and Kishino 1994). After all, it is true that VR is constituted on the borderline of the real and the non-real or virtual (Heim 1993, 128).[9] The common understanding of the virtual as non-real leads one to the center of today's debates: Are virtual objects real or fictional?

This question was recently addressed by Chalmers (2017; 2022, 105–123) who spelled out two main theoretical positions in this regard as *virtual realism* and *virtual irrealism*. The former generally holds that (1) virtual objects really exist, (2) events in VR are equally real, whereas (3) experiences in VR are non-illusory, and (4) as valuable as non-virtual experiences. The latter holds opposite theses (Chalmers 2017, 310). Against this background, Chalmers (2017, 311) identifies his position as a sort of *virtual realism* which holds that:

1 Virtual objects really exist and are digital objects.
2 Events in virtual worlds are largely digital events that really take place.

3 Experiences in virtual reality involve non-illusory perception of a digital world.
4 Virtual experiences of a digital world can be about as valuable as non-virtual experiences of a non-digital world.

As a result, Chalmers (2017, 312) holds that VR digitally generates a VR environment, which is *immersive*, *interactive*, and *computer-generated*. Chalmers's point is that even if virtual objects are objects of a VR environment, one *really* encounters these objects and *really* copes with them. For him, it is possible to comprehend virtual objects as real, as they are digital objects which means that they are constituted by computational processes on a (real) computer. But if digital processes on a computer are real, and VR is founded on these processes, VR is also real.[10] In this context, Chalmers holds that properties of virtual objects which are perceived by the user reflect the properties of data structures which are causally related to real computers in the real world. In a word, there is a strict correlation between properties of relevant objects. Virtual properties correspond to non-virtual properties. While virtual properties are determined to play certain functional roles, non-virtual properties are functionally equivalent. Consequently, if there is functional equivalence, both properties are real.

Chalmers's ideas were critically discussed, among others, by Beisbart (2019) and McDonnell and Wildman (2019). Beisbart (2019, 300) classifies Chalmers's theory as *virtual digitalism*, which comprehends virtual objects as *digital objects*. Following Beisbart, however, the consequences of Chalmers's ideas are puzzling and unclear in regard to the category of virtual objects. Beisbart rightly shows that there are different digital objects, e.g., digital simulations, which can hardly be classified as real. The problem arises since Chalmers does not specify the category of virtual objects (Beisbart 2019, 320). As shown above, for Chalmers, virtual objects are generated by digital processes on a computer. These very computer processes are performed by different data structures depending on the program generating the data. However, asks Beisbart, what are digital objects at the level of hardware? Moreover, what does it mean that properties of computer properties are correlated with properties of virtual objects? In his reply, Chalmers (2019, 455, 460) indicates that, first, virtual objects have double foundation in data structures *and* mental objects, and, second, he specifies the category of data structures as "a *physical data structure*, which is a *realization* of the abstract data structures specified in computer programs and the like." In addition, Chalmers (2019, 467) is skeptical about Beisbart's argument involving virtual simulation, as a simulation generates a huge number of digital objects; this consequence is consistent with Chalmers's point.

McDonnell and Wildman's (2019, 379) train of thought also starts with an idea similar to Beisbart's – that differences in hardware make the issue of the identity of digital objects controversial. Unlike Beisbart, McDonnell and Wildman (2019, 391) offer to use Walton's (1990; 2013) idea of make-believe in order to describe how one is engaged with virtual objects. According to McDonnell and Wildman, one is engaged with VR as VR generates specific props which are used in a game of make-believe. Consequently, the user comprehends virtual objects as they virtually are. In the end, McDonnell and Wildman (2019, 391) label their proposal "virtual walt-fictionalism" and hold that its central claim "is that virtual reality is a kind of walt-fiction, and our engagement with VR is not different in kind from our engagement with other forms of walt-fiction." In response, Chalmers (2019, 472–473) partly accepts McDonnell and Wildman's proposal, yet as a limited thesis. "Virtual walt-fictionalism" applies namely only in regard to VR and roleplaying, and it is not true of virtual worlds in general. Rather, as Chalmers concludes, virtual digitalism is an adequate position. All things considered, the main disagreement here lies in the question of the reality or non-reality of virtual objects. Some realists *à la* Chalmers may argue that these objects are real, as they are constituted in digital objects generated by process on a computer which exists in the real world. By contrast, irrealists – *pace* Chalmers – may refer to Walton's idea of make-believe to defend the position that virtual objects are fictions based in imaginative though structured experiences.

The second topic I want to highlight now concerns the phenomenology of virtuality. In his texts, Chalmers (2017, 2022) discusses the non-illusionary character of the perception of virtual objects, and he draws a distinction between *naïve* and *sophisticated* users of VR. Whereas the former may falsely believe that what one perceives is real, the latter knows that one in fact perceives virtual objects, and thus one perceives them *as* virtual. This division leads to the issue of the phenomenology of virtuality: What is it like to perceive virtual objects? At bottom, this question concerns experiences considered from the users' perceptual point of view. In this regard, McDonnell and Wildman (2019, 394–395) offer to describe this type of experience as an imaginative activity which consists in playing a game of make-believe. In today's literature, scholars in phenomenology generally agree that experiencing the virtual is founded on imagination. To begin with, de Warren (2014) explicitly binds the lived experience of virtuality with imagination, wherein the category of virtuality is broadly understood, e.g., computer games and fictional environments. It seems that de Warren's (2014, 103) key insight lies in the thesis that virtuality is mediated with images which, in turn, possess "a virtual manner of givenness." To describe what it is

like to experience the virtual, de Warren refers to Husserl's analysis of image-consciousness,[11] and he argues that the distinction of the real and the virtual (non-real) reflects the way objects are given in perception and imagination respectively. De Warren's idea to use Husserl's conceptual framework in describing VR-experiences is developed, for instance, by Liberati (2018, 217) and O'Shiel (2019; 2020; 2022). Liberati and O'Shiel seem to agree that it is important to take into account Husserl's differentiations of perception and imagination in order to understand how different virtual technologies shape our lived experiences by generating various ways of presence. At the same time, Liberati and O'Shiel use other phenomenological theories in this regard. Whereas Liberati (2018, 221–225, esp. 223) adopts Schütz's idea of paramount reality to show that AR produces the world as virtual with time and space on its own, O'Shiel (2019, 29–31, 37) confronts Sartre's view on imagination, i.e., the imaginary makes reality irreal, with that of Deleuze's, i.e., the virtual is not opposed to reality but to the actual. And Breuer (2020) comments on Deleuze's identification of the virtual with the possible, yet the virtual still has a reality of its own. Moreover, she juxtaposes Deleuze's approach with Merleau-Ponty's account of the virtual as embedded in embodied actions.[12]

So far, the divide is clear, i.e., whereas Chalmers offers a classification of the phenomenology of virtuality in terms of perception only, a closer investigation – carried out by de Warren and others – unfolds different layers of presence of virtual objects, and thus virtual worlds. As a consequence, de Warren and others speak of the *imaginative* character of VR- and AR-experiences. For them, virtual objects are given in a distinct way which is irreducible to objects given in perception. Thus, the main disagreement here concerns the question of perception *or* imagination. Of course, if one opts for the latter, it would be necessary to adopt a kind of fictionalism in regard to VR; after all, from a phenomenological point of view, virtual objects are commonly regarded as *irreal*,[13] and thus non-real. However, two problems arise in this context. First, it is difficult to keep the language of imagination in regard to MR-technologies; after all, MR-technologies make virtual objects indistinguishable from the physical world (e.g., O'Shiel 2020, 2022; Turner 2022). Here, one still perceives virtual objects, and does not imagine them. How shall one describe this type of experience? Second, phenomenologists use the term "virtuality" in a very broad sense, which does not correspond to the common usage within VR-studies. At bottom, if the virtual is bound with imagining, and thus with non-perceiving, all objects which are not present in perception are virtual.[14] This conclusion, however, is counter-intuitive, and moreover this broad sense of the term "virtuality" does not seem to fit today's debates on the virtual.

9.4 Confronting Blaustein with the Virtual

Blaustein formulated the basics of his phenomenology of mediated experiences in the 1930s. Of course, he did not explore the virtual in the sense discussed in the previous section of this study. With this in mind, a preliminary methodological remark would be instructive. If one wishes to confront Blaustein with the virtual, one shall focus on the topics in his thought which can be rephrased and adopted in the context of recent debates; against this background, one can draw parallels between compared elements, and by doing this, one attempts to determine a conceptual framework – rooted in Blaustein's philosophy – which enables one to precise some nuances in the phenomenology of virtuality. It is clear that such parallels can be drawn at least in regard to a few points. To begin with, (1) the term "VR" was first introduced by Artaud in 1938 in French as "la réalité virtuelle" in regard to theater; for him, the stage generates a non-real environment for presenting fictional characters and internal dramas. Indeed, some scholars notice a clear parallel between theater-experiences and VR-experiences (e.g., Chalmers 2017; Metzinger 2018, 313–314). Given that Blaustein also explored this type of experience, one can expect that his language may fit VR-experiences as well. Next, (2) one of the central controversies today over the virtual lies in the question of virtual objects; Blaustein explored the structure of objects constituted in complex mediated experiences. Finally, (3) he widely investigated the phenomenon of imaginative perception, which intentionally indicates so-called imaginative objects; in this regard, scholars have recently, as shown above, examined imagination as an important factor which shapes experiences of virtuality. Of course, to reiterate, the sketched parallels give some clues as to how Blaustein's philosophy can be used in today's debates, yet further explorations seem to be necessary. Let me begin with theater-experiences.

(1) The parallel between theater- and VR-experiences lies in the plain fact that spectators in a theater, just like VR-users, perceive objects, say, stage props or characters performed by actors, as non-natural objects. Theater (and VR) generates an environment in which a spectator treats objects as "something more" than what is actually given. Blaustein (2011a), as shown above, claims that one observes in theater objects as a mixture of what is intuitively and non-intuitively given. Stage props or actors on stage are *not* just ordinary objects; their meaning is co-constituted by the world generated by the play. Therefore, a spectator shall change her attitude, which is normally focused on *closer* or *proper* objects, and instead focus on *distant* or *improper objects* (in Blaustein's sense). The same holds for VR-experiences in which objects are comprehended *as* virtual only if the user is focused on them *as* virtual. In this regard, virtual objects would be classified as *distant* or *improper* objects which are *non-identical* with

ordinary objects. Just like theater-goers, VR-users are generally aware that the objects they perceive require a specific attitude. This leads to another point. Blaustein seems to agree that one perceives stage props or characters performed by actors on stage only if she adopts a specific attitude directed toward the imaginative objects and, correlatively, the imaginative world. While describing plinths elevating sculpted figures, Blaustein (2011a, 219) compares this function with a curtain in a theater as follows:

> [the plinth] performs a function analogous to the frame of a painting or a curtain in the theatre, since raising the curtain perfectly symbolizes the shift from the attitude towards the real world to the attitude towards the imaginary world.

The shift of attitudes mentioned here can be described as a phenomenon of *immersion*: a theater-goer is entirely immersed in the imaginative world which includes the characters' psychic life, drama and stage props. Blaustein (2005, 97–99; 2011a, 219–220) goes a step further by claiming that an analogous phenomenon is experienced by cinema-goers watching a movie. In this case, the immersion is intensified by, e.g., music which ushers goers into a certain emotional mood (Blaustein 2005, 111–115).

The phenomenon of *immersion*, specified by Blaustein as the shift of attitudes, is widely discussed in VR-studies (e.g., Chalmers 2017, 312; Metzinger 2018; Liberati 2018, 217). However, an important nuance has to be noticed. Theater- or cinema-experiences take place in a specific environment and circumstances; after all, goers perceive imaginative worlds on stage or on screen. The phenomenon of the limits of these worlds are often unnoticed, e.g., theater-goers do not observe a stage which is physically distanced but rather are immersed in the drama performed by actors; the same holds for cinema-goers who are aware of the darkness around the screen but are focused on the imaginative world represented "on" the screen (Blaustein 2005, 101). In the case of VR- or AR-devices, the framework of the virtual is usually determined by the screens of these devices (Liberati 2018, 217). The problem arises with MR-devices when virtual objects are indistinguishable from the physical world (O'Shiel 2022; Turner 2022). It seems that the latter case would delineate the limit of a possible use of Blaustein's phenomenology in regard to the topic discussed here; after all, if virtual objects are indistinguishable from the physical world, in contrast to objects given in a theater on stage or in a cinema on screen, no shift of attitude can take place.

(2) According to Blaustein, objects given in mediated aesthetic experiences are complex and multi-layered. For him, imaginative objects are founded on non-imaginative ones. He classifies the former as *distant* or *improper* while the latter are described as *closer* or *proper*. These distinctions

are made by Blaustein in regard to different ways of experiencing objects: some are experienced as fully adequate (closer or proper objects, e.g., phantoms on the cinema screen), and some are experienced as *quasi*-adequate (distant or improper objects, e.g., characters performed by actors on stage). For him, imaginative objects and these imaginative worlds are *quasi*-real. With these ideas in mind, how shall one describe the structure of virtual objects? While experiencing virtual objects, one adequately experiences only groups of some colorful surfaces and shapes; as such, they manifest themselves in a way which is mediated by relevant devices, say, a VR-headset or a phone-screen. Even though one experiences here phantoms, in the end, one perceives virtual objects. To use Blaustein's language, in this case, phantoms are apprehended by the VR-user, and constituted *as* virtual objects. As such, they are *distant* or *improper* objects, which are constituted on the basis of *closer* or *proper* objects. To phrase it differently, virtual objects are given as mediated objects in the sense that they are constituted on other objects. As such, they are *quasi*-real; this means that they form an environment distinct from the ordinary world, and finally, virtual objects are parts of imaginative worlds.[15]

Blaustein's ideas are marked by two clear reformulations in relation to recent debates on VR. First, Blaustein would not agree – as I suppose – with Chalmers (2017, 312; 2022) that virtual objects are digital objects what means that the distinctive feature of them is being constituted by computational processes on a computer. For Chalmers, this leads to the claim that virtual objects are real. By contrast, for Blaustein, the causal basis of virtual objects is irrelevant as in the end they are constituted as *quasi*-real. So even though the digital processes on a computer are an existential basis for virtual objects, so-to-speak, Blaustein adopts a kind of metaphysical neutrality by suggesting that they become partly autonomous; after all, virtual objects are experienced by the user, and without the user only digital processes exist. Second, as Beisbart (2019) shows, Chalmers's point of view is problematic, as it is hard to understand that properties of virtual objects which are perceived by the user reflect the properties of data structures which are causally related to real computers in the real world. For Blaustein, the process of ascribing properties is simpler, as virtual objects arise as distant or improper objects which, from a mereological viewpoint, are new subjects of properties generated by relevant acts of perception.

(3) For Blaustein, the way objects are given in mediated experiences differs from non-mediated experiences. For instance, while perceiving a tree in a garden, the relation between the presenting content and the object is adequate, i.e., all elements of the former relate to elements of the latter. By contrast, while perceiving an actor on stage, the relation between the presenting content and the represented object is *quasi*-adequate (Blaustein

2011a, 216). Perceiving actors *as* performing characters is not intuitively given. In any case, while observing the play as a fictional object, this act constitutes the latter as the intentional object of theater-goers' experiences. Thus, finally, one perceives, following Blaustein's example discussed above, Caesar who is talking with Cleopatra, and *not* an actor imitating Caesar, etc. An analogous change is noticed by him in film-experiences. According to Blaustein (2005, 97), reference to an object (instead of the contents of experiences) is founded in the cinema-goer's *attitude* toward the object: if one perceives a movie *as* a movie, one apprehends objects "in" the movie, but if one focuses on the objects *represented* by the movie, as, e.g., in a documentary, one is directed toward objects "in" the real world. To describe this phenomenon, which is specific to a cinema-goer, Blaustein coined the term "observation," described as attentive living through the presentations. Next, in the field of the radio-listener's experiences, Blaustein (1939, 111; 2005, 149) shows that the phenomenon of acousion (as analyzed above) consists in directly perceiving objects of a radio drama. Sounds which are identified by Blaustein as the content of acousion serve to present the objects *of* these very sounds. Sounds intentionally indicate "their" objects, but the listeners of a radio drama are directed towards the represented objects instead of the presenting sounds. One "really" perceives sounds *as* the sounds *of* represented objects, e.g., the sound *of* a telephone ring or the murmurs *of* a departing train. Of course, "[i]f there are no 'reproducing objects,' if there is no representation, then of course no imaginative world is constituted" (Blaustein 1939, 112; 2005, 157). Blaustein notices a clear shift in focus: one does not hear "mere" sounds but objects themselves "directly," for instance, a phone ringing or a train. How do these ideas of Blaustein's correspond with VR-experiences?

First, VR-experiences can be described as essentially imaginative, i.e., they present their objects *quasi*-adequately. The user does not comprehend digitally generated images as groups of colorful shapes, but one perceives virtual objects constituted on the basis of the very presenting contents. Second, VR-experiences are possible due to the shift of attitude of the user. Just as theater-, cinema-goers, or radio-listeners (in Blaustein's view) adopt a specific attitude which makes imaginative objects accessible, VR-users are aware that one shall be directed toward virtual objects while using relevant devices. Third, this shift is also possible because of a specific environment. VR-users usually operate with relevant mediatory devices in order to experience VR. In this regard, Liberati (2018, 218) shows that AR-devices are tools which function in a specific way: "[b]y looking through the device the subjects turn their surroundings into a surface where to visualize digital objects." Fourth, VR-experiences are in the end constituted on devices and digital processes generated by computers (Chalmers 2017; 2022). This, however, does not mean that VR-experiences stop

being imaginative; to be precise, even though experiences are imaginative, one "really" experiences *quasi*-real (virtual) objects.

9.5 Conclusions

To conclude this study, recall that my aim here was not to define a full-scale theory of the virtual. Given the undeniable fact of still-developing VR-technologies which are close to being indistinguishable from the real world (e.g., O'Shiel 2020), scholars shall seek flexible languages in an attempt at describing VR-experiences. With this in mind, in the present study, I discussed parts of Blaustein's aesthetics of media experience, and I examined to what extent his conceptions can be useful in the context of today's studies on the virtual. I proceed by drawing parallels between Blaustein's analysis of lived experiences (in the fields of perception of theater plays, cinema movies, and radio dramas) on the one hand, and recent readings in VR, mainly Chalmers and his critics, de Warren, O'Shiel, and Liberati, on the other. To do so, I have looked closer at the main ideas of Blaustein's aesthetics: (1) the distinction of *closer-distant* or *proper-improper* objects, (2) the phenomenon of *imaginative attitude*, (3) his examination of *directedness of perception*, and (4) the concept of *imaginative worlds* that exist as *quasi*-real. In addition, I have argued that two factors are decisive in classifying Blaustein's analysis as phenomenological: (1) he adopts descriptive tools to analyze the first-person perspective of mediated experiences, and (2) he emphasizes the ways in which the objects of these experiences are experienced, i.e., he asks about the "as" of the objects manifested in these experiences. This approach enables Blaustein to comprehend a radio drama or a movie displayed in a cinema as non-neutral, technically mediated objects which determine one's way of experiencing. Keeping this important point in mind, I have suggested that Blaustein's ideas can be helpful in describing VR-experiences which, to summarize my results so far, were described as (1) *intentionally* directed toward *distant* or *improper* objects, (2) requiring a shift into an *imaginative* attitude, (3) *direct* and *immersed*, and, finally, (4) constituting the *imaginative world* which is *quasi*-real, though constituted on the basis of real *digital processes*. Thus, how would one classify Blaustein if he were confronted with the issue of VR-experiences?

To address this question, in conclusion, recall that today's debates on VR are basically divided into realism-or-fictionalism camps (e.g., Chalmers 2017; 2022; McDonnell and Wildman 2019). In this regard, Blaustein may seem to argue in favor of a kind of fictionalism as he saw objects perceived in mediated (virtual) experiences as *imaginative* and thus *quasi*-real. This classification is even more evident if one refers to Blaustein's (2005, 125) approval for Geiger's (1908, 320) idea that

an illusionary (or irreal) character of objects is the basis for specifically aesthetic experiences; for Blaustein, who followed Geiger in this respect, the illusion in media experiences is founded in the phenomenon of the pictorial nature of what is presented. If VR-experiences are illusionary in this sense, Blaustein should be classified as a proponent of fictionalism. However, as I demonstrated in this study, this classification is only partly adequate. Even though he opted to understand mediatory experiences as *imaginative*, he did not refuse a kind of (phenomenological) realism according to which one "really" perceives distant or improper objects of mediated experiences. After all, Blaustein comprehended perception as *direct*, i.e., he described it as *de re* intentionality. In closing, however, let me note that the use of Blaustein's ideas have to be limited to VR-, AR-, and MR-technologies which are distinguishable from the real world. From Blaustein's point of view, the shift of attitude so central to his approach would be impossible if the user were not aware of their being immersed in VR.

Acknowledgments

This work was supported by the National Science Centre, Poland as a part of the SONATA BIS program (No. 2021/42/E/HS1/00108) within the research project on "The Philosophy of Leopold Blaustein in Contexts: Brentano, Gestalt Psychology, Lvov–Warsaw School and Early Phenomenology."

Notes

1 On Blaustein's biography, see Miskiewicz (2009); Pokropski (2015), 93–94; Płotka (2020b).
2 More on Blaustein's position in the contexts of the Lvov-Warsaw School, Husserl and Ingarden, see Citlak (2019); Płotka (2020a; 2021).
3 "Given this primacy of the lived experience of fictions, the relative paucity of philosophical reflection on the imagination – without which the experience of fiction is impossible – in the sprawling literature on videogames and virtual reality is striking. The exercise of the imagination is repeatedly acknowledged, yet rarely receives the sustained analysis required for understanding its role in the constitution of simulated realities, interactive fictions, and virtual lives as meaningful lived experiences" (de Warren 2014, 92). For an overview, see also O'Shiel (2019; 2022).
4 Blaustein (2011b). For discussion, see Rosińska (2013).
5 "Even though Blaustein did not use the tools of *epoché*, imaginative variation (like Husserl), or the investigation of the content of ideas (like Ingarden), he followed the basic intuition that analysis should be focused on an object as it is presented or manifested in experience. This is why [...] Blaustein's project of descriptive psychology is *phenomenologically* oriented" (Płotka 2021, 265).

6 By claiming this, Blaustein followed neither Husserl, who would prefer to talk about essences, nor Ingarden, who comprehended ideas as proper objects of phenomenological analysis. His position is close to Twardowski's.
7 Blaustein's idea that some objects are *quasi*-real stems from Ingarden's philosophy. For the latter, the character of "fiction" is given in a modification of the "thetic" moment of the given object. In the end, this modification presents the "fiction" as *quasi*-reality. See, e.g., Ingarden (1973, 48).
8 For an overview of Blaustein's view of the phenomenon of listening to the radio drama, see Ciccotti (2014, 147–161).
9 "Ultimately, anything on the reality-virtual continuum, if it is not a simple case of reality or virtuality, is a mixture between the two" (O'Shiel 2020, 5).
10 Thus, "[t]he virtual world, virtual dragons and all, is part of the real world, in virtue of existing on real computers" (Chalmers 2017, 320).
11 Husserl's phenomenology of image-consciousness is widely discussed in the literature; see recent studies by Jansen (2016); Summa (2021); Geniusas (2022, 37–73); Hui (2022).
12 "Merleau-Ponty conceives of the virtual in terms of the possible, while virtuality is defined as a potentiality that indicates the direction of possible actualisation of actions. The virtual is thus a projection or extension of the phenomenal fields, which supplies the subject with an open horizon of possibilities and meaning" (Breuer 2020, 9).
13 To be precise, the Husserlian category of irreality does not oppose the real; instead, it refers to a way of presentation which is ontologically neutral. On this issue, see Aldea (2013, 373–381).
14 This consequence is spelled out explicitly by O'Shiel (2020, 2): "Although Husserl never uses the term 'virtual' himself, it quite clearly fits perfectly for that which is not but could easily become present in my perceptual field within his highly detailed phenomenological theory. A very simple example is the implicit virtuality of the room next to from where I currently am, and then the exterior of the house beyond that, then the road, the town and so on."
15 As already noticed above, the idea of *quasi*-reality stems from Ingarden's philosophy. For use of Ingarden's ideas in VR-studies, see Skowron (2020, 619–639).

References

Aldea, Andreea Smaranda. 2013. "Husserl's Struggle With Mental Images: Imaging and Imagining Reconsidered." *Continental Philosophy Review* 46(3): 371–94. https://doi.org/10.1007/s11007-013-9268-7

Beisbart, Claus. 2019. "Virtual Realism: Really Realism or Only Virtually so? A Comment on D. J. Chalmers's *Petrus Hispanus* Lectures." *Disputatio* 11(55): 297–331. https://doi.org/10.2478/disp-2019-0008

Blaustein, Leopold. 1931. *Przedstawienia schematyczne i symboliczne. Badania z pogranicza psychologji i estetyki*. Lvov: Nakładem Przeglądu Humanistycznego.

Blaustein, Leopold. 1937. "Das imaginative Kunstwerk und seine Gegebenheitsweise." In *deuxième Congrès International d'esthetique et de science de l'art, Paris, 1937*, edited by Paul Valéry, Paul Claudel and Victor Basch, 245–49. Paris: Librairie Félix Alcan.

Blaustein, Leopold. 1939. "Étude sur la perception des pièces radiophoniques." *Kwartalnik Psychologiczny* 11: 105–61.

Blaustein, Leopold. 2005. *Wybór pism estetycznych*, edited by Zofia Rosińska. Krakow: TAiWPN Universitas.
Blaustein, Leopold. 2011a. "Imaginary Representations: A Study on the Border of Psychology and Aesthetics." Trans. Monika Bokiniec. *Estetika: The Central European Journal of Aesthetics* 48(2): 209–34.
Blaustein, Leopold. 2011b. "The Role of Perception in Aesthetic Experience." Trans. Monika Bokiniec. *Estetika: The Central European Journal of Aesthetics* 48(2): 235–43.
Breuer, Irene (2020). "Body and World: The Correlation Between the Virtual and the Actual Through Phenomenological Reflections via Merleau-Ponty and Deleuze." *Indo-Pacific Journal of Phenomenology* 20(1): 1–10. https://doi.org/10.1080/20797222.2020.1863564
Chalmers, David J. (2017). "The Virtual and the Real." *Disputatio* 9(46): 309–52. https://doi.org/10.1515/disp-2017-0009
Chalmers, David J. (2019). "The Virtual and the Digital." *Disputatio* 11(55): 453–86. https://doi.org/10.2478/disp-2019-0022
Chalmers, David J. (2022). *Reality +: Virtual Worlds and the Problems of Philosophy*. New York, NY: W.W. Norton & Company.
Ciccotti Eusebio. 2014. "The Philology and Aesthetics of Radio According to Leopold Blaustein." In *Philology and Performing Arts: A Challenge*, edited by Mattia Cavagna and Costantino Maeder. Louvain-la-Neuve: Presses universitaires de Louvain, 147–61.
Citlak, Amadeusz. 2019. "The Problem of Mind and Mental Acts in the Perspective of Psychology in the Lvov-Warsaw School." *Philosophical Psychology* 32(7): 1049–77. https://doi.org/10.1080/09515089.2019.1647410
De Warren, Nicolas. 2014. "Towards a Phenomenological Analysis of Virtual Fictions." *Metodo. International Studies in Phenomenology and Philosophy* 2(2): 91–112.
Flavián, Carlos, Sergio Ibáñez-Sánchez, and Carlos Orús. 2019. "The Impact of Virtual, Augmented and Mixed Reality Technologies on the Customer Experience." *Journal of Business Research* 100: 547–60. https://doi.org/10.1016/j.jbusres.2018.10.050
Geiger, Moritz. 1908. "Ästhetik." In *Systematische Philosophie*, edited by Paul Hinneberg, 311–51. Berlin: B. G. Teubner Verlag.
Geniusas, Saulius. 2022. *Phenomenology of Productive Imagination: Embodiment, Language, Subjectivity*. Stuttgart: ibidem Press.
Heim, Michael. 1993. *The Metaphysics of Virtual Reality*. New York, NY, Oxford: Oxford University Press.
Hui, Ka-yu. 2022. "The *Hyle* of Imagination and Reproductive Consciousness: Husserl's Phenomenology of Phantasy Reconsidered." *Husserl Studies* 38: 273–92. https://doi.org/10.1007/s10743-022-09308-2
Ingarden, Roman. 1973. *The Cognition of the Literary Work of Art*. Translated by Ruth Ann Crowley and Kenneth R. Olson. Evanston: Northwestern University Press.
Jansen, Julia. 2016. "Husserl." In *The Routledge Handbook of Philosophy of Imagination*, edited by Amy Kind, 69–81. New York, NY, London: Routledge.
Liberati, Nicola. 2018. "Phenomenology, *Pokémon Go*, and Other Augmented Reality Games: A Study of a Life Among Digital Objects" *Human Studies* 41: 211–32.
McDonnell, Neil, and Nathan Wildman. 2019. "Virtual Reality: Digital or Fictional?" *Disputatio* 11(55): 371–97. https://doi.org/10.2478/disp-2019-0004

Metzinger, Thomas K. 2018. "Why Is Virtual Reality Interesting for Philosophers?" *Frontiers in Robotics and AI* 5: 101. https://doi.org/10.3389/frobt.2018.00101

Milgram, Paul, and Fumio Kishino. 1994. "A Taxonomy of Mixed Reality Visual Displays." *IEICE Transactions on Information and Systems* 77(12): 1321–29.

Miron, Ronny. 2021. *Hedwig Conrad-Martius: The Phenomenological Gateway to Reality*. Cham: Springer. https://doi.org/10.1007/978-3-030-68783-0

Miskiewicz, Wioletta. 2009. "Leopold Blaustein's Analytical Phenomenology." In *The Golden Age of Polish Philosophy. Kazimierz Twardowski's Philosophical Legacy*, edited by Sandra Lapointe, Jan Woleński, Mathieu Marion, and Wioletta Miskiewicz. Dordrecht: Springer, 181–88. https://doi.org/10.1007/978-90-481-2401-5_13

Moran, Dermot, and Rodney K.B. Parker. 2015. "Editors' Introduction: Resurrecting the Phenomenological Movement." *Studia Phaenomenologica* 15: 11–24. https://doi.org/10.5840/studphaen2015151

O'Shiel, Daniel. 2019. "Phenomenology and the Challenge of Virtuality." In *Conceiving Virtuality: From Art to Technology*, edited by Joaquim Braga. Cham: Springer, 21–43. https://doi.org/10.1007/978-3-030-24751-5_2

O'Shiel, Daniel. 2020. "Disappearing Boundaries? Reality, Virtuality and the Possibility of 'Pure' Mixed Reality (MR)." *Indo-Pacific Journal of Phenomenology* 20(1): 1–8. https://doi.org/10.1080/20797222.2021.1887570

O'Shiel, Daniel. 2022. *The Phenomenology of Virtual Technology: Perception and Imagination in a Digital Age*. London, New York, NY, Oxford, New Delhi, Sydney: Bloomsbury Academic.

Parker, Rodney K.B. (ed.). 2021. *The Idealism-Realism Debate Among Edmund Husserl's Early Followers and Critics*. Cham: Springer. https://doi.org/10.1007/978-3-030-62159-9

Płotka, Witold. 2020a. "From Psychology to Phenomenology (and Back Again): A Controversy over the Method in the School of Twardowski." *Phenomenology and the Cognitive Sciences* 19(1): 141–67. https://doi.org10.1007/s11097-019-09620-x

Płotka, Witold. 2020b. "Leopold Blaustein's Descriptive Psychology and Aesthetics in Light of His Criticism of Husserl." In *Early Phenomenology in Central and Eastern Europe: Main Figures, Ideas, and Problems*, edited by Witold Płotka & Patrick Eldridge. Cham: Springer, 163–85. https://doi.org/10.1007/978-3-030-39623-7_10

Płotka, Witold. 2021. "A Critical Analysis of Blaustein's Polemic Against Husserl's Method". *Husserl Studies* 37(3): 249–70. https://doi.org/10.1007/s10743-021-09292-z

Płotka, Witold, and Patrick Eldridge (eds.). 2020. *Early Phenomenology in Central and Eastern Europe: Main Figures, Ideas, and Problems*. Cham: Springer. https://doi.org/10.1007/978-3-030-39623-7

Pokropski, Marek. 2015. "Leopold Blaustein's Critique of Husserl's Early Theory of Intentional Act, Object and Content." *Studia Phaenomenologica* 15: 93–103. https://doi.org/10.5840/studphaen2015156

Rauschnabel, Philipp A. 2021. "Augmented Reality Is Eating the Real-World! The Substitution of Physical Products by Holograms." *International Journal of Information Management* 57: 102279. https://doi.org/10.1016/j.ijinfomgt.2020.102279

Rosińska, Zofia. 2011. "Leopold Blaustein's Aesthetics." *Estetika: The Central European Journal of Aesthetics* 48(2): 199–208.

Rosińska, Zofia. 2013. "The Model of Aesthetic Experience. The Theory of Leopold Blaustein." In *20th Century Aesthetics in Poland. Masters and Their Followers*, edited by Krystyna Wilkoszewska, 74–94. Warszawa: Semper.

Skowron, Bartłomiej. 2020. "Virtual Objects: Becoming Real." *Horizon. Studies in Phenomenology* 9(2): 619–39. https://doi.org/10.21638/2226-5260-2020-9-2-619-639

Summa, Michela. 2021. "How Are Fictions Given? Conjoining the 'Artifactual Theory' and the 'Imaginary-Object Theory'." *Synthese* 199: 13749–69. https://doi.org/10.1007/s11229-021-03396-x

Turner, Cody. 2022. "Augmented Reality, Augmented Epistemology, and the Real-World Web." *Philosophy & Technology* 35: 19. https://doi.org/10.1007/s13347-022-00496-5

Vendrell Ferran, Íngrid. 2008. *Die Emotionen. Gefühle in der realistischen Phänomenologie*. Berlin: Akademie Verlag.

Walton, Kendall L. 1990. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press.

Walton, Kendall L. 2013. "Fictionality and Imagination Reconsidered." In *Fictionalism to Realism: Fictional and Other Social Entities*, edited by Carola Barbero, Maurizio Ferraris and Alberto Voltolini, 9–26. Newcastle upon Tyne: Cambridge Scholars Publishing.

10 From Immersive Body Swapping to Apprehending the Other's Emotions

Perspective-Taking and Levels of Empathy in Embodied Virtual Reality

Íngrid Vendrell Ferran

10.1 Introduction

Natural scientists working at the intersection of virtual reality (VR), psychology, and computer science have recently explored the question of whether embodied virtual reality (EVR) can be employed to train empathy.[1] Yet, opinions are divided in this debate between those who establish a link between EVR and empathy and those who regard this link with skepticism.

On the one hand, authors such as Bertrand et al. (2018) argue that EVR can enhance empathy. In their view, EVR leads users to adopt the other's perspective and resonate with her experience. For these authors, applying multisensory and motor stimuli synchronically with the first-person perspective of an avatar, immersive VR might lead to a series of perceptual illusions, which generate in its users the impression of swapping bodies with the avatar (Lenggenhager et al. 2007; Maselli and Slater 2013). In this vein, Maister et al. (2014) exposed participants to bodily illusions that induced ownership over the body of another person pertaining to an outgroup regarding gender, age, or race. As a result, participants experienced a reduction in biases against that outgroup. This shift of perspective enabled by EVR makes it a powerful tool to reduce negative biases or increase positive responses toward individuals who are very different from us. As Bertrand et al. have put it: "Experiences of EVR allow users to literally step into the shoes of others and see the world from their perspective. Research on EVR has explored how manipulations of the senses can be used to modulate empathic responses" (Bertrand et al. 2018, 10).

On the other hand, authors such as Sutherland (2016) and Sora-Domenjó (2022) have been more skeptical about the powers of EVR to increase empathy in its users. In an attempt to demystify the powers of VR to foster empathy, Sora-Domenjó has argued that

> there is little empirical evidence of a correlation between VR exposure and an increase in empathy that motivates pro-social

DOI: 10.4324/9781003312284-11

> behavior, and a lack of research covering VR films exposure eliciting empathy. Furthermore, the results show an alarming lack of research into the long-term effects of VR films and other VR immersive experiences.
>
> (Sora Domenjó 2022)

Despite acknowledging that VR is a powerful tool for perspective-shifting, it is still subject to cultural biases, stress, and so on. He concludes that "based on current research it's premature at this early stage to consider VR as a medium that elicits empathy over other media such as cinema, television or photography. Empirical evidence supporting the claim that immersive storytelling experiences enhances empathy is limited" (Sora Domenjó 2022, 10). For him, the mechanisms involved in VR neither lead to a better understanding of another person's feelings nor do they lead to a greater arousal of empathy. In his view, the contemporary debate on the link between VR and empathy, which is marked by the lack of a solid understanding of what empathy means, should acknowledge the complexity of empathy, and the term should not be deployed in a blurry and superficial sense.

Taking Sora-Domenjó's desideratum as a point of departure, this chapter explores how perspective-taking contributes to empathy in EVR.[2] In my view, the recent debate on EVR can be enriched with phenomenological insights on how perspective-taking takes place at different levels of empathy. Drawing on classical and contemporary phenomenology and in particular the works of Stein (1989) and Svenaeus (2018), the chapter applies to the EVR debate a distinction between sensual and emotional empathy. While in sensual empathy, the user empathizes with the other's lived body, in emotional empathy, she apprehends the other's emotional states and their related values. It is argued that while the perceptual illusions elicited by EVR can lead the user to the impression of a body swap and, in so doing, foster sensual empathy, for the apprehension of the other's emotions in emotional empathy, EVR must be supplemented with a series of narrative devices that make the other's emotions and values accessible to her.

The next section begins by exploring the role of perspective-taking in contemporary accounts of empathy (section 10.2). Next, I offer a micro-analysis of perspective-taking in non-virtual (section 10.3) as well as in virtual environments (section 10.4). The phenomenological distinction between sensual and emotional empathy as levels of empathy is then introduced (section 10.5). I argue that EVR fosters a form of perspective-taking that leads to sensual empathy but that to foster emotional empathy, narrative techniques must be involved (section 10.6). The conclusions summarize the main findings (section 10.7).

10.2 In the Other's Shoes: Perspective-Taking in Contemporary Accounts of Empathy

Philosophical debate about empathy has been marked by strong controversies regarding how the apprehension of the other's experience takes place, and not all theories draw on perspective-taking to explain empathy.[3]

During the 1990s, the debate was dominated mainly by the Theory-Theory (TT) and the Simulation Theory (ST). The TT argues that empathy requires that we have a folk psychological theory of mind about the other with whom we empathize (Carruthers and Smith 1996). For the TT, perspective-taking is not necessary for empathy. By contrast, according to the ST, which draws on Lipps' imitation theory, we need to re-create, re-enact, or simulate what the other is going through (Goldman 2006; Stueber 2006). Traditionally, as understood by Lipps (1903) in his imitation theory, which can be regarded as a proto-simulationist account of empathy, empathy requires "feeling into" the empathized object and imitating its feelings. The idea of "feeling into" can be explained in terms of a projection into the other's situation. Most of today's proponents of ST argue that for empathy to take place, the subject must put herself in the other's shoes.[4] In other words, empathy requires that we adopt the other's perspective.

Of the two theories, TT has lost momentum given that it leads to an over-intellectualization of the understanding of other people. Indeed, it is often the case that we can understand what the other is going through without having a theory of mind about it. Moreover, TT conflates empathy with mindreading. Unlike TT, ST remains an option in current research. As a result, perspective-taking is regarded as a necessary step in the empathic process.

In the last couple of decades, however, the Direct Perception Theory (DPT) has emerged as an alternative to the ST. According to the DPT, which builds on Scheler's (2008) phenomenological views, we can directly perceive in the other's expressions what she is going through (Zahavi 2011). For the DPT, perspective-taking is not necessary for empathy insofar as we can directly see the other's emotions in their expressive movements.

Though in the current debate proponents of the ST and the DPT present both mechanisms as offering alternative explanations of the apprehension of the other's experience, it is also possible to regard them as complementary. While the DPT can explain basic forms of empathy in which we immediately see what the other is experiencing (e.g., we see the sadness in the other's face), more complex forms of empathy, as argued by the ST, require us to project ourselves into the other's experience and re-enact what the other is going through. In this vein, there is a distinction made between basic empathy and re-enactive empathy.

Moreover, there are theories which can be regarded as endorsing a hybrid model in which both mechanisms are involved. For instance, Stein (1989) and some contemporary proponents of her view (e.g., Magrì 2018; Svenaeus 2018; Vendrell Ferran 2015) regard empathy as a process that evolves in stages. It starts with a – usually vague – perception of the other's experience. This first apprehension of what the other is going through can be made more accurate by means of transferring ourselves into the other's situation, i.e., by means of adopting the other's perspective. Finally, we come to grasp in a more nuanced way what the other is going through. In this vein, perspective-taking can enable a more accurate apprehension of the other's experience.[5]

To sum up, in today's research, perspective-taking is central for proponents of the ST and for those endorsing a hybrid model of empathy, such as the view of empathy as a process which starts with a perception of what the other is going through but whose completion requires the empathizer to put herself in the other's shoes.

10.3 A Micro-Analysis of Perspective-Taking in Non-Virtual Environments

Perspective-taking is not a homogeneous mental state. In fact, what in the debate is usually called "perspective-taking" and which traditionally has been described as "putting oneself in the other's shoes" entails different stages and imaginative activities, which I will refer to here as perspective-building, projection, and resonance.

10.3.1 Perspective-Building

Perspective-shifting requires first of all that we re-present to ourselves what the other is going through. In representing what the other is going through, the empathizer reconstructs in her mind the other's perspective.[6] This re-presentation necessarily entails imaginings of different types: propositional imaginings such as imagining that X is the case, imagining the other realizing an action, imagining the other experiencing something, and so on. In this re-presentation of the other's experience, we imagine "from the outside" and do not adopt the other's perspective. However, we cannot imagine "from the inside" without first re-presenting to ourselves what the other is going through.

Importantly, the re-presentation of the other's experience might differ in its "scope" (Vendrell Ferran 2023).[7] It can re-present a single experience such as a perception, imagining, emotion, belief, etc., or a combination of experiences such as a set of perceptions and emotions or a set of beliefs and desires. It can also re-present a set of possible experiences which are

characteristic of the way in which the other engages with her environment (the other's heart).

In addition, the re-presentation of the other's experience can be motivated by different cognitive phenomena such as a perception, an imagining, belief, memory, etc. Thus, it can be the case that the empathizer sees the other's situation, hears about it, imagines it, knows about it, remembers it, etc.

10.3.2 Projection

Once the other's perspective is re-presented to our mind, the empathizer can project herself into it. This is what the expression "in the other's shoes" literally means. The empathizer here not only imagines what the other is going through but she adopts the other's perspective. This imagining being in the other's place is an imagining "from the inside."[8] In the relevant literature, there is a distinction between "self-" and "other-oriented perspective-shifting" (Coplan 2011). While in self-oriented perspective-shifting, I imagine how it would be for me to be in the other's situation, in other-oriented perspective-shifting, I imagine how it is for the other.

Some authors have argued that other-oriented perspective-shifting is not possible, either because we do not have information about the other or because in order to construct the other's experience we have to resort to our own experience and our psychology can be very different from that of the other so that we end up imagining the other, but this imagined other has nothing to do with the real other. Moreover, for Goldie (2011), other-oriented perspective-shifting is not desirable because if we imagine to be the other, then in empathy the distinction between self and other vanishes, so that empathy as such does not take place. Though I think that we should be aware of these criticisms, in this chapter, I will adopt the distinction between both forms of perspective-shifting for analytical purposes. To begin, though other-oriented perspective-shifting requires that we leave aside our individual empirical self and we cannot renounce the person we are, a step toward other-oriented perspective-shifting can be made possible by gathering information about the other and by abstracting momentarily from our selves. Moreover, we can keep the self–other differentiation necessary for empathy by means of switching perspectives.

10.3.3 Resonance

In adopting the other's perspective, we imagine "from the inside" what the other is going through. Imagining from the inside is the hallmark of

"experiential imagining" (Kind 2016). This means that we not only imagine the other's experience but we imagine it experientially, i.e., we imagine undergoing it. In this respect, the empathizer comes to re-create, re-enact, or simulate the other's experience by re-living it. In recreating the other's experience, the empathizer simulates the other's perceptions, emotions, beliefs, and so on. These recreated mental states are often described in terms of "quasi-perceptions," "quasi-emotions," "quasi-beliefs," etc. This experiential imagining leads to a vicarious experience of what the other is going through. Experientially imagining X generates an X-like state.[9] Thus, if we imagine experiencing the other's perceptions, emotions, desires, etc., we generate a similar state in us.[10] This re-creation of the other's experience motivated by perspective-shifting is what leads us to resonate with the other's experience.[11]

I will come back to this distinction between different steps in the process of perspective-taking below when analyzing the role of EVR for empathy.

10.4 Perspective-Taking in EVR: The Illusions of Presence, Embodiment, and Agency

Let's take the established view on perspective-taking in non-virtual environments and apply it to perspective-taking in virtual contexts. As noted in the Introduction, EVR enables us to put ourselves in the other's shoes. By means of different audiovisual and motor devices, such as an oculus rift, headphones, etc., we can adopt the other's perspective, creating the impression of swapping perspectives with the other. These technological devices enable the user to "feel into" the other's situation so that the EVR user has the impression that a real perspective-shifting is occurring. How does perspective-shifting happen in virtual environments compared to in non-virtual environments?

10.4.1 Perspective-Building in EVR

In virtual environments, the perspective-shifting exhibits a series of particularities. To begin with, in virtual environments, the first step of perspective-shifting, which I referred to above as perspective-building, requires a particular interplay of perception and imagination. Indeed, in EVR, some aspects of what the other is going through are made accessible by means of the audiovisual and sensory devices. These techniques enable the user to perceive what the avatar is perceiving. We have access to what the other sees, hears, and so on. However, this access is limited to certain sensations and, relatedly, to the spatial localization of the other.

For mental states other than perceptions, the user is required to imagine. For instance, using an oculus rift, the user can see what the other is supposed to see. Yet, she does not have access to the other's beliefs, expectations, or emotions. Such mental states have to be imagined by the user in order to re-present to herself what the other is going through. As a result, the construction of the other's experience takes place with the aid of technical devices in EVR. By contrast, in non-virtual environments, the empathizer can build the other's experience by means that do not necessarily involve perception. For instance, the empathizer can imagine, remember, suppose, etc., what the other is going through but does not necessarily have to perceive the other's situation. In addition, the use of technical devices is less prominent in non-virtual environments than in virtual ones.

The involvement of technical devices plays an important role in determining what I called above the scope of the other's experience. More precisely, the transmitted scope is limited to perceptual experiences. Basically, EVR will provide us with single perceptual experiences or combinations of perceptual experiences. For instance, by means of an oculus rift, the user sees what the other sees, or by means of headphones, she hears what the other hears, etc., or an oculus rift and headphones can be combined. These perceptions give the user the impression of being "on the inside," i.e., of embodying the other's perspective. As argued in the literature, this sense of embodiment entails the sense of self-location, the sense of agency, and the sense of body ownership (Bertrand et al. 2018; Kilteni, Groten, and Slater 2012). For proponents of simulationism, this means that perspective-taking takes place in virtual environments on the basis of perceptions. For proponents of the hybrid model of empathy, this means that, at the first step, what is presented by our perception is not vague but accurate, albeit limited in its scope.

10.4.2 Projection in EVR

The perceptions mediated by the different technical devices enable the user to adopt the other's perspective. The user sees, hears, etc. the world as the other does. The user "feels into" the other's experience. At this stage, the perceptions enabled by EVR are responsible for a series of illusions, which Bertrand et al. (2018) have classified as illusions of presence, embodiment, and agency.

10.4.2.1 Illusion of Presence

In VR, in the expression "illusion of presence," presence is not necessarily related to having a body, but as the feeling of "being there." Slater (2009) distinguishes between "Perceptual Illusion" (PI) as the illusion of being in a place which is constrained by sensorimotor contingencies of VR, and

the "Plausibility Illusion" (Psi), which is the credibility of the scenario and concerns the illusion that the scenario is taking place.[12]

10.4.2.2 Body Ownership Illusion

The body ownership illusion consists in the feeling of owning an artificial body which acts as a substitute for the real body as the origin of perceptual sensations (Maselli and Slater 2013). For the body ownership illusion, multisensory and motor stimuli have been applied in synchronicity with the first-person perspective of an avatar. This can occur by means of computer-generated imagery (Maselli and Slater 2013) or through the image of a real human (Petkova and Ehrsson 2008). As argued by Bertrand et al., the most explored mechanism is visuomotor synchronicity (here we see ourselves in the body of an avatar who mimics our movements in real time) and visuotactile synchronicity (we see tactile stimuli applied to the avatar at the same time which is applied to the user's congruent body part). Importantly, as shown by Maselli and Slater (2013) in their study of the perceptual components of the full body ownership illusion as a specific type of bodily illusion in which a user experiences an artificial body as if it were her own, a combination of stimuli can promote embodiment illusions but do not always require visuomotor or visuotactile stimulation. Incongruent perception does not break the bodily illusion.[13] Embodiment might occur voluntarily or involuntarily.

10.4.2.3 Agency Illusion

Agency illusion is different from the embodiment illusion (Sato and Yasuda 2005). Agency is voluntary. Bertrand et al. (2018) describe the illusion in the following terms: "By embodying a digital avatar that could be controlled by the user's movements, researchers observed self-attribution of agency to subjects over actions taken by the avatar, even without any prior intention, prediction, priming, and cause preceding effect" (Banakou and Slater 2014). According to their experiment, the digital avatar would speak independently of the user's action creating not only the perception that subjects were themselves talking, but also changing the fundamental frequency of the user's voice after the experience. This illusion was found to be even stronger when a vibration stimulus was applied to the user's throat in synchronicity with the avatar's voice.

Importantly, in contrast to what occurs in non-virtual environments in which perspective-taking relies heavily on imagination, given these illusions, in virtual contexts, the user does not imagine how it would be for her to see, hear, etc., if she were in the other's situation; rather, she really sees, hears, etc., what the other is supposed to see, hear, etc. This leads

to an interesting interplay of self- and other-oriented perspective-taking. In having the illusions, the empathizer in virtual contexts might have the impression that she is undergoing the avatar's experience. Yet, for empathy to take place, other-oriented perspective-taking is necessary. Thus, the empathizer must realize each time that not only is she having these experiences but that these experiences belong to the other. She must have the impression of swapping perspectives with the other while maintaining the self–other differentiation.

10.4.3 Resonance in EVR

The empathizer does not just experience a recreation of what the other is going through; she has the impression of actually undergoing the same perceptions as the other. Thus, we do not have a quasi-perception of the other's visual, auditory, etc., perception; we have the illusion of perceiving like the other does.

In this respect, EVR offers a powerful tool to experience what the other is going through. The user ends up having the impression of putting herself in the other's shoes and generates in this way a body swap. Yet, is this enough to argue that EVR is helpful to train empathy? As I shall argue, we need an accurate analysis of how perspective-taking plays a role in empathy.

10.5 Sensual and Emotional Empathy: Perspective-Taking and Levels of Empathy in Phenomenology

Having described the phenomenon of perspective-shifting in the previous section, in this section, I introduce and elaborate on a distinction between two levels of empathy found in classical and contemporary phenomenology, which Stein (1989) and Svenaeus (2018) refer to as sensual and emotional empathy. This distinction will be useful for understanding empathy in EVR. Here, my aim is to argue that the phenomenon we call "perspective-taking," which is central to putting ourselves in the other's shoes, can take place at two different levels.

10.5.1 Sensual Empathy

The term "sensual empathy" (Empfindungseinfühlung) (Stein 1989, 65) is introduced by Stein to describe empathy with the other's lived body. With the concept of sensual empathy, she describes the possibility of empathizing with the sensual experience of another living being, something that has been often overlooked in the research. Indeed, the debate on empathy has been mainly focused on the possibility of empathizing with the

other's affective states.[14] Yet, for Stein, it is possible for a form of empathy that targets the other's lived bodily experiences such as perceptions and sensations.

In phenomenology, the lived body describes how the body is given in the first-person perspective. As such, it is distinguished from the physical body, which is the body given in the third-person perspective. While the lived body indicates the subjective experience of one's own body and its surroundings independently of sensory perception, the physical body is the body given through external perception by means of seeing, hearing, touching, etc. While the lived body is a particular form of consciousness of our body and its surroundings (Scheler 1973) which is not measurable, the physical body can be measured and is observable.

According to Stein, the other is given to us as a lived body with her fields of sensation located at a zero point of orientation in her spatial world, a field of expression of the experiences of the "I" and an instrument of the will (Stein 1989, 57). In her view, in perceiving the fields of sensation of the other, we are able to grasp the implicit tendencies in her movements. In so doing, we can apprehend what the other is going through at the sensory level. It is precisely because the other is given to us as a lived body like ours that sensual empathy is possible. In Stein, we can modify the real properties of our physical body in the imagination. In this regard, Svenaeus (2018, 748) writes:

> Sensual empathy is a process of recognition and understanding that takes place on the level of embodied existence when one lived body feels and perceives the presence of another lived body and follows its experiences through in a spontaneous manner.

Sensual empathy in this sense is not to be conflated with motor empathy. Motor empathy consists in unconsciously and automatically adopting the other's expressions. By contrast, sensual empathy is a much more complex process of projecting ourselves in the other's lived body while maintaining the distinction between self and other.

In Stein's model, we are able to empathize not just with the other's field of sensations, but also with her position as a zero point of orientation and her will. Empathizing with the other's body as a living body placed in the spatial outer world presupposes that we "transfer" ourselves, i.e., put ourselves in the other's zero point of orientation. In so doing, the empathizer gets a new image of the spatial world which corresponds to the other's position within it and becomes a sense of agency.

In Stein's account, this transfer and the recreation of what the other is going through is not merely an imaginative exercise. The new

orientation and the image of the world gained through it are not merely the product of imaginings. In fact, as she puts it, it is "con-primordial, because the living body to which it refers is perceived as a physical body at the same time and because it is given primordially to the other 'I', even though non-primordially to me" (Stein 1989, 61–62). As a result, once the empathizer has projected herself into the other's perspective, she does not imagine how it is the world for the other, but she is able to live it from the other's point of view. In this respect, there is a significant difference between Stein and proponents of simulationism for whom empathy enables us to experientially imagine how the world looks from the other's perspective. As Stein puts it: "The world I glimpse empathically is an existing world, posited as having being like the world primordially perceived" (1989, 63–64). Thus, the perspective of the other's world obtained through empathy is not just a modification of one's own perspective by means of imaginings. What we glimpse is the other's real world.

Sensual empathy can take place in different degrees of accuracy. Human beings are able to empathize better with other human beings by virtue of having a human body than with non-human ones: "the further I deviate from the type 'human being' the smaller does the number of possibilities of fulfillment become" (Stein 1989, 59). Attempts to empathize with the other's human hand will probably be more accurate than our attempts to empathize with a dog's paw. Given that we have a human hand ourselves, we are familiar with the kind of sensations the other is experiencing in the first-person perspective. By contrast, though I can imagine how it might feel to have a dog's paw, my empathy will be less accurate because we do not have paws ourselves. In fact, for Stein, certain movements and positions are given to me as "empty presentations without the possibility of fulfillment" (1989, 59).

In my view, for sensual empathy to take place with accuracy, perspective-taking plays a crucial role. As mentioned above, the direct perception of the other's experience might remain quite vague, so that for an accurate and comprehensive understanding of what the other is going through, the empathizer has to put herself in the other's place. This is certainly true if we adopt Stein's model of empathy as a process insofar as the second stage involves what I call here perspective-taking. But it is also true in the case of the pure simulationist for whom empathy can be explained via simulation and does not involve a first-level perception of the other's experience.

10.5.2 Emotional Empathy

Stein uses the term "emotional empathy" to describe empathy with the other's affective experiences. Emotional empathy is what contemporary

research usually refers to as empathy. While the question of whether or not we can empathize with the other's cognitive states such as perceptions and beliefs is controversial, philosophers of empathy usually think of it in terms of targeting the other's affective states.

An important aspect of Stein's account is that, for her, affective states are linked to values. This aspect of her work is also a cornerstone of today's philosophy of emotion. Indeed, it has been widely argued that affective states do not present their object as being neutral but as embodying an evaluative property. For instance, in fear, the feared target is presented as dangerous, in disgust as disgusting, and so on. Though the particular form of this connection is a matter of dispute among contemporary philosophers (e.g., Massin 2023), the idea that emotions are connected with values is widely accepted.

Yet, Stein not only acknowledges a link between affective states and values; she also offers a nuanced picture of how different types of affective states relate to different types of values. In particular, while general feelings (e.g., tiredness) and moods (e.g., cheerfulness) are able to tincture all the mental states of a person and in so doing present the world under a certain evaluative light, emotions (e.g., joy) are directed toward epistemic, moral, and aesthetic values, and sentiments (e.g., love) to personal values.

It is by virtue of this connection between affective states and values that, according to Stein, the other is presented in this kind of empathy as a "spiritual being." In empathizing with the other's affective states, we are able to discover the other's world of values:

> Similarly, in every literal act of empathy, i.e., in every comprehension of an act of feeling, we have already penetrated into the realm of the spirit. For, as physical nature is constituted in perceptual acts, so a new object realm is constituted in feeling. This is the world of values (Stein 1989, 92)

As a result, in empathy, we apprehend not only the other's affective state but also how the world is evaluatively presented to her. In this respect, emotional empathy makes us accessible the other's world of values, and this is a dimension of the human being which was not accessible by means of sensual empathy limited to the other's lived body. Emotional empathy can be based on sensual empathy so that here we are dealing with different levels of the empathic experience. Though we can be reluctant to characterize the other with whom we empathize as a spiritual being and to speak of entering into the realm of the sprit, Stein's idea that it is by virtue of a link between affective states and values that we come to know how the world is presented to the other as having particular evaluative properties remains valid in current research.

For emotional empathy, it is my view that perspective-taking plays a crucial role. This is the case for Stein, too, given that she understands empathy as a process. For her, it can be the case that we vaguely perceive an affective state in the other's expressions, and that in order to make this apprehension of the other's experience, perspective-taking is necessary. For instance, we can see that the other is suffering. This might give us an initial but vague grasp that for this person the world is presented under a negative light. Yet, if we are able to adopt the other's perspective and imagine how it is for the other to be in her particular situation, this might offer us a more nuanced view of what the other is going through. For Stein, what I call here an imagining enables the empathizer to really grasp how the world looks for the other. We might then realize that the other is suffering because she is anxious about her future and therefore that the world appears to her as a menacing and dangerous place. In so doing, we can come to grasp the evaluative dimension of the other's experience in more detail. Yet, the role of perspective-taking for emotional empathy remains central, even for the pure simulationist, which does not require a stage of perception in order to have an empathic experience. Indeed, for the pure simulationist, it is by means of perspective-taking that we come to imagine the evaluative properties with which the world is presented to the other.

To sum up, and to adopt more contemporary terms, the distinction between sensual and emotional empathy concerns the *object* of empathy. In sensual empathy, the object is the lived body as a field of sensations and placed at a zero point of orientation. In emotional empathy, the object is the other as a spiritual being with its affective states and the corresponding world of values. In my view, at both levels of empathy, perspective-taking performs a crucial role, since otherwise the apprehension of the other's experience would likely remain rather vague. Moreover, both levels of empathy remain in a particular and functional order. Sensual empathy might serve as a basis for emotional empathy (Svenaeus 2018). Importantly for my account is that what we call perspective-taking can in fact take place at two different levels: there is a perspective-taking at the sensory level and a perspective-taking at the emotional level. In each of these levels, perspective-taking presupposes perspective-building, projection, and resonance.

Stein's distinction has been reintroduced into current research by Svenaeus to explore empathy in the medical context. My aim in this chapter is to show that this distinction can also be useful in understanding empathy in virtual contexts. In fact, with the exception of Svenaeus's work, the distinction is absent from current research in which the idea of sensual empathy has been scarcely examined. In addition, the phenomenological distinction between levels of empathy fills a lacuna in contemporary research insofar as the differences it captures cannot be explained in terms of other distinctions that are in current usage. On the one hand, the distinction

should not be conflated with the distinction between cognitive and affective empathy (see Maibom 2017b, 1). The term cognitive empathy refers to mere perspective-shifting or mindreading, while the term affective empathy presupposes that the empathizer comes to feel something similar to what the other is going through. In contrast, the terms sensual and emotional empathy refer to two different objects of the empathic experience. On the other hand, the distinction between sensual and emotional empathy should not be conflated with the distinction between low-level and enactive (or high-level) empathy. Low-level empathy is an automatic response to the other's bodily expression, while enactive (or high-level) empathy entails imagination-like states and perspective-shifting. Yet, as explained above, sensual empathy entails perspective-shifting and as such imagining.

10.6 From Immersive Body Swapping to Apprehending the Other's Emotions in EVR

Having argued that the phenomenological distinction between two levels of empathy entails distinguishing between two different forms of perspective-taking, let's go back to the initial question of this chapter regarding the role of perspective-taking to explain how EVR might help to foster empathy. My claim is that EVR enables sensual empathy but not necessarily emotional empathy. Yet, given that emotional empathy is often based on sensual empathy, when complemented with other tools, EVR can be used to promote emotional empathy as well.

As we have seen in section 10.4, in EVR, perspective-taking is limited to a particular or a set of particular perceptual experiences of an avatar. By means of different devices, the user comes to perceive a particular aspect of the other's reality. These perceptions generate in the empathizer a series of perceptual illusions of presence, embodiment, and agency. This enables her to adopt the other's perspective "from the inside" in a very precise and realistic way. These sensations can generate the impression in the user that there is a real body swap with the avatar.

However, other mental states of the avatar, such as her emotions, thoughts, beliefs, memories, etc., are not made available for these different technical devices. In fact, the scope of the other's experiences made available to the user is limited to the perceptual experiences transmitted, thanks to the oculus rift, headphones, etc. Therefore, non-perceptual states need to be presented to the user by other means. For this to occur, the user has to imagine the elements of the other's experience which are not accessible via the aforementioned techniques. Once imagined, the user can then adopt the other's perspective "from the inside" and imaginatively recreate what the other is going through. The empathizer can do so in a self- and in

an other-oriented perspective-taking. For the latter to be the case and for empathy to be possible, the empathizer has to gather as much information as possible about the other. However, unlike what occurs with the perceptual experiences made immediately available to the user, she does not have the impression of really undergoing these non-perceptual experiences.

According to these results, perspective-taking in EVR makes available to the user the avatar's perceptions in an immediate and accurate way. EVR generates the impression of a body swap enabling the empathizer to put herself in the other's shoes. Yet, perspective-taking is limited to those of the avatar's perceptions transmitted by different technical devices. By contrast, non-perceptual mental states cannot be made available by the technical devices used in EVR. The user has to imagine in order to build the avatar's perspective, project herself into it, and resonate with the other's experience. How to explain the role of perspective-taking for empathy in EVR given these results?

To begin with, according to the description of perspective-shifting in EVR elaborated above, the techniques employed foster sensual but not necessarily emotional empathy. There are different arguments in favor of this claim. First, as we have seen, EVR enables its users to adopt the perceptual experiences of another being and, in so doing, it generates perceptual illusions of presence, embodiment, and agency, in which the user has the impression of perceiving the world from the other's perspective. Drawing on the phenomenological tradition, I characterized sensual empathy as empathy with the sensual experiences of another being. I underscored that this involves adopting the other's sensations, her point of orientation in the world, and her will. Clearly, the kind of perspective-taking that takes place in EVR fits into the model of sensual empathy.

Second, focusing on the second illusion, authors such as Maselli and Slater (2013) describe the illusion of body ownership in terms of healthy subjects experiencing the other's body as if it were "their own physical body," and they argue that this illusion is particularly interesting for the study of self-consciousness as it relies on an altered representation of the entire body. We can interpret this illusion in phenomenological terms and claim that the illusion does not concern the "physical body" but what I called above the "lived body," i.e., the consciousness of one's own body and its surroundings. In my view, this consciousness would affect those perceptions made available to the user as well as the sense of agency. Indeed, the empathizer has the impression of feeling into the other's lived body and perceiving aspects of reality as the other does.

This speaks in support of the claim that perspective-taking in EVR might lead to sensual empathy. However, while in phenomenology sensual empathy occurs via a direct perception of the other's lived body, in EVR it takes place by experiencing ourselves the other's perceptions. In other

words, in EVR, the user has the illusion of experiencing herself the other's lived body. By contrast, the kind of perspective-taking enabled by EVR does not make the other's emotions, and thereby her values, directly accessible. In EVR, non-perceptual aspects of the other's mind remain hidden to the user. Thus, to access the other's emotions, the empathizer has to resort to imaginings of different kinds. In brief, the kind of empathy enabled by EVR consists in swapping bodies rather than minds.

That said, emotional empathy can be fostered in EVR on the basis of sensual empathy. For instance, after having an experience of body swap with the other, a user can come to imagine how it is for the other to experience a certain situation. Yet, for this to occur, strategies other than the techniques mentioned above should be involved. Here, for instance, narrative devices, such as stories or short narrations about what the other is going through, play an important role. Otherwise, the other's emotional life remains unavailable to the user. Thus, for EVR to foster emotional and not merely sensual empathy, it has to be supplemented with the usage of narratives.

In my view, "The Machine To Be Another" (TMBA) is a good example of how perspective-taking in EVR can be employed to foster sensual as well as emotional empathy. TMBA is an example of artistic work outside the laboratory which combines different techniques and which, in my view, has the potential to foster empathy at the two levels mentioned above. The machine has been used in a different context to reduce negative stereotyping. The use of an oculus rift and headphones enables a body swap and can present real narratives from individuals acting as performers. The machine has been described as an "Embodied system" because there is no virtual world or 3D model but rather a body swap (Sutherland 2016). The machine consists of an EVR System that facilitates body transfer.[15] TMBA also resorts to the use of narratives to make empathy possible. Combining VR with performance art, the machine enables us to perceive ourselves in a different body, fostering active perspective-taking and empathic concern. The use of technical devices and narratives means, in my view, that the machine is capable of promoting forms of perspective-taking which might lead to sensual and emotional empathy.

To sum up, when contemporary researchers claim that EVR can be used to train empathy, in fact, they should be aware that the kind of empathy at work here is sensual but not emotional. Indeed, EVR facilitates body swap and, in so doing, it enables sensual empathy, but EVR alone cannot foster emotional empathy. Yet, since sensual empathy can be the basis for emotional empathy, EVR can potentially play a role in fostering emotional empathy as well. However, for this to occur, EVR must be supplemented with narratives.

10.7 Concluding Remarks

This chapter has examined the role of perspective-taking for empathy in EVR. My aim has been to show how the contemporary debate among natural scientists can benefit from the phenomenological distinction between levels of empathy. After elaborating an account of perspective-shifting in non-virtual as well as in virtual environments, I introduced the phenomenological distinction between sensual and emotional empathy. I argued that this distinction enables us to differentiate between two different forms of perspective-taking. I argued that EVR can be used to train sensual empathy, but that for fostering emotional empathy it is necessary to introduce narrative devices, which make the other's emotions and values accessible to the user. These results offer a contra-argument to those who are skeptical of the link between EVR and empathy. Indeed, as we have seen, for Sora Domenjó (2022), VR does not increase our understanding of the other's feelings. Yet, while supporting this view, this chapter has argued that there is a link between EVR and sensual empathy, and that, on the basis of sensual empathy and with the help of narratives, emotional empathy could be engendered. More generally, these results show that to understand the link between EVR and empathy, the introduction of phenomenological concepts can help to achieve a nuanced view on how perspective-taking, a core element of complex forms of empathy, works in EVR.

Acknowledgments

Work on this project was funded by the DFG Project "Mental Images and Imagination" (Project no. 417829707).

Notes

1 There are different forms of Virtual Reality (VR) (see Bertrand et al. 2018). In non-immersive VR, tridimensional environments created by computer-generated imagery use two-dimensional visual interfaces, such as computer screens and projectors. In immersive VR, tridimensional environments use immersive interfaces, such as VR glasses, and immersive projections, such as a CAVE system. As observed by Bertrand et al. (2018), some of these technologies are more immersive than others.
2 Other issues to be analyzed (some of them already mentioned by Sora-Domenjó) are how empathy promotes pro-social behavior and whether its effects are short- or long-term.
3 For a discussion, see: Feagin (1996); Gibson (2016); Stueber (2016).
4 For an exception: Walton (2015). For him, empathy does not require perspective-taking.

5 In this respect, she claims that the first and the third stages resemble perception, while the second stage, which involves perspective-taking, in my view is based on imagining.
6 I employ the term "re-presenting" as used by phenomenologists to refer to imaginings in which the object is presented in image and not in person. It translates the German term "Vergegenwärtigung." See: Cavallaro (2017).
7 Differences in the scope can be reflected, too, at the phenomenal level. See: Werner (2023).
8 For the view that imagining from the inside involves perspective-taking, see: Williams (1973).
9 This idea that imagining-experiencing generates experience-like states can be found in: Arcangeli (2020). Note that I adopt a quite liberal view here according to which we are able to generate imaginative counterparts to all our mental states and the character's ways of engaging with the world.
10 It is a matter of controversy if this recreated state is an actual or an imagined state.
11 Here, some authors argue that for this we have to experience the same (Coplan 2011; Feagin 1996), while for others, it suffices that we experience something similar (Stueber 2006).
12 The interrelation of presence and empathy has been analyzed by Schutte and Stilinović (2017).
13 Lenggenhager et al. (2007) designed an experiment using conflicting visual-somatosensory input in virtual reality in order to disrupt the spatial unity between the self and the body. They found that during a multisensory conflict, individuals feel as if the virtual body seen in front of them is, in fact, their own body. This leads them to mislocalize themselves toward the virtual body, i.e., outside their own bodily borders. These authors work with the idea of a bodily self-consciousness in terms of a non-conceptual and pre-reflective processing and representation of body-related information. In cases of multisensory conflict, they found that vision dominates over proprioception and touch.
14 In current research, sensual empathy has not been a focus of concern either in the phenomenological or in the analytical tradition (see: Maibom 2017a; Englander and Ferrarello 2023). Even in Stein's research, there has been a focus on emotional empathy rather than sensual empathy (Vendrell Ferran 2015; Magrì 2018).
15 For a critique of TMBA, see: Sutherland (2016).

References

Arcangeli, Margherita. 2020. "The Two Faces of Mental Imagery." *Philosophy and Phenomenological Research* 101 (2): 304–22.

Banakou, Domna, and Mel Slater. 2014. "Body Ownership Causes Illusory Self-Attribution of Speaking and Influences Subsequent Real Speaking." *Proceedings of the National Academy of Sciences* 111 (49): 17678–83.

Bertrand, Philippe, Jérôme Guegan, Léonore Robieux, Cade Andrew McCall, and Franck Zenasni. 2018. "Learning Empathy Through Virtual Reality: Multiple Strategies for Training Empathy-Related Abilities Using Body Ownership Illusions in Embodied Virtual Reality." *Frontiers in Robotics and AI* 5: 26.

Carruthers, Peter, and Peter K. Smith. 1996. *Theories of Theories of Mind*. Cambridge: Cambridge University Press.
Cavallaro, Marco. 2017. "The Phenomenon of Ego-Splitting in Husserl's Phenomenology of Pure Phantasy." *Journal of the British Society for Phenomenology* 48 (2): 162–77.
Coplan, Amy. 2011. "Understanding Empathy: Its Features and Effects." In *Empathy. Philosophical and Psychological Perspectives*, edited by Amy Coplan and Peter Goldie, 3–18. Oxford: Oxford University Press.
Englander, Magnus, and Susi Ferrarello. 2023. *Empathy and Ethics*. London: Rowman & Littlefield.
Feagin, Susan. 1996. *Reading with Feeling. The Aesthetics of Appreciation*. Ithaca and London: Cornell University Press.
Gibson, John. 2016. "Empathy." In *The Routledge Companion to Philosophy of Literature*, edited by Noël Carroll and John Gibson, 200–19. London: Routledge.
Goldie, Peter. 2011. "Anti-Empathy." In *Empathy: Philosophical and Psychological Perspectives*, edited by Amy Coplan and Peter Goldie, 302–17. Oxford: Oxford University Press.
Goldman, Alvin. 2006. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: Oxford University Press.
Kilteni, Konstantina, Raphaela Groten, and Mel Slater. 2012. "The Sense of Embodiment in Virtual Reality." *Presence* 21: 373–87.
Kind, Amy. 2016. "Introduction – Exploring Imagination." In *The Routledge Handbook of Philosophy of Imagination*, edited by Amy Kind, 1–12. London and New York: Routledge.
Lenggenhager, Bigna, Tej Tadi, Thomas Metzinger, and Olaf Blanke. 2007. "Video Ergo Sum: Manipulating Bodily Self-Consciousness." *Science* 317 (5841): 1096–99.
Lipps, Theodor. 1903. *Leitfaden der Psychologie*. Leipzig: Wilhelm Engelmann.
Magrì, Elisa. 2018. "Stein and the 'Rainbow of Emotions': Empathy and Emotional Experience." *Studia Phenomenologica* 16: 198–212.
Maibom, Heidi. 2017a. *The Routledge Handbook of Philosophy of Empathy*. London and New York: Routledge.
Maibom, Heidi. 2017b. "Introduction to Philosophy of Empathy." In *The Routledge Handbook of Philosophy of Empathy*, edited by Heidi Maibom, 1–9. London and New York: Routledge.
Maister, Lara, Mel Slater, Maria V. Sanchez-Vives, and Manos Tsakiris. 2014. "Changing Bodies Changes Minds: Owning Another Body Affects Social Cognition." *Trends in Cognitive Sciences* 19 (1): 6–12.
Maselli, Antonella, and Mel Slater. 2013. "The Building Blocks of the Full Body Ownership Illusion." *Frontiers in Human Neuroscience* 7: 83.
Massin, Olivier. 2023. "The Reactive Theory of Emotions." *European Journal of Philosophy* 31 (3): 785–802.
Petkova, Valeria I., and H. Henrik Ehrsson. 2008. "If I Were You: Perceptual Illusion of Body Swapping." *PLoS ONE* 3 (12): e3832.
Sato, Atsushi, and Yasuda Asako. 2005. "Illusion of Sense of Self-Agency: Discrepancy between the Predicted and Actual Sensory Consequences of Actions Modulates the Sense of Self-Agency, but Not the Sense of Self-Ownership." *Cognition* 94 (3): 241–55.
Scheler, Max. 1973. *Formalism in Ethics and Non-Formal Ethics of Values*. Evanston, IL: Northwestern University Press.

Scheler, Max. 2008. *The Nature of Sympathy*. New York: Routledge.
Schutte, Nicola S., and Emma J Stilinović. 2017. "Facilitating Empathy Through Virtual Reality." *Motivation and Emotion* 41: 708–12.
Slater, Mel. 2009. "Place Illusion and Plausibility can Lead to Realistic Behavior in Immersive Virtual Environments." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364: 3549–57.
Sora-Domenjó, Carles. 2022. "Disrupting the "empathy Machine": The Power and Perils of Virtual Reality in Addressing Social Issues." *Frontiers in Psychology* 13: 814565.
Stein, Edith. 1989. *On the Problem of Empathy: The Collected Works of Edith Stein*. Washington: ICS Publications.
Stueber, Karsten. 2006. *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*. Cambridge, MA: MIT Press.
Stueber, Karsten. 2016. "Empathy and the Imagination." In *The Routledge Handbook of Philosophy of Imagination*, edited by Amy Kind, 368–79. London: Routledge.
Sutherland, Ainsley. 2016. "The Limits of Virtual Reality: Debugging the Empathy Machine." *MIT – Docubase*. Available at: https://docubase.mit.edu/lab/case-studies/the-limits-of-virtual-reality-debugging-the-empathy-machine/
Svenaeus, Fredrik. 2018. "Edith Stein's Phenomenology of Sensual and Emotional Empathy." *Phenomenology and the Cognitive Sciences* 17 (4): 741–60.
Vendrell Ferran, Íngrid. 2015. "Empathy, Emotional Sharing and Feelings in Stein's Early Work." *Human Studies* 38 (4): 481–502.
Vendrell Ferran, Íngrid. 2023. "Fictional Empathy, Imagination, and Knowledge of Value." In *Empathy and Ethics*, edited by Magnus Englander and Susi Ferrarello, 375–98. London: Rowman & Littlefield.
Walton, Kendall L. 2015. *In Other Shoes: Music, Metaphor, Empathy, Existence*. Oxford: Oxford University Press.
Werner, Christiana. "'Tell Me, How Does It Feel?' Learning What It Is Like Through Literature." In *Empathy's Role in Understanding Persons, Literature, and Art*, edited by Thomas Petraschka and Christiana Werner, 174–96. New York: Routledge.
Williams, Bernard. 1973. *Problems of the Self*. Cambridge: Cambridge University Press.
Zahavi, Dan. 2011. "Empathy and Direct Social Perception." *Review of Philosophy and Psychology* 2 (3): 541–58.

Index