
Dieter Sturma [Ed.]

Mind and Time

Philosophical and
Psychological Perspectives

ALBER PHILOSOPHIE



Alber Philosophie

Dieter Sturma [Ed.]

Mind and Time

Philosophical and
Psychological Perspectives

ALBER PHILOSOPHIE



The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>

ISBN 978-3-495-99406-1 (Print)
978-3-495-99407-8 (ePDF)

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

ISBN 978-3-495-99406-1 (Print)
978-3-495-99407-8 (ePDF)

Library of Congress Cataloging-in-Publication Data

Sturma, Dieter
Mind and Time
Philosophical and Psychological Perspectives
Dieter Sturma (Ed.)
178 pp.
Includes bibliographic references.

ISBN 978-3-495-99406-1 (Print)
978-3-495-99407-8 (ePDF)



Online Version
Nomos eLibrary

1st Edition 2023

© Verlag Karl Alber within Nomos Verlagsgesellschaft, Baden-Baden, Germany 2023.
Overall responsibility for manufacturing (printing and production) lies with Nomos Verlagsgesellschaft mbH & Co. KG.

This work is subject to copyright. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage or retrieval system, without prior permission in writing from the publishers. Under §54 of the German Copyright Law where copies are made for other than private use a fee is payable to "Verwertungsgesellschaft Wort", Munich.

No responsibility for loss caused to any individual or organization acting on or refraining from action as a result of the material in this publication can be accepted by Nomos or the author.

Table of Contents

<i>Dieter Sturma</i>	
Mind & Time: Interdisciplinary Division of Labor	7
Mind	15
<i>André Beauducel</i>	
Consciousness and Psychometric Modeling	17
<i>Bert Heinrichs</i>	
Artificial Intelligence and Rational Discourse	45
<i>Michael H. Herzog</i>	
The Temporal Microstructure of Conscious Perception . . .	55
Time	71
<i>Charlotte Gauvry</i>	
How to explain the experience of the passage of time? . . .	73
<i>Christoph Horn</i>	
Husserl's Phenomenology of the Inner Time Consciousness: Achievements and Limits	95
<i>Marius Bartmann</i>	
Wittgenstein on the Language of Time and the Solipsism of the Present	109

Table of Contents

Mind & Time 129

Dieter Sturma

The Language of Time and Self-Consciousness 131

Ulrich Ettinger, Kristof Keidel, Carsten Murawski

Temporal Cognition, Impulsivity and Intertemporal Choice 151

Contributors 177

Mind & Time: Interdisciplinary Division of Labor

1. Approaching Mind and Time

In answering the question of what it means to live the life of a person, one cannot help but to refer to the experience of the passage of time as well as to self-consciousness with its deep questions of existence, mortality, and the meaning of life. The experience of time is clearly at the very center of the workings of the mind. We experience the passing of time on very different levels, from brief moments to whole lives, to historical epochs that exceed the lifetime of any person. On closer examination, it appears more and more puzzling what actually connects these different moments and periods of time. The more so as humans have dramatically expanded the temporal horizons throughout the cultural history. The human time horizon leads not least to the problem of the place that human consciousness occupies in the fabric of the world in contrast to everything else that senses and experiences.

The widening of the temporal horizon is the result of technology, which includes above all language—the technique of expression—and formal techniques such as counting and calculation. In particular, drawing a line and its partitioning belongs to the primal scenes of technology. This simple technique of linearity has put us on a path to the modern scientific image of the world and to the measurement of time. The primal scene is the reconstructed process that stands at the beginning of that path that eventually leads to the modern scientific image of the world. Along this path, the expectation has been formed that insights into the fabric and workings of the universe can be gained from the mathematical measurement of space and time. Even today, this expectation has lost none of its relevance.

There are apparently two different ways to approach mind and time. We can either explicitly take the perspective of a person with her subjective point of view or an objective standpoint, which is formed, as it were, by a view from nowhere on the sequence of events.

Perspective-independent descriptions register the succession of events without reference to the passage of time. The objective description of the centerless view lacks the experiential presence, which has such an overriding importance in conscious life. It therefore comes as no surprise that the terms »mind« and »time« have been associated with so-called hard problems again and again—this applies both in semantic and explanatory respects as well as in systematic respects.

When we deal with manifestations of the human mind, we soon stumble upon temporal aspects of perceiving, knowing, behaving, deciding, and acting. These mental acts stand in a peculiar connection with the course of events. Therefore, the question arises whether time is something that directly depends on these processes or whether it runs independently of them. This question has often been answered by saying that time is ultimately an illusion, or at best that it can be understood as a certain perspective on the course of events.¹ In any case, the modern philosophy of time takes its starting point from the thesis of the unreality of time as it has been developed, for instance by John Ellis McTaggart.

In the context of the development of his naturalistic system, Spinoza pointed out that time is not a determination of things, but only a way to think about them that enables us to explain duration.² The metaphor of the passing of time, which we so frequently encounter in everyday life, misleads us about a puzzling phenomenon, especially since there is no direct evidence for the experience of time as such. Rather, it seems that the subjective experience of time depends on the propositional content of the respective perceptions of the course of events. Temporal processes are accordingly inseparable from spatial processes.

Nevertheless, it would be hasty to regard time as not real or as a mere illusion. A person who expresses the words »here« and »now« or conceives them in the mind, makes her respective position in the world explicit and takes by this mental act a new attitude to her reality with new options for action. A person who masters the language of time is capable of perceiving and changing the course of events in a specific way. With persons, a perspective appears in the world, which joins the

¹ Cf. Rovelli 2018, 40: »The idea that a well-defined now exists throughout the universe is an illusion, an illegitimate extrapolation of our own experience.«

² Cf. Spinoza 1972, 244: »Quare tempus non est affectio rerum, sed tantum merus modus cogitandi [...] durationi explicandae inserviens.«

sequences of »before,« »simultaneous,« and »after,«—McTaggart's B-series—but consistently occupies a new present with a new past and future—McTaggart's A-series.³ With the terms »here« and »now« a unique constellation manifests itself in the B-series with elements of the A-series.⁴ This complicated relationship is likely to contribute significantly to the impression of the passage of time.

For us, the passage of time always points to the future, of which we have no knowledge, but about which we can make plausible conjectures. The plausibility of these conjectures depends to a not inconsiderable degree on the information we have about relevant events of the past and present. These events are determined by causal relations which follow the B-series. An event is never followed by its cause. Our understanding of time accompanies the direction of causality.⁵ At least in this respect, time is not a mere illusion, but an expression of real processes.

2. What is Time?

The consciousness of the human life-form is characterized by its own way of dealing with temporal processes. It is often assumed that the analysis of the experience of time is the key to the understanding of human consciousness. But both fields of investigation are still dominated by profound controversies, and it looks as if the difficulties would be doubled if they were brought together.

Physicalism as well as most naturalistic approaches deny the possibility of speaking of a passing of time. The physicalist and naturalist positions are confronted with the phenomenological view, which understands the passage of time as the core of the experiential reality of persons. This opposition still characterizes current research. Yet, in recent debates, conceptions are emerging in which the physical

³ Cf. McTaggart 1927, 9–31.

⁴ Cf. Prosser, 3: »There is [...] an ever-changing succession of A-series, always consisting of the same series of physical events but with a different time being present on each series, and with other time's A-series positions differing accordingly.«

⁵ Cf. Prosser 204: »Earlier events cause later ones, but not vice versa. [...] A person's brain, at any given time, contains informational traces of the past (in memory), but not the future (the subject can, of course, have true beliefs about the future, but these are not the result of any causal influence by future events).«

view and the phenomenological view no longer appear to be mutually exclusive.⁶

What we seem to know is that we can measure time with a variety of techniques. These techniques, however, say little about the experiential awareness of temporal consciousness as distinct from the experiences of other forms of life. Further, we presuppose that the consciousness of a person is for the most part transparent and self-referential. We can ask persons what they are thinking about or what their thoughts are at this specific moment. But we also imply that not every mental process is explicitly experienced. Not least, we are skeptical that we can measure consciousness as such.⁷

In the history of philosophy, there are early philosophical approaches to dealing with the difference between subjective and objective time. Aristotle and the Stoics bring time into a direct connection with movement and process. This is in contrast with the later view that time emerges from activities of consciousness. This position is associated with Descartes and sometimes with Augustine, but especially with Kant, Brentano, and Husserl. However, this constellation of theories is not as clear as is often suggested.⁸ It is widely accepted that Husserl's phenomenological method marks an innovative turn for the 20th century in the philosophy of the time.⁹

3. The Semantic and Methodical Challenges

In methodological and semantic terms, it hardly escapes the attention that conceptions or ideas of time are referred to using metaphorical expressions, such as »time flows,« »time drags on,« or »time stands still.« These metaphorical expressions, as familiar as they may seem to us, are anything but harmless from a philosophical point of view, and it is of utmost necessity that they be subjected to semantic analysis.¹⁰ Without conceptual analysis and linguistic control, metaphors generate inconsistent or contradictory ideas, especially when they are taken directly as representations of the way things are. Particularly, in the case of expressions like »mind« and »time« linguistic clarifica-

⁶ Cf. Chauvry 2023.

⁷ Cf. Herzog 2023, Beauducel 2023.

⁸ Cf. Horn 2023, 99–102.

⁹ Cf. Horn 2023, 96–99.

¹⁰ Cf. Bartmann 2023.

tions and conceptual analysis are indispensable. Linguistic vagueness is impressively revealed by the supposedly simple question »When is now?«.¹¹ Not only is it impossible to answer it in a sufficiently discriminating way, but it confronts us directly with the fundamental problem of the pros and cons of discrete temporal units.¹²

Empirical studies indicate that the threshold for the distinction of temporal units is measurable. This does not mean that no mental activities take place beyond these boundaries.¹³ We are dealing with different scientific resolutions of time relations. The sequences of objective or physical time run to a large extent below the threshold of subjective awareness.¹⁴

There is no consensus on the use of the concepts »mind« and »time«—neither interdisciplinary nor intradisciplinary. Contrary to what is widely implied, the absence of a consensus need not necessarily be understood as a semantic disaster. It is quite possible that it is a manifestation of an inevitable incoherence, which could result from the phenomenon as such or from limited epistemic access. Furthermore, in the case of the concepts »mind« and »time« there are established practices of language use, which also allow for methodical comparisons and identifications of relationships. In this way, no unified picture emerges, but at least a retraceable linguistic constellation of definitions, connections, discontinuities, and oppositions. In this situation, a method is appropriate that is oriented towards semantic diversity and division of labor—as is indicated for instance in Wittgenstein’s reflections on language games and family resemblance.¹⁵

4. Interdisciplinary Diversity and Division of Labor

The different approaches to »mind« and »time« open a wide interdisciplinary field. In its everyday research, a single discipline does not aspire to cover a wide research field in its entirety—although some public statements seem to suggest otherwise. A single discipline does not have the instruments and methods needed for all aspects of the research area at its disposal. At any rate, the focus is usually on only

¹¹ Cf. Gale 1964.

¹² Cf. Herzog 2023, 56.

¹³ Cf. Herzog 2023, 58–62.

¹⁴ Cf. Herzog 2023.

¹⁵ Cf. Wittgenstein 1958. 5–13, 32, 33.

one approach. No single discipline is able to encompass the entire scope and complexity of the field of investigation of »mind« and »time.«¹⁶

In the thematic connection of »mind« and »time,« we are dealing with a phenomenon which is in each case only accessible in certain aspects. This leads to the challenging interdisciplinary task of elaborating the corresponding relations and connections—such as between neuroscientific findings, psychometric approaches, and investigations in the disciplines of psychology, philosophy of science, epistemology, philosophy of language, and philosophy of mind. This constellation obstructs from the outset dualistic and eliminative positions that are still widespread in theories of mind and time. Interdisciplinary diversity and division of labor can only come about within a worldview that is, in the broadest sense, both naturalistic and open for revision.¹⁷

For example, in investigations of decision-making, results have been worked out in different scientific disciplines, such as philosophy, psychology, neuroscience and economics, which in turn are, or can be made, the subject of interdisciplinary research. In specific decisions, the respective attitude of a person, which she adopts towards immediate consequences or later consequences of her actions, is of critical importance.¹⁸ This means that a person's subjective attitude is the origin of processes that can be identified, measured, and evaluated by various disciplines.

For all the mystery and indeterminacy that surrounds the expressions »mind« and »time,« they always refer to ways in which people deal with reasons and their temporal perspectives. This practice ultimately determines who we are and who we can be. These days, however, one wonders whether our use of reason, mind, and time is unique once and for all or whether biological, social, and technical processes can lead to radical changes.¹⁹

In personal and institutional decision-making, developments of Artificial Intelligence are becoming more and more important and obviously influence future manifestations of the human mind and knowledge.²⁰ Regardless of the expansion of technical automation, persons can deliberately change the course of events with their deci-

¹⁶ Cf. Beauducel 2023, 17–18.

¹⁷ Cf. Beauducel 2023, 19–21, Sturma 2023, 131–132.

¹⁸ Cf. Parfit 1984, ch. 8: Different Attitudes to Time, 149–186.

¹⁹ Cf. Heinrichs 2023, 53–54.

²⁰ Ettinger, Ulrich, Keidel, Kristof & Murawski, Carsten 2023.

sions. But they need to ask about impacts when technical systems, designed to execute processes previously initiated only by a person, change the course of events independently of human control. We must even consider the possibility that the proliferation of automated technical interventions will require us to expand the semantic field of the expressions »mind« and »knowledge.« The way we deal with time and especially with self-consciousness seems to stand in the way of such an expansion. The time has not yet come for us to accept full-heartedly systems of Artificial Intelligence as one of us²¹—as beings we can question about their conception of future and past.

The contributions in this volume explore in an interdisciplinary division of labor the semantic, systematic, and empirical constellations of mind and time by using their respective scientific methods. They dispense with grand theory and limit themselves to tracing those contours of mind and time accessible from their respective methods from which, under favorable circumstances, a larger picture emerges.

References

- Bartmann, Marius (2023): Wittgenstein on Time, in this volume.
 Beauducel, André (2023): Consciousness and Psychometric Modeling, in this volume.
 Ettinger, Ulrich, Keidel, Kristof & Murawski, Carsten (2023): Intertemporal Choice, Impulsivity and Temporal Cognition, in this volume.
 Gale, Richard M. (1964): Is It Now Now? in: *Mind* 73.
 Gauvry, Charlotte (2023): What is it like to experience the passage of time? in this volume.
 Heinrichs, Bert (2023): Artificial Intelligence and Rational Discourse, in this volume.
 Herzog, Michael (2023): Discrete Conscious Perception, in this volume.
 Horn, Christoph (2023): Husserl's Phenomenology of the Inner Time Consciousness: Achievements and Limits, in this volume.
 McTaggart, John McTaggart Ellis (1927): *The Nature of Existence, Volume II*, Cambridge: Cambridge University Press.
 Nagel, Thomas (1986), *The View from Nowhere*, Oxford.
 Parfit, Derek (1984): *Reasons and Persons*, Oxford.
 Prosser, Simon (2016): *Experiencing Time*, Oxford: Oxford University Press.
 Rovelli, Carlo (2018): *The Order of Time*, London: Allen Lane.

²¹ Cf. Heinrichs 2023.

Dieter Sturma

Spinoza, Baruch de (1972): *Renati Des Cartes Principiorum Philosophiae Pars I. & II. / Cogitata Metaphysica*, in: *Opera I*, Heidelberg: Carl Winter Universitätsverlag.

Sturma, Dieter (2023): Self-Consciousness and the Language of Time, in this volume.

Wittgenstein, Ludwig (1958): *Philosophical Investigations*, Oxford: Blackwell.

Mind

Consciousness and Psychometric Modeling

Entering ›consciousness‹ into Google Scholar on Dec. 16th 2020 yielded 4.500.000 hits. There is a plethora of researches, paradigms, and results related to consciousness. One might conclude that this is an especially productive and innovative area of research. However, the wealth of issues, results, and paradigms on specific conscious processes as well as on general issues of consciousness may also indicate that some scientific disciplines (e.g., neuroscience, theology, philosophy, psychology) are partly unable to cope with the broadness, heterogeneity, and complexities implied by the topics related to ›consciousness‹. It seems that research on consciousness is still quite a challenge. In the following, a transdisciplinary perspective on consciousness will be tried out. The transdisciplinary perspective is based on the analysis of a possible parallelism between (1) some neurocognitive results on consciousness, (2) some philosophical accounts on consciousness, and (3) psychometric modeling and its possible relationship to consciousness. The arguments from each discipline are outlined in a separate section below. The parallelism that will be outlined below is based on the observation that some neurocognitive results indicate (1) that a localization of specific brain regions of consciousness might be difficult, that (2) a philosophical demonstration of consciousness of individual reactions might be impossible, and (3) that consciousness can be distributed on ›true‹ and ›error‹ components of measurement so that psychometric measurement cannot unambiguously determine the latent variables that are the basis for (conscious) behavior. Thus, the parallelism implies that consciousness occurs at the borders of neurocognitive science, philosophy, and psychometrics. If consciousness is placed at the borders of the disciplines, one may ask why humans use the term ›consciousness‹ in order to provide statements on a scientifically rather intangible phenomenon. A tentative answer to this question is presented in the final part of the third, psychometric section: In a simulation study, a model

whose number of parameters was larger than the number of measured variables was more robust against minor and irrelevant changes in the data. The phenomenon of a model having more parameters than measured variables is termed ›indeterminacy‹ in the psychometric literature. It was proposed that there might be an adaptive advantage for organisms to be regulated by a system that can be described as an indeterminate model that is robust against minor data changes because a robust model may enhance the continuity of behavior across time. If consciousness is an aspect of such a model, an organism equipped with consciousness may have an advantage in producing substantial continuity of behavior. Moreover, one may use the term ›consciousness‹ in order to describe a system providing behavior that is consistent across time but that can nevertheless not be unambiguously determined from available data. This is, of course, a tentative interpretation of the outlined parallelism of neurocognitive, philosophical, and psychometric perspectives. Certainly, other transdisciplinary perspectives on consciousness are possible. Nevertheless, in order to provide a basis for further research, the details of a perspective on consciousness being related to complex neuronal networks, being nearly intangible in the philosophical sense, and being psychometrically indeterminate will be presented in the following.

1. Some Neurocognitive Results on Consciousness

One strategy to cope with the complexities of research on consciousness might be to investigate manifestations and neural correlates of consciousness to manipulate conditions of conscious awareness, of conscious stimulus processing, or to use the natural variation of consciousness that occurs with sleep or brain injuries.¹ By means of this strategy investigators emphasized testable hypotheses on conscious processes. Testable hypotheses on conscious processing have regularly been investigated by means of different neurocognitive methods. For example, event-related potentials (ERP) of the electro-encephalogram have been investigated because of their high temporal resolution and functional magnetic resonance imaging (fMRI) has been per-

¹ Sohn 2019.

formed because of its high spatial resolution.² Several experimental paradigms in the context of cognitive neuroscience and psychology have been used for the investigation of conscious stimulus processing. For example, the attentional blink paradigm refers to temporary impairment of attention occurring when two or more target stimuli have to be processed in very close temporal proximity.³ Target stimuli are those stimuli that ask participants to react when they occur. When there is temporal proximity of two target stimuli, the second target stimulus is not always detected, which might be regarded as an ›attentional blink‹ or as a temporary impairment of conscious perception. It has been found that the accuracy in detecting two successive targets is positively correlated with intelligence.⁴ The detection of two successive targets implies that the attention of participants does not ›blink‹ during the period of successive target presentation. This phenomenon has been named as (suppression of) attentional blink. The attentional blink has been addressed in ERP studies in order to investigate at which processing stage the impairment of attention occurs.⁵ There are several ERP studies based on manipulations of conditions for stimulus detection that have been related to conscious awareness. Many studies investigate the amplitude of the P300 component, i.e., the size of a positive deflection occurring about 300 ms after a stimulus⁶ as an indicator of conscious perception even when earlier ERPs could also mark conscious perception.⁷ Although this is an example for a neurocognitive correlate of consciousness, results based on stimulus detection paradigms are typically restricted to conscious stimulus detection and do not refer to the concept of consciousness as a whole.

Although the use of experimental paradigms in cognitive neuroscience allows to measure conscious processes, the results of several experimental studies depend on the specific design and should perhaps not be generalized to consciousness as a whole. An example for the problem of specificity could be the debate on whether neural correlates of consciousness primarily activate frontal or more parietal

² Goense, Bohraus & Logothetis 2016.

³ Shapiro, Arnell & Raymond 1997.

⁴ Klein, Arend, Beauducel & Shapiro 2011.

⁵ E.g. Kranczioch, Debener & Engel 2003.

⁶ E.g. Kranczioch et al. 2003 and Lamy, Salti & Bar-Haim 2008.

⁷ Rutiku, Martin, Bachmann & Aru 2015.

parts of the cerebral cortex.⁸ When conscious perception is measured, regions in more parietal parts of the cortex were shown to be activated. Koch et al. summarized: »... we describe recent findings showing that the anatomical neural correlates of consciousness are primarily localized to a posterior cortical hot zone that includes sensory areas, rather than to a fronto-parietal network involved in task monitoring and reporting.«⁹ In contrast, when conscious motor reactions or active responses are measured, the prefrontal cortex turns out to be important. Accordingly, Odegaard et al. conclude: »The literature highlights [Prefrontal Cortex] PFC's essential role in enabling the subjective experience in perception, contra the objective capacity to perform visual tasks; conflating the two can also be a source of confusion.«¹⁰ Thus, experiments inducing responses resulting from consciousness emphasize frontal regions and experiments inducing conscious perception are more related to activity in parietal regions of the cortex. This is not surprising as motor responses are more fronto-centrally located whereas sensory perceptions are more parietally located in the brain.¹¹ This demonstrates that the frontal or parietal localization of consciousness depends on the experimental paradigms. However, some neurocognitive results do not relate consciousness to specific brain areas. For example, Agnati et al. proposed mosaic networks made of hierarchically organized functional modules as a basis of consciousness. The mosaic networks allow for the realization of a wide spectrum of different neuronal assemblies, leading to the emergence of neural correlates of consciousness. Agnati et al. relate consciousness to very basic processes of neuronal activation that are widely distributed in the brain.¹² Demertzi et al. report results indicating that broad and complex neural networks can be related to consciousness: »We identified a pattern of positive and negative long-distance coordination, high modularity, with low similarity to the anatomical connectivity, potentially relevant for the support of conscious cognition (pattern 1). We also identified a pattern of low interregional dynamic coordination, low efficiency, with high similarity to anatomical connectivity, potentially specific to reduced or absent conscious

⁸ Koch et al. 2016 and Odegaard et al. 2017.

⁹ Koch et al. 2016, 307.

¹⁰ Odegaard et al. 2017, 9593.

¹¹ Gray & McNaughton 2000.

¹² Agnati et al. 2012.

processing...«.¹³ These and other results on a global neuronal workspace¹⁴ indicate that consciousness is related to a complex coordination of processes that is not necessarily limited to a specific brain region. The idea that a complex long-distance coordination might be essential for consciousness might fit with the philosophical perspectives on consciousness summarized in the next paragraph.

2. Some Philosophical Views on Consciousness

Specific neuronal regions that are related to specific conscious processes may be compatible with a philosophical mind-brain dualism in that they allow for the identification of functional-anatomic substrates underlying specific mental processes. In contrast, neurocognitive results indicating the relevance of complex distributed neuronal networks as a basis for consciousness imply that large parts of the brain are the relevant functional-anatomic substrate of consciousness. The hypothesis that large parts of the brain are the basis for consciousness is certainly not a ›bold hypothesis‹ in the sense of Popper.¹⁵ In order to avoid trivial predictions, it might be helpful to combine the results indicating the relevance of large neuronal networks with philosophical positions that overcome the mind-brain dualism as well as neuroanatomical reductionism. According to Hagberg Wittgenstein provides a philosophical perspective on consciousness that overcomes the mind-brain dualism and the neuroanatomical reductionism.¹⁶ It is therefore interesting to relate the neurocognitive perspective on consciousness as resulting from a complex coordination pattern with the philosophical perspective of Wittgenstein: »But isn't it our meaning it that gives sense to the sentence? (And here, of course, belongs the fact that one cannot mean a senseless series of words.) And ›meaning it‹ is something in the sphere of the mind. But it is also something private! It is the intangible something; only comparable to consciousness itself.«¹⁷ From some neurocognitive models, consciousness is a complex pattern and from Wittgenstein's philo-

¹³ Demertzi et al. 2019, 3–4.

¹⁴ Dehaene, Charles, King & Marti 2014.

¹⁵ Popper 1979.

¹⁶ Hagberg 2018.

¹⁷ Wittgenstein 1958a, 358.

sophical perspective, it is an ›intangible something‹. Lehrer's more recent philosophical approach on consciousness may also fit to these ideas as he relates consciousness to the freedom of representation.¹⁸ Lehrer assumes that conscious states provide the individual with the freedom of representation of sensory material. Accordingly, consciousness allows for a reconfiguration of sensory appearances. Lehrer wrote: »Such reconfiguration reflects our plasticity and autonomy in how we represent the external world, including the world of science. Attention to sensory detail in scientific photography enables us to distinguish artefacts of the process from features of the object photographed, for example. Exemplar representation converts input into a represented term of representation. Once the conversion takes place, we note our freedom, our autonomy, in how we represent the world and ourselves in terms of sensory materials.«¹⁹ If consciousness is necessary for the reconfiguration of sensory representations of the world, it might also be possible that some representations are processed without consciousness, i.e., without the necessity of a reconfiguration. This can again be related to Wittgenstein: »And I want to give you the following rule of thumb: If you are puzzled about the nature of thought, belief, knowledge, and the like, substitute for the thought the expression of the thought, ... This, of course, doesn't mean that we have shown that peculiar acts of consciousness do not accompany the expressions of our thoughts! Only we no longer say that they must accompany them.«²⁰ Thus, the perspective of consciousness as a system used in order to obtain the freedom of reconfiguration for sensory material would perfectly match with the idea that this system must not always be in action. In this sense, consciousness must not accompany the expressions of our thoughts. To sum up, Lehrer's idea that consciousness may facilitate the reconfiguration of sensory material fits to Wittgenstein's idea of consciousness as an ›intangible something‹ that can accompany the expressions of our thoughts.

Jacquette's dynamic attribution model of consciousness might also be related to this perspective.²¹ Jacquette describes consciousness as follows: »Consciousness is the brain's unconscious (autonomic)

¹⁸ Lehrer 2018.

¹⁹ Lehrer 2018, 105.

²⁰ Wittgenstein 1958b, 42.

²¹ Jacquette 2018.

dynamic attribution of cognitive, including perceptual and affective data as properties to passing moments of objective mind-independent real time.«²² Starting from this idea it may be possible to merge Lehrer's idea of consciousness that allows for a reconfiguration of sensory material with Jacquette's model of dynamic attribution of data to passing moments of time. Figure 1 represents either a cup or two faces depending on how information is integrated. If consciousness interprets the perceptual data presented in Figure 1 as a cup in one moment of time, it might be possible that consciousness interprets these data as two faces in another moment of time. Thus, the idea of time attribution may allow for the separation of moments during a constantly given configuration of sensory input. In this case attribution of time moments as proposed by Jacquette would allow to explain the reconfiguration of sensory material proposed by Lehrer by means of a second moment of sensory attribution. Moreover, a perceived change of the stimulus meaning (i.e., seeing either a cup or two faces) and knowing that one has seen two different things based on the same overall sensory material should be conceived as a result of a conscious process. So, if an individual indicates that she/he has seen two different objects in one and the same picture, one may expect that a higher form of conscious processing has occurred.



Figure 1. A cup or two faces.

Although such processes are impossible without some sort of consciousness that might be conform to some philosophical or neurocognitive model, any measurement of consciousness has to face the problem that we need a verbal indication of an individual that conscious

²² Jacquette 2018, 261.

processing has occurred. However, even explicit verbal indications may occur without a conscious understanding of the meaning of the indications. Consider, for example, Wittgenstein:

Someone says irrelevantly »That's a tree«. He might say this sentence because he remembers having heard it in a similar situation; or he was suddenly struck by the tree's beauty and the sentence was an exclamation; or he was pronouncing the sentence to himself as a grammatical example; etc., etc. And now I ask him »How did you mean that?« and he replies »It was a piece of information directed at you«. Shouldn't I be at liberty to assume that he doesn't know what he is saying, if he is insane enough to want to give me this information?²³

Being suspicious regarding the presence of consciousness even in a rather complex verbal response, also follows from the Turing test, which is based on the rating whether a written sentence has been produced by a computer or by an individual.²⁴ Even when more complex versions of the Turing test have been discussed meanwhile,²⁵ it is sufficient in the present context to acknowledge that some sentences might be produced automatically by a computer, so that we will not expect them to be the result of human conscious processing. However, the same sentences may also be produced by a human being without knowing what he is saying, i.e., perhaps without consciousness.²⁶ The Turing test underlines that the occurrence of verbal sentences is no guarantee that conscious processing has occurred.

3. Representing Consciousness in Psychometric Models

3.1. Measurement Error and Consciousness

It follows from the previous paragraphs that the measurement of consciousness is quite a challenge. The reasons are: (1) Neurocognitive results indicate that consciousness is related to complex, large neuronal networks;²⁷ (2) verbal expressions that are typically attributed to a conscious individual may appear without conscious processing by

²³ Wittgenstein 1969, 61.

²⁴ Turing 1950.

²⁵ Penco 2012.

²⁶ See the above mentioned example provided by Wittgenstein 1969.

²⁷ Demertzi et al. 2019.

the individual or may be produced by a computer,²⁸ (3) consciousness may be related to the attribution of perceptions to moments of time,²⁹ and (4) consciousness may facilitate the reconfiguration of stimulus material.³⁰ It should be clear from this list of arguments and results that although a large number of interesting results were obtained for specific conscious processes like perception or attention, the measurement of consciousness itself constitutes a major problem. Nevertheless, since more than one century, psychological research is devoted to the prediction of individual differences, i.e., to the identification of stable and transsituational consistent determinants of individual differences of behavior. Determinants of behavior that are stable across time and consistent across situations are often termed personality traits or individual differences of abilities.³¹ It is, however, not clear how consciousness can be related to personality traits or abilities. Individual differences of consciousness are hard to conceive at the subjective level. For example, when individuals respond to questionnaires, we will typically assume that they provide perfectly conscious responses. We will usually not assume that some individuals are more conscious than others when they respond to questionnaires. It is not clear whether more careful responses are more conscious responses because individuals may consciously be careless. This shows that things are rather complex as will be illustrated when we consider the psychometric modelling of presumably conscious responses to questionnaires.

Interestingly, a substantial advancement in the measurement of traits and abilities occurred when Spearman introduced the difference between the common ability factor ›general intelligence‹ g and the specific abilities s .³² Although the terminology was not fully developed at this time, it was already clear for Spearman that behavioral prediction will mainly be based on g and that s may be related to unpredictable aspects of behavior. Later, when factor analysis has been fully developed, the difference between common factors as determinants of behavior and unique factors comprising specific variance and

²⁸ Wittgenstein 1958 and Turing 1950.

²⁹ Jacquette 2018.

³⁰ Lehrer 2018.

³¹ E.g. intelligence or knowledge, cf. Cattell 1987, Messick 1989 and Stern 1911.

³² Spearman 1904.

measurement error has been more clearly presented.³³ In the common factor model, the measured variables representing individual behavioral responses \mathbf{x} are decomposed by the common factors \mathbf{f} and the unique factors \mathbf{u} .

$$\mathbf{x} = \mathbf{\Lambda}\mathbf{f} + \mathbf{\Psi}\mathbf{u}. \quad (1)$$

The weights of the common factors are called >common factor loadings<. They are given in the loading matrix $\mathbf{\Lambda}$ and the weights of the unique factors are in the matrix $\mathbf{\Psi}$. There are some additional assumptions of the common factor model that are not discussed here.³⁴ It should be noted that \mathbf{u} can be decomposed into the specific variance \mathbf{s} and the measurement error \mathbf{e} . This yields

$$\mathbf{x} = \mathbf{\Lambda}\mathbf{f} + \mathbf{\Psi}(\mathbf{s} + \mathbf{e}). \quad (2)$$

The idea of a measurement error that occurs with every psychological measurement and that is partly represented by the unique factors has also been emphasized in the context of classical test theory.³⁵ Classical test theory can be regarded as a simplification of the common factor model because only a single common factor, the so-called >true score< \mathbf{t} and the so-called >measurement error< \mathbf{e} is assumed.

$$\mathbf{x} = \mathbf{t} + \mathbf{e}. \quad (3)$$

However, the discussions on the measurement error are relevant for the common factor model as well as for classical test theory since \mathbf{e} occurs in Equation 2 and 3. There have been several refinements and specific improvements over classical test theory that have been subsumed under the term item-response theories (Hambleton & Swaminathan, 2013). However, although mathematical and statistical refinements of measurement error are treated in a plethora of papers and books, it is hard to find a comprehensive semantic description or a conceptual, psychologically meaningful definition of measurement error.

³³ Mulaik 2012.

³⁴ Mulaik 2012.

³⁵ Lord & Novick 1968/2008.

However, a few descriptions of the psychological meaning of measurement error are available. For example, when Lord and Novick introduced classical test theory, they started with the description of two final course examinations leading to the different results.³⁶ They explain that it would be a problem when students get different scores and even different orderings of scores for the same course. This leads to the assumption that there is some ›true‹ score and that the difference between the two course examinations is due to measurement error. Referring to Lazarsfeld, they consider that the trait or the ability is constant (in this sense it is regarded as ›true‹) and that some transient state of the person, resulting in differences between the examinations, is random.³⁷ On this basis the differences between the so-called ›true score‹ and the so-called ›error score‹ was introduced in the context of classical test theory.³⁸ There are, however, two possible shortcomings or misunderstandings resulting from this perspective. First, the term ›error score‹ may induce the idea that this term represents something intrinsically wrong, a real noise component. However, it has been shown that what has been termed ›error score‹ only represents the variance that is not focused in aggregation and generalization.³⁹ It has therefore been proposed to replace the term ›true‹ variance by ›wanted‹ or ›intended‹ variance and the term ›error‹ by ›unwanted‹ or ›non-intended‹ variance.⁴⁰ The second issue is that classical test theory and item response theories aim at describing the relationship between wanted and unwanted variance for all kinds of psychological or sociological data. They do not contain specific interpretations or methodological parameters for specific domains of measurement. The example used by Lord and Novick was from the domain of achievement and ability.⁴¹ In their book, the common rank order of scores for the two examinations represent the wanted variance and the differences between the rank order represent the unwanted variance. It is probably quite compelling to regard the unwanted variance as a form of error variance in this specific context

³⁶ Lord & Novick 1968.

³⁷ Lazarsfeld 1959.

³⁸ Lord & Novick 1968.

³⁹ Humphreys 1962, Cronbach, Gleser, Nanda & Rajaratnam 1972 and Wittmann 1988.

⁴⁰ Cattell & Radcliffe 1962 and Beauducel & Leue 2014a.

⁴¹ Lord & Novick 1968.

of achievement test. Test theories are, however, also applied for questionnaire data in the domain of personality research.⁴²

As nearly all areas of psychological research implicitly or explicitly refer to some test theory, it is important to note that ›measurement error‹ or ›non-intended variance‹ can represent several aspects, which may relate to consciousness in different ways. According to Magnusson the ›measurement error‹ term may represent variance from the (1) administration of an instrument, (2) guessing, (3) scoring, and (4) lack of agreement between ›true scores‹ of different measurements.⁴³ The lack of agreement of true scores may be due to fluctuation of the true scores or memory effects. Other relevant aspects, not mentioned by Magnusson, are (5) social desired responding, impression management or response bias,⁴⁴ (6) states, moods, and attention/concentration during responding, and (7) previous experiences with assessment instruments. All of these aspects, with exception of (3) scoring, can be related to some conscious processing. These aspects may therefore not be random from the perspective of the individual although they are treated as if they were random in the test theoretical models. In order to provide examples for this, we will demonstrate the relationship of (4) lack of agreement between ›true scores‹ of different measurements as well as (5) response bias with the measurement of consciousness in more detail.

Lack of Agreement Between ›True Scores‹ of Different Measurements

It is worth to consider possible alterations between two measurements in the domain of questionnaire research more closely. For example, consider the sentence that might be an item of a personality questionnaire: »I can let go myself and enjoy myself at a lively party.« An individual may respond with »I strongly agree«, »I agree«, »I disagree«, and »I strongly disagree«. Let the individual respond with »I agree«. Now, the same item is presented one week later and the individual responds with »I strongly agree«. The concept of measurement error implies that there is a ›true‹ tendency to agree that lies between »I agree« and »I strongly agree« and that the measurement error explains the difference between the two measurement occasions.

⁴² Beauducel & Leue 2014b, Table 1.

⁴³ Magnusson 1967.

⁴⁴ Beauducel & Leue 2014a.

However, what does this imply for research on consciousness? One will usually assume a conscious answering of questionnaires although the mass of items that individuals have to respond to in several contexts may lead to careless responding (perhaps even without reading the sentence). Starting from the assumption that the sentence has been read and understood by the individual which provides a conscious response, measurement error implies that the individual was not completely inclined to respond with »I agree« at the first occasion because the individual was partly inclined to respond with »I strongly agree«. Nevertheless, the individual response was »I agree«, just in order to reach a decision. At the second occasion, the individual is again inclined to respond somewhere between »I agree« and »I strongly agree« and decided to respond with »I strongly agree«. One may figure out more complex settings with an individual that remembers the first response one week before and tries to compensate the first response by means of the second response in order to reach a true averaged response. Alternatively, the individual was sober at the first measurement occasion with some social inhibition leading to the »I agree« response and a reduced social inhibition under the influence of alcohol leading to the »I strongly agree« response at the second measurement occasion.⁴⁵ As a third idea, the individual enjoyed a nice party between the first and the second measurement occasion. Finally, the individual met someone between the measurement occasions, with whom s(he) would like to enjoy a lively party, etc. Thus, the item may represent a trait that determines that the individual tends to agree with the item and it may also represent a state to agree more or less strongly. The combined measurement of traits and states has been acknowledged and represented in Latent-State-Trait-Models.⁴⁶ It should be noted that states may change rapidly when they depend on environmental differences. For example, an individual may have low agreement with the item »I am satisfied with myself« immediately before feedback of an examination result and may have high agreement with this sentence two seconds later, when a positive feedback of the examination-result was given. Then, the difference between the responses of the individual mainly represents the feedback effect. The difference may be regarded as unwanted variance, if one is interested

⁴⁵ See Lazarsfeld 1959 for a similar example.

⁴⁶ Steyer, Schmitt & Eid 1999.

in an overall evaluation of self-satisfaction, but this does not imply that the variance represents an intrinsic error.

The common aspect of these examples is that the conscious response to the items might be altered. The conscious processing of the item (i.e., reading and responding) does not preclude response alterations and we expect that the item response provides us with the current conscious individual appraisal of the item content. This idea can be related to *Jacquette's* notion of consciousness as an attribution of data to passing moments of time. In this sense, the response alterations represented by measurement error can be related to the fidelity of conscious responses to the respective moment of time.⁴⁷ It follows from these considerations that consciousness is at least as present in the response alterations that may be represented by as the unwanted variance of unique factors as in the constant or common factor variance of the responses. Whereas psychological research on personality traits and abilities has focused on the constant part of the variance that is typically represented by the common factors of factor analysis,⁴⁸ the unique factors representing the unwanted variance or measurement error have typically been ignored. This does, however, not preclude that there are domains where especially the unique factors represents essential aspects of conscious processing. Whenever generalization of individual behavioral tendencies is not the focus and when unique responses regarding attitudes, perceptions, cognitions, or emotions need conscious processing, the unique variance or the unique factors may be of special interest for research on consciousness. Unique conscious responses could be those that are based on a specific, individual integration and interpretation of stimuli, and individual response options. To sum up, models of consciousness referring to the attribution of cognitive, perceptual, and affective data as properties to passing moments⁴⁹ as well as the perspective of consciousness as the basis for the reinterpretation of perceptions⁵⁰ underline the uniqueness of conscious experiences. Psychometric models comprising a unique term for each measurement are compatible with this perspective.

⁴⁷ *Jacquette* 2018.

⁴⁸ *Mulaik* 2012.

⁴⁹ *Jacquette* 2018.

⁵⁰ *Lehrer* 2018.

Response Bias

Presuming that items in a questionnaire do not only represent systematic variance measuring the conceptually intended construct (e.g., extraversion) but also a systematic variance that could be named as a ›bias‹ (e.g., the tendency to agree with the item content in purpose of impression management), we should aim at disentangling different types of conscious processing. According to classical test theory, we do not have a variable that corresponds to a systematic measurement component for ›bias‹ (**b**) because it is neither a ›true‹, intended variance component (**t**) nor an unsystematic error (**e**). When individuals answer items as truly as possible with regard to the instruction and with regard to memorizable situations and contexts they possibly add intentions to their answers (e.g., presenting one-self as favorable as possible). These intentions could be added consciously or unconsciously. It is part of factor analysis to disentangle **t**, as a systematic common factor, **b** as a systematic (intentional bias), and **e** as unique factor or measurement error.

3.2 Indeterminacy of Scores and Indeterminacy of Consciousness

From a formal point of view *e* cannot be directly determined from Equations 1–3 because there is only one measurement *x* but there are two numbers to be obtained: One number for the common part and one number for the unique part. Therefore, Equations 1–3 are indeterminate without further assumptions. However, Guttman introduced a definition of *t* for Equation 3 that has regularly been used in representations of classical test theory,⁵¹ that is

$$E(\mathbf{x}) = E(\mathbf{t}) + E(\mathbf{e}) = E(\mathbf{t}) + 0 = \mathbf{t}, \quad (4)$$

where *E* denotes the expectation (the average value in the population). Equation 4 defines the ›true score‹ or wanted score of an individual as the expected value of the measured variables *x*. Thus, the average of the values of an infinity of measurements *x* yields the true score *t*, which is a constant in all measurements. However, in empirical settings the population of an infinity of measurements will never be

⁵¹ Guttman 1945. E.g. Lord & Novick 1968 and Zimmerman 2011.

reached, so that the average value of the available measurements will be used instead of the expectation as an estimate of the ›true score‹.⁵² For p variables, this yields

$$\frac{\sum_{i=1}^p \mathbf{x}_i}{p} = \hat{\mathbf{t}}, \tag{5}$$

and,

$$\mathbf{x}_i - \hat{\mathbf{t}} = \mathbf{x}_i - \frac{\sum_{i=1}^p \mathbf{x}_i}{p} = \hat{\mathbf{e}}. \tag{6}$$

Although $\hat{\mathbf{t}}$ is an estimator of \mathbf{t} , entering Equation 3 at the left hand side of Equation 5 yields

$$\frac{\sum_{i=1}^p \mathbf{x}_i}{p} = \frac{\sum_{i=1}^p \mathbf{t}_i}{p} + \frac{\sum_{i=1}^p \mathbf{e}_i}{p} = \mathbf{t} + \frac{\sum_{i=1}^p \mathbf{e}_i}{p} = \hat{\mathbf{t}}. \tag{7}$$

As long as $\frac{\sum_{i=1}^p \mathbf{e}_i}{p} \neq 0$, the resulting $\hat{\mathbf{t}}$ is an indeterminate composite of \mathbf{t} and \mathbf{e} because any value of \mathbf{t} can be combined with any average value of \mathbf{e}_i in order to get a given $\hat{\mathbf{t}}$. The indeterminacy of classical test theory is obvious for a finite number of measurements since the error term does not vanish. However, the indeterminacy is typically not noted when classical test theory is presented for the population of measurements (in the form of Equation 4) where the error term vanishes.

In the same line, Beauducel and Leue argued that scales based on unit-weighted (e.g., personality) questionnaire items imply models that should be tested.⁵³ Thus, the fact that a unit-weighted sum is computed does not imply that a ›true score‹ is unequivocally determined by means of this procedure. Thus, the model implied by the unit-weighted item sum could be wrong, i.e., might not fit to the data. It is also shown in Beauducel and Leue that even unit-weighted sum scales typically imply that items are differentially important.⁵⁴ Moreover, Loevinger’s critique that it is not compelling to conceive items

⁵² Beauducel & Leue 2014b.

⁵³ Beauducel & Leue 2013.

⁵⁴ Beauducel & Leue 2013.

(i.e., measurements) as random samples of a population of measurements has never been completely refuted.⁵⁵ From this point of view the differential weights of the measurements that have been introduced in factor analysis (Equations 1 and 2) might allow for a more realistic measurement model.

It has been shown here that the ›true score‹ estimates or wanted score estimates in the context of classical test theory are indeterminate for a finite number of measurements. For factor analysis, the indeterminacy of the common factor scores has already been noted by Wilson and has been discussed repeatedly.⁵⁶ Although the discussion has focused on the common factor scores, the unique factor scores are as indeterminate as the common factor scores.

Moreover, in conventional factor models, the indeterminate factors are considered as latent variables, as the causes of the measured variables which are considered as the effects.⁵⁷ Thus, the indeterminacy is related to modelling of a latent variable as the cause of the measured variables. Taking Equations 1–3 as models for the generation of the measurements by means of common and unique factors implies that the generation process contains information that cannot be completely reproduced by means of an analysis of the measured variables. As it is impossible to reconstruct the original scores of the generating common and unique factors from the measured variables is exactly what is implied by factor score indeterminacy. Factor score indeterminacy implies that the generation model contains more variables (factors) than the generated data set of measured variables. This complex relationship may be similar to the relationship between the observed reactions of individuals on the one hand and the internal processes, considerations, and behavioral determinants producing the observed reactions. If this complex relationship and the resulting indeterminacy is a property of conscious individuals this would imply that it would be, in principle, impossible to determine definitely whether a conscious reaction has occurred or not. It would than be easier to build up a model generating a set of measured variables for which it is impossible to reconstruct the scores of the generating variables than to reconstruct the scores of the generating variables of any

⁵⁵ Loevinger 1965.

⁵⁶ Wilson 1929. E.g. Guttman, 1955, Schönemann & Wang 1972 and Beauducel & Hilger 2015.

⁵⁷ Bollen & Lennox 1991.

living individual. For example, it is no problem to generate a data set of artificial measured variables that are conform to the indeterminate factor model,⁵⁸ even when it is impossible to perfectly reconstruct the original generating common and unique factor scores. This would imply that the situation for research on consciousness, even when based on neurocognitive and behavioral data, is similar to the situation in the Turing test, where a number sentences are available as a basis for the attribution of (un-)consciousness to the communicator. In this sense, neurocognitive correlates of consciousness cannot definitely indicate whether consciousness was present or not. It should, however, be noted that behavioral predictions that are based on the indeterminate factor model and the indeterminate test theories are generally quite successful, as for example when intelligence scales are used for the prediction of job success.⁵⁹

3.3. Advantages of Indeterminate Models and their Relationship to Models of Consciousness

It follows from the previous paragraph that indeterminate models contain a fundamental limitation for research because the exact scores of the original variables cannot be exactly reproduced from the number of measured variables. However, as has been noted in Section 3.1., the success of psychological behavioral predictions in several applied fields⁶⁰ are related to the use of indeterminate models since Spearman.⁶¹ One may therefore ask whether the indeterminate factor model and the indeterminate test theories have properties that make them suitable for behavioral research. It has been noted that indeterminacy is based on the fact that the number of factors (comprising common and unique factors) is larger than the number of measured variables. It has also been noted that consciousness may operate on the basis of a system that is more complex than the resulting behavior. Therefore, possible advantages of an indeterminate psychometric model may also be advantages of a complex and, possibly indeterminate, system based on consciousness.

⁵⁸ Beauducel & Hilger 2017.

⁵⁹ Schmidt & Hunter 1998.

⁶⁰ Schmidt & Hunter 1998.

⁶¹ Spearman 1904.

One reason why indeterminate models comprising more factors (latent variables) than measured variables could be advantageous is that common and unique factors can be equally distributed across all measured variables. Consider, for example Equations 1, where each measured variable is decomposed into at least one common factor and one unique factor. It follows that there are p unique factors when there are p measured variables (one unique factor for each variable). Since there is at least one common factor in the factor model, this implies that the sum of common and unique factors n_f is greater or equal than $p + 1$.

But why is this an advantage? The advantage of the large number of factors is, that each measured variable, is treated equally within the model. That is, the model allows each measured variable to contain specific or unique variance as well as common variance. If a model is determinate, it cannot contain more variables than measured variables. Principal component analysis⁶² is presented as an example for such a model. The aim of principal component analysis is to find components representing a maximum of the variance of the measured variables, regardless whether the variance is common or unique. In the principal component model, the number of components equals the number of measured variables, that is, $n_c = p$. Therefore, the principal component scores are determinate, i.e., they can unambiguously be computed from the measured variables. When there is at least one common component, it follows from $n_c = p$ that the number of components representing the unique variance is smaller than the number of measured variables. In consequence, it is impossible that the unique variance of each measured variable is represented by a unique component. The consequences of $n_c = p$ for the representation of the common and unique variance of the measured variables in the principal component model is demonstrated by means the following example based on artificial data.

Example: Comparison of the (Indeterminate) Factor Model and the (Determinate) PCA

A first artificial data set (Sample 1) was based on $n = 2,000$ cases and five normally distributed, standardized measured variables. The sample was generated by means of a random number generator of IBM

⁶² Harman 1967.

SPSS (Version 26). The measured variable was designed to represent 50 % of common variance and 50 % of unique variance in the population. Such a perfectly symmetric pattern will practically never occur but it is informative to see how the variances are represented when these five variables are represented in the factor model and in the principal component model. The amount of variance that the measured variables share with the unrotated common and unique factors and the unrotated principal components is represented by the squared loadings. The sample size of 2,000 cases implies that only a minimal amount of sampling error may lead to departures from squared equal population loadings of the measured variables (x1 – x5) on the common and unique factors (see Table 1).

Table 1. Squared sample loadings for the factor model and for the principal component model

Measured variables	Sample 1 (n = 2,000)						Sample 2 (n = 2,000)					
	squared factor loadings						squared factor loadings					
	common			unique			common			unique		
x1	.50	.50	.00	.00	.00	.00	.50	.50	.00	.00	.00	.00
x2	.50	.00	.50	.00	.00	.00	.50	.00	.50	.00	.00	.00
x3	.50	.00	.00	.50	.00	.00	.50	.00	.00	.50	.00	.00
x4	.50	.00	.00	.00	.50	.00	.50	.00	.00	.00	.50	.00
x5	.50	.00	.00	.00	.00	.50	.50	.00	.00	.00	.00	.50
	squared component loadings						squared component loadings					
x1	.60	.00	.03	.37	.00	-	.60	.39	.00	.00	.01	-
x2	.60	.00	.03	.04	.33	-	.60	.02	.33	.03	.01	-
x3	.60	.25	.03	.04	.08	-	.60	.05	.03	.00	.32	-
x4	.60	.00	.40	.00	.00	-	.60	.02	.00	.33	.04	-
x5	.60	.25	.03	.04	.08	-	.60	.01	.13	.13	.13	-

Note. All variances greater than zero are given in bold face.

As can be seen in Table 1, the partition of the variance (squared loadings) on the common and unique factors was as expected for the factor model in Sample 1. However, in the component model the squared loadings on the first component were considerably larger than the squared loadings on the common factor. Although this has been

regarded as an indication that the principal component model overestimates the loadings,⁶³ the focus of the present demonstration is on the difference between the unique variance which is equally distributed across the measured variables in the factor model whereas the distribution of the variance on the components 2–5 is rather unequal in the component model.

Sample 2 was generated in order to investigate the robustness of the variances explained by the factor model and the principal component model. Therefore, Sample 2 was based on the same population model. Moreover, Sample 2 was identical to Sample 1 with the only exception that the value of the first case in the first value was increased by 0.01 (it was -0.03 in Sample 1 and -0.02 in Sample 2). No other values were changed so that 99.99 % of the values in Sample 2 were identical to the values in Sample 1. As expected, the variance distribution on the common factor and the unique factor was not altered by this minimal modification. The variance explained by the first component also remained unchanged. However, the variance explained by the components 2–5 was quite different. Since the modification of the sample was minimal, this indicates that the component model does not result in robust estimations of the more specific variances.

The variability of parameters of the indeterminate factor model and of the determinate principal component model can also be shown in a more complex simulation study. First, a population of three common factors was generated that can also be described by three salient principal components. The corresponding population loadings are given in the Appendix (Table A1). From this population 1,000 samples with $n = 400$ cases were drawn and submitted to common factor analysis as well as to principal component analysis. For each sample, a factor analysis and a principal component analysis was performed and -as in the previous example with two samples- the variability (standard deviation) of the unique factor loadings and the standard deviation of the non-salient principal component loadings was investigated. The resulting standard deviation of the unique factor loadings (Ψ) was $s = .033$ whereas the standard deviation of the non-salient principal component loadings (N) was $s = .200$. For the common factor loadings, the standard deviation was $s = .301$ whereas it was $s = .364$ for the salient principal component loadings. Thus, the standard deviations of the factor loadings were smaller than the stan-

⁶³ Snook & Gorsuch 1989.

dard deviations for the principal component loadings. This indicates that the principal component loadings depend more on sampling error than the factor loadings. As sampling error represents a variability that is due to noise, a model that is less affected by sampling error can be regarded as statistically more robust. Although more complex simulation studies would be necessary in order to explore the conditions of the instability of the component variances/loadings, it is demonstrated here that the indeterminate factor model may yield more robust model parameters than the determinate component model.

The loadings of the common factors have been compared with the corresponding principal component loadings in more comprehensive studies on the effect of changing contexts of variables on results. These studies have found that the common factor loadings were more robust against changes of the context of variables than the respective component loadings.⁶⁴ Although these studies did not investigate the robustness of the unique loadings and their corresponding principal components, they also found that the indeterminate factor model was more robust than the determinate principal component model.

The initial question of this section was, whether there could be possible advantages of indeterminate models. In light of the example shown here and in light of previous research one may conclude that the robustness of model parameters may be enhanced for indeterminate models. It could be that the larger number of model parameters allows to represent small empirical variations more conveniently, which may explain the enhanced robustness of indeterminate models. This may—in turn—explain the success of such models in the prediction of human behavior.

4. Conclusion

Since indeterminate models are robust across minor parameter variations, it is possible to assume that consciousness—which may help to provide behavior that is consistent across time—can be described by models having more parameters than can be measured by means of observed variables. Such models of consciousness are indeterminate in that the original scores representing consciousness cannot be

⁶⁴ Widaman 1993 and Beauducel 2000.

completely reproduced from the measured data. As mentioned in the introduction, one may speculate whether it is adaptive for organisms to be based on more parameters than necessary for the generation of immediate behavioral outcome. Being robust against minor changes of the observed variables may facilitate temporal continuity of behavior. If consciousness is related to the fact that the number of model parameters is larger than the number of observed outcome variables, consciousness may be related to behavioral continuity. The perspective outlined in this chapter implies that neurocognitive correlates of consciousness cannot definitely indicate whether consciousness was present or not. Moreover, this perspective is compatible with the results of the previous sections, namely (Section 1.) that consciousness is related to a complex coordination of processes that is not necessarily limited to a specific brain region, that (Section 2.) consciousness is related to very basic processes like the dynamic attribution of data to passing moments of time and to the reconfiguration of stimulus material while the consciousness of an individual reaction cannot unambiguously be assumed, and (Section 3.) that conscious reactions can be the basis of variance components representing the true part of the variance as well as of variance components representing measurement error. This distribution of consciousness on several variance components implies that the indeterminate factor model may be appropriate for the description of conscious behavior. It remains to be explored whether the ideas presented here can be related to aspects of indeterminacy that are discussed in other fields,⁶⁵ especially in the field of artificial intelligence.⁶⁶

References

- Agnati, Luigi Francesco, Guidolin, Diego, Cortelli, Pietro, Genedani, Susanna, Cela-Conde, Camilo, & Fuxe, Kjell (2012): Neuronal correlates to consciousness. The »Hall of Mirrors« metaphor describing consciousness as an epiphenomenon of multiple dynamic mosaics of cortical functional modules, in *Brain Research* 1476, 3–21. <https://doi.org/10.1016/j.brainres.2012.01.003>
- Beauducel, André (2001): On the generalizability of factors: The influence of changing contexts of variables on different methods of factor extraction, in *Methods of Psychological Research* 6, 1–24.

⁶⁵ del Val 2020.

⁶⁶ Wit et al. 2018.

- Beauducel, André & Hilger, Norbert (2015): Extending the debate between Spearman and Wilson 1929: When do single variables optimally reproduce the common part of the observed covariances?, in *Multivariate Behavioral Research* 50, 555–567.
- Beauducel, André & Hilger, Norbert (2017): On the bias of factor score determinacy coefficients based on different estimation methods of the exploratory factor model, in *Communications in Statistics – Simulation and Computation* 46, 6144–6154. <https://doi.org/10.1080/03610918.2016.1197247>
- Beauducel, André & Leue, Anja (2013): Unit-weighted scales imply models that should be tested!, in *Practical Assessment, Research & Evaluation* 18(1).
- Beauducel, André & Leue, Anja (2014a): *Grundlagen der Psychologischen Diagnostik*, Göttingen: Hogrefe.
- Beauducel, André & Leue, Anja (2014b): Testing the assumption of uncorrelated errors for short scales by means of structural equation modeling, in *Journal of Individual Differences* 35, 201–211.
- Bollen, Kenneth & Lennox, Richard (1991): Conventional wisdom and measurement: A structural equation perspective, in *Psychological Bulletin* 110, 305–314. <https://doi.org/10.1037/0033-2909.110.2.305>
- Casali, Adenauer G., Gosseries, Olivia, Rosanova, Mario, Boly, Mélanie, Saasso, Simone, Casali, Karina R., Casarotto, Silvia, Bruno, Marie-Aurélié., Laureys, Steven, Tononi, Giulio, & Massimini, Marcello (2013): A theoretically based index of consciousness independent of sensory processing and behavior, in *Science Translational Medicine* 5, 198: 198ra105.
- Cattell, Raymond Bernard & Radcliffe, John A. (1962): Reliabilities and validities of simple and extended weighted and buffered unifactor scales, in *The British Journal of Statistical Psychology* 15, 113–128.
- Cronbach, Lee J., Gleser, Goldine C., Nanda, Harinder & Rajaratnam, Nageswari (1972): *The Dependability of Behavioral Measurements: Theory of Generalizability of Scores and Profiles*, New York: Wiley.
- Dehaene, Stanislas, Charles, Lucie, King, Jean-Rémi, & Marti, Sébastien (2014): Toward a computational theory of conscious processing, in *Current Opinion in Neurobiology* 25, 76–84.
- Del Val, Jaime (2020): Beyond error: Philosophy of indeterminacy in the age of algorithms, in Popat, Sita and Whatley, Sarah (eds.): *Error, ambiguity, and creativity: A multidisciplinary reader*. Cham, Switzerland: Springer Nature, 85–106. <https://doi.org/10.1007/978-3-030-39755-5>
- Demertzi, Athena, Tagliazucchi, Enzo, Dehaene, Stanislas, Deco, Gustavo, Bartfeld, Pablo, Raimondo, Federico, Martial, Charlotte, Fernández-Espejo, Davinia, Rohaut, Benjamin, Voss, Henning U., Schiff, Nicholas D., Owen, Adrian M., Laureys, Steven, Naccache, L., Sitt, Jacobo Diego (2019): Human consciousness is supported by dynamic complex patterns of brain signal coordination, in *Science Advances* 5: eaat7603.
- Goense, Jozien, Bohraus, Yvette, & Logothetis, Nikos K. (2016): fMRI at high spatial resolution: Implications for BOLD-Models, in *Frontiers in Computational Neuroscience* 10, 66. <https://doi.org/10.3389/fncom.2016.00066>

- Gray, Jeffrey A., & McNaughton, Neil (2000): *The Neuropsychology of Anxiety*, Oxford: University Press.
- Guttman, Louis (1945): A basis for analyzing test-retest reliability, in *Psychometrika* 10, 255–282.
- Guttman, Louis (1955): The determinacy of factor score matrices with applications for five other problems of common factor theory, in *British Journal of Statistical Psychology* 8, 65–82. <https://doi.org/10.1111/j.2044-8317.1955.tb00321.x>
- Hagberg, Garry L. (2018): Wittgenstein and the concept of consciousness, in Jacquette, D. (ed.), *The Bloomsbury Companion to the Philosophy of Consciousness*, Bloomsbury Publishing, Kindle-Version, 57–77.
- Hambleton, Ronald K. & Swaminathan, Hariharan (2013): *Item Response Theory: Principles and Applications*, New York: Springer.
- Harman, Harry H. (1976): *Modern Factor Analysis* (3rd ed.), Chicago, : The University of Chicago Press.
- Humphreys, Lloyd G. (1962): The organization of human abilities, in *American Psychologist* 17, 475–483.
- Jacquette, Dale (2018): Cracking the Hard Problem of Consciousness. In Jacquette, D. (ed.), *The Bloomsbury Companion to the Philosophy of Consciousness*, Bloomsbury Publishing, Kindle-Version, 258–287.
- Klein, Christoph, Arend, Isabel C., Beauducel, André, & Shapiro, Kimron L. (2011): Individuals differ in the attentional blink: Mental speed and intra-subject stability matter, in *Intelligence* 39, 27–35.
- Koch, Christof., Massimini, Marcello, Boly, Melanie, & Tononi, Giulio (2016): Neural correlates of consciousness: Progress and problems, in *Nature Reviews Neuroscience* 17, 307–321.
- Kranczioch, Cornelia, Debener, Stefan, & Engel, Andreas K. (2003): Event-related potential correlates of the attentional blink phenomenon, in *Cognitive Brain Research* 17, 177–187.
- Lamy, Dominique, Salti, Moti, Bar-Haim, Yair (2009): Neural correlates of subjective awareness and unconscious processing: an ERP study, in *Journal of Cognitive Neuroscience* 21, 1435–1446.
- Lazarsfeld, Paul Felix (1959): Latent structure analysis, in Koch, S. (ed.), *Psychology: a study of a science* Vol. 3. New York: McGraw-Hill.
- Lehrer, Keith (2018): Conscious exemplars as freedom of representation, in Jacquette, D. (ed.), *The Bloomsbury Companion to the Philosophy of Consciousness*, Bloomsbury Publishing, Kindle-Version, 104–122.
- Loevinger, Jane (1965): Person and population as psychometric concepts, in *Psychological Review* 72, 143–155.
- Lord, Frederic M. & Novick, Melvin R. (1968/2008): *Statistical Theories of Mental Test Scores*, Addison-Wesley, Menlo Park.
- Magnusson, David (1967): *Test Theory* (2nd ed), London: Addison-Wesley.
- Mulaik, Stanley A. (2012): *Foundations of Factor Analysis*, New York: Taylor & Francis.

- Odegaard, Brian, Knight, Robert Thomas, & Lau, Hakwan (2017): Should a few null findings falsify prefrontal theories of conscious perception?, in *The Journal of Neuroscience* 37 (40), 9593–9602.
- Penco, Carlo (2012): Updating the Turing test. Wittgenstein, Turing and symbol manipulation, in *Open Journal of Philosophy* 2, 189–194.
- Popper, Karl Raimund (1979): *Objective Knowledge. An Evolutionary Approach* (revised ed.), Oxford: Oxford University Press.
- Rutiku, Renate, Martin, M., Bachmann, Talis, & Aru, Jaan (2015): Does the P300 reflect conscious perception or its consequences?, in *Neuroscience* 298, 180–189.
- Schönemann, Peter H., & Wang, Ming-Mei (1972): Some new results on factor indeterminacy, in *Psychometrika* 37, 61–91. <https://doi.org/10.1007/BF02291413>
- Schmidt, Frank L. & Hunter, John E. (1998): The validity and utility of selection methods in personnel psychology. Practical and theoretical implications of 85 years of research findings, in *Psychological Bulletin* 124, 262–274.
- Shapiro, Kimron L., Arnell, Karen A., & Raymond, Jane E. (1997): The attentional blink: A view on attention and a glimpse on consciousness, in *Trends in Cognitive Science* 1, 29–296.
- Snook, Steven C. & Gorsuch, Richard L. (1989): Component analysis versus common factor analysis: A Monte Carlo study, in *Psychological Bulletin* 106, 148–154.
- Sohn, Emily (2019): Decoding Consciousness, in *Nature* 571, S1-S5. <https://doi.org/10.1038/d41586-019-02207-1>
- Spearman, Charles (1904): »General intelligence.« Objectively determined and measured, in *The American Journal of Psychology* 15, 201–292.
- Steyer, Rolf, Schmitt, Manfred & Eid, Michael (1999): Latent state-trait theory and research in personality and individual differences, in *European Journal of Personality* 13, 389–408.
- Turing, Alan Mathison (1950): Computing Machinery and Intelligence, in *Mind* 236, 430–460.
- Widaman, Keith F. (1993): Common factor analysis versus principal component analysis: Differential bias in representing model parameters?, in *Multivariate Behavioral Research* 28, 263–311.
- Wilson, Erik B. (1929): Comment on Professor Spearman's note, in *Journal of Educational Psychology* 20, 217–223. <https://doi.org/10.1037/h0071925>
- Wit, Andrew John, Vasey, Lauren, Parlac, Vera, Marcu, Mara, Jabi, Wassim, Gerber, David, Daas, Mahesh, & Clayton, Mark (2018): Artificial intelligence and robotics in architecture: Autonomy, agency, and indeterminacy, in *International Journal of Architectural Computing* 16, 245–247. <https://doi.org/10.1177/1478077118807266>
- Wittgenstein, Ludwig (1958a): *Philosophical Investigations*, Oxford, UK: Blackwell.
- Wittgenstein, Ludwig (1958b): *The Blue and Brown Books: Preliminary Studies for the "Philosophical Investigations"*, Oxford, UK: Blackwell.
- Wittgenstein, Ludwig (1969): *On Certainty*, Oxford, UK: Blackwell.

- Wittmann, Werner W. (1988): Multivariate reliability theory: Principles of symmetry and successful validation strategies, in Nesselroade, J. R. & Cattell, R. B. (eds.): *Handbook of Multivariate Experimental Psychology* 2nd ed. New York: Plenum, 505–560.
- Zimmerman, Donald W. (2011): Sampling variability and axioms of classical test theory, in *Journal of Educations and Behavioral Statistics* 36, 586–615. <https://doi.org/10.3102/1076998610397052>

Appendix

Table A1. Population common factor loadings and population principal component loadings

	Common factor loadings			Salient component loadings		
	F1	F2	F3	C1	C2	C3
x1	.55	.00	.00	.66	.00	.00
x2	.55	.00	.00	.66	.00	.00
x3	.55	.00	.00	.66	.00	.00
x4	.55	.00	.00	.66	.00	.00
x5	.55	.00	.00	.66	.00	.00
x6	.55	.00	.00	.66	.00	.00
x7	.00	.53	.00	.00	.65	.00
x8	.00	.53	.00	.00	.65	.00
x9	.00	.53	.00	.00	.65	.00
x10	.00	.53	.00	.00	.65	.00
x11	.00	.53	.00	.00	.65	.00
x12	.00	.53	.00	.00	.65	.00
x13	.00	.00	.50	.00	.00	.63
x14	.00	.00	.50	.00	.00	.63
x15	.00	.00	.50	.00	.00	.63
x16	.00	.00	.50	.00	.00	.63
x17	.00	.00	.50	.00	.00	.63
x18	.00	.00	.50	.00	.00	.63

Note. Salient loadings are given in bold face.

Artificial Intelligence and Rational Discourse

1. Introduction

The term ›artificial intelligence‹ (AI) which John McCarthy invented for the famous Dartmouth Summer Research Project on Artificial Intelligence proposal he submitted to the Rockefeller Foundation in 1955 with his colleagues Marvin Minsky, Nathaniel Rochester, and Claude Shannon, has always been controversial. It might, therefore, seem idle to make it the subject of an examination again. However, the term has been almost omnipresent for several years now and it has significantly shaped debates about many technical developments. It is not only the word ›artificial intelligence‹, but also the divergent meanings, or rather, the promises, that many associate with it that have made it controversial. In this paper, I aim at presenting some thoughts on the notion of artificial intelligence and finally place a second notion—that of rational discourse—alongside the former one. In this way, I hope to introduce a perspective that moves the infamous, but in my view over-emphasized question ›When will machines have surpassed us in X?‹ aside. The more interesting question, I submit, is ›Who is ›us?‹.

2. Defining Artificial Intelligence

Some argue that AI is simply the field devoted to building a machine that can pass the notorious Turing test.¹ Though highly controversial, the Turing test certainly is still an important benchmark.² For my considerations, however, other attempts to define AI are more relevant in the first place. A frequently read—though deliberately tenta-

¹ Turing 1950.

² Moor 2003.

tive—definition of AI is as follows: »Artificial Intelligence is the study of how to make computers do things at which, at the moment, people are better.«³

According to this approach, AI research is about human capacities, or at least capacities that humans have, and about implementing them in artificial systems, and possibly even about designing these systems to outperform humans in terms of these capacities. Of course, this is still rather vague and imprecise.

In their influential textbook on AI, Stuart Russell and Peter Norvig distinguish two dimensions in AI definitions that can help get a slightly better and more accurate handle on the matter. Accordingly, definitions can be distinguished along two different types of goals, first, human-based vs. ideal rationality-based definitions, and second, reasoning-based vs. behavior-based definitions.⁴ This results in a total of four types of definitions, which apparently capture a large proportion of all proposed definitions of AI. The above-mentioned definition by Rich, Knight and Nair, for example, clearly falls into the category human-based / behavior-based, for it highlights that AI research aims at making computers *do* certain things in which *humans* are currently especially good. In contrast to Rich, Knight and Nair, Russell and Norvig themselves emphasize *perfect rationality*, and therefore fall into the category ideal rationality-based / behavior-based. They state the goal of their approach to AI as designing »successful agents«.⁵ Later, they define an ideal rational agent as follows:

For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has.⁶ From this it should be clear that the authors focus on perfect rationality and behavior.

In their introduction to the *Cambridge Handbook of Artificial Intelligence*, Keith Frankish and William M. Ramsey observe:

Very generally, artificial Intelligence (AI) is a cross-disciplinary approach to understanding, modeling, and replicating intelligence and

³ Rich, Knight & Nair 2010, 3.

⁴ Russell & Norvig 2010, 4–8.

⁵ Russell & Norvig 2010, 34.

⁶ Russell & Norvig 2010, 37.

cognitive processes by invoking various computational, mathematical, logical, mechanical, and even biological principles and devices.⁷

The lack of reference to actions makes this definition fall into the reasoning-based category and the lack of reference to humans puts it into the ideal rationality-based subfield. This shows that there are also definitions belonging to this category in use.

However, the majority of AI research today seems to focus on behavior, i. e. on how to make computers *do* certain things. One may actually assume that it is never exclusively about reasoning processes, so that behavior-based definitions—at least if one takes a broad understanding as a basis—can serve as a comprehensive definition of AI in the one dimension.

This is in line with an expansion of the term ›agent‹ that can be observed in the context of AI research, but also in philosophy of AI. Again, Russell and Norvig are a case in point. In fact, they consider AI as the field concerned with the development of intelligent *agents*.⁸ They do so against the background of a very broad understanding of ›agency‹:

An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators. [...] A human agent has eyes, ears, and other organs for sensors and hands, legs, vocal tract, and so on for actuators. A robotic agent might have cameras and infrared range finders for sensors and various motors for actuators. A software agent receives keystrokes, file contents, and network packets as sensory inputs and acts on the environment by displaying on the screen, writing files, and sending network packets.⁹

This understanding of the term ›agent‹ is by no means self-evident. On the contrary, it even contradicts an established way of using it. It therefore seems appropriate to take a closer look at the concepts of agent and agency.

3. Agents

Traditionally, the term ›agency‹ was used as an opposite term to behavior: While non-human animals behave, humans act. In this sense,

⁷ Frankish & Ramsey 2018, 1.

⁸ Russell & Norvig 2010, 4f.

⁹ Russell & Norvig 2010, 3.

agency is closely linked to other notions, especially to the notion of responsibility. Thus, the concept of responsibility is often brought into play when it comes to the argument that non-humans (non-human animals and artificial systems alike) cannot be agents, because they cannot assume moral responsibility. However, this conventional understanding of ›agency‹ has been replaced for some time by a much broader and more general understanding.

At a sufficiently high level of abstraction,¹⁰ an agent is someone who (or something that) can act, and to be able to act is to be able to effect changes in the world.¹¹ As Floridi correctly observes, this abstract definition includes not only ordinary people but also earthquakes since they, apparently, are able to effect changes in the world. Floridi indicates that the definition's level of abstraction is too high for his purposes, and this holds true also for the present purpose. He therefore suggests turning to a lower level of abstraction so that the following criteria are included: interactivity, autonomy, and adaptability. Note that autonomy is not meant as a philosophically rich notion here, but merely designates the ability of a system »to change its state without direct response to interaction: it can perform internal transitions to change its state.«¹²

The crucial consequence of this understanding of the notion of an agent and agency respectively is that it makes it possible to distinguish agency from *moral* agency. While all moral agents are, naturally, agents, not all agents need to be moral agents. As mentioned, this stands in contrast to a long philosophical tradition that conceives of agency exclusively as moral agency. However, it allows entities that share interesting properties to be grouped under one term. If these entities fulfill the three criteria mentioned above, namely interactivity, autonomy, and adaptability, they all belong together and are, in turn, distinct from entities which do not meet the criteria. In this perspective, there are consequently two basic categories: agents and non-agents. If one follows this approach, then it should be uncontroversial that some AI systems are agents—and in fact, this is what Floridi and many scholars from the field of AI argue for. There is nothing wrong with that. In some ways, AI systems and humans share more than AI systems and chairs or humans and chairs respectively. In short, both humans and AI are agents—or at least can be agents.

¹⁰ Cf. Floridi 2015, Chap. 3.

¹¹ Cf. Floridi 2015, 140.

¹² Floridi 2015, 140.

4. Intelligence

If the term ›agent‹ is open to several interpretations, then the term ›intelligence‹ is even more so. In a statement first published in December 1994 in the *Wall Street Journal* and later as an editorial in the journal *Intelligence*, 52 experts suggested the following broad understanding of the concept:

Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings—›catching on,‹ ›making sense‹ of things, or ›figuring out‹ what to do.¹³

This definition is primarily directed against racist appropriations of the term and explicitly interprets intelligence as a human capability. However, there is nothing to be said against detaching this general definition from the historical context of its origin and applying it to current developments. If one does so, then one will have to concede that some modern computer systems are quite capable of meeting the requirements described. They are certainly able of comprehending their surroundings, catching on, making sense of things, and figuring out what to do. The same will be said of many animal species. This simply means that there are many entities to which the characterization ›intelligent‹ fits.

What follows from all this? First of all, it makes sense to speak of artificial intelligent agents, i.e., one must concede that there are artificial systems that are agents and that they are intelligent. Conversely, it follows that the question ›Can machines act intelligently?‹ must be answered positively. On this very question Russell and Norvig remark:

Our definition of AI works well for the engineering problem of finding a good agent, given an architecture. Therefore, we're tempted to end this section right now, answering the title question in the affirmative. But philosophers are interested in the problem of comparing two architectures—human and machine. Furthermore, they have traditionally

¹³ Gottfredson 1997, 13.

posed the question not in terms of maximizing expected utility but rather as, »Can machines think?«¹⁴

In the end, then, the definitional question of what AI is leads back to a comparison of two architectures, humans, and artificial systems. Even if we concede that both architectures allow intelligent agents to be implemented, there still seems to be an open question.

Russell and Norvig discuss a whole series of philosophical arguments and finally conclude that (phenomenal) consciousness marks a significant difference that continues to be difficult, but not really of central importance from an engineering perspective.

Running through all the debates about strong AI—the elephant in the debating room, so to speak—is the issue of consciousness. [...] Qualia are challenging not just for functionalism but for all of science. Turing himself concedes that the question of consciousness is a difficult one, but denies that it has much relevance to the practice of AI [...]. We agree with Turing—we are interested in creating programs that behave intelligently. The additional project of making them conscious is not one that we are equipped to take on, nor one whose success we would be able to determine.¹⁵

One can summarize this train of thought in the following way: The comparison between artificial systems and humans ultimately boils down to the question Turing rejected as too vague: ›Can machines think?‹ The answer to this question depends on how one deals with the problem of consciousness: If ›thinking‹ is understood to mean intelligent action, then they can; if it is understood to mean conscious action, then they cannot. AI in practice is primarily concerned with the former and not with the latter. But perhaps the question whether machines can think is just a proxy, at least for those who don't take an engineering perspective. The real question might be ›Can machines be like us?‹

If you actually think this question is interesting, then another traditional term might help: rationality. Authors like Russell and Norvig and others use ›intelligent action‹ and ›rationality‹ almost synonymously. They maintain: »A rational agent is one that acts so as to

¹⁴ Russell & Norvig 2010, 1021.

¹⁵ Russell & Norvig 2010, 1033.

achieve the best outcome or, when there is uncertainty, the best expected outcome.«¹⁶

A broader understanding of rationality as opposed to intelligence is mediated by the notion of reasons or, more precisely, the exchange of reasons. Therefore, I suggest instead of asking ›Can machines act intelligently?‹ (Yes) or ›Can machines think?‹ (Unclear) focusing on the question ›Can machines be rational?‹—if only as an intermediate step to addressing the question, ›Are they like us?‹

5. Rationality as the Giving and Asking for Reasons

Among others, Robert Brandom has proposed and detailed an approach of rationality as the giving and asking for reasons. Following Wilfrid Sellars, Brandom has used the metaphor of the space of reasons for this. Famously, Sellars argued in his *Empiricism and the Philosophy of Mind*:

The essential point is that in characterizing an episode or a state as that of knowing, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says.¹⁷

Brandom subsequently highlighted, among other things, the social structure of the space of reasons. He writes in this regard:

Thinking of things this way, assessing someone as having successfully achieved the status or standing of a knower involves adopting three different attitudes: *attributing* a commitment, *attributing* an entitlement, and *undertaking* a commitment. There is nothing in principle mysterious about such assessments, nor, therefore, about the standing being assessed. Knowledge is intelligible as a standing in the space of reasons, because and insofar as it is intelligible as a status one can be taken to achieve in the game of giving and asking for reasons. But it is essentially a *social* status because it incorporates and depends on the social difference of perspective between *attributing* a commitment (to another) and *undertaking* a commitment (oneself). If one *individualizes* the space of reasons, forgetting that it is a shared space within which we adopt attitudes towards *each other*—and so does not think about standings in the space of reasons as socially articulated, as potentially

¹⁶ Russell & Norvig 2010, 4.

¹⁷ Sellars 1997, § 36.

including the social difference of perspective between attributing and undertaking commitments, that is, between your standing and mine—then one will not be able to understand knowledge as a standing in the space of reasons.¹⁸

If one understands rationality in this way, i. e. in terms of giving and asking for reasons, it is essentially a social practice, or a game played by normative beings, i. e. beings who can undertake commitments.

This brings us back to Russell and Norvig's as well as Floridi's wider notion of agents and agency. It is possible, as I said, to decouple the notion of agency from that of responsibility so that AI systems also fall under it. And it is also possible to understand intelligence in a somewhat neutral sense. What is not possible, according to Brandom, is to decouple the notion of rationality from that of epistemic responsibility, for to be rational is to undertake commitments.

Now, moral responsibility and epistemic responsibility share the feature of being normative. Thus, it is impossible to conceive of rationality in non-normative terms. Moreover, if we widen the notion of agent and agency to the effect that humans and AI systems can fall under it, we end up with a notion that includes both normative and non-normative entities. This is no problem, if putting the two in one category (for certain purposes) does not obscure this important difference. What is most important however is that admitting the existence of moral and non-moral agents is not enough for it can lead to the view that AI is like us in every respect. True, AI systems can be like us in terms of certain capacities, namely those capacities that fall into the category ›intelligence‹, but this is not the same as rationality. AI systems cannot give reasons for what they are doing. Of course, AI is not like us in another respect: we are alive. I leave this aspect out of consideration here, although of course it should be included in a more extensive account of the differences between humans and AI systems.

The upshot of all this is that notions such as ›intelligence‹ and ›agent‹ can be used to denote both humans and AI systems and by using them we pick out interesting features of both humans and AI systems. It would, however, be wrong to conclude that AI systems are almost like us in every respect, except that for the moment they are not ›moral‹. They are not ›normative‹ and this is a much more far-reaching claim, for it implies that they are not ›rational‹.

¹⁸ Brandom 1995, 903–904. See also Brandom 1994, 199–206.

6. Who are ›we‹?

In the opening passage of the first chapter of *Making it Explicit* Robert Brandom remarks: »We are the ones for whom reasons are binding, who are subject to the peculiar force of the better reason.«¹⁹ This ›we‹ is not, as Brandom highlights, exclusionary or disparaging and certainly not simply limited to humans. This ›we‹ is basically open for other lifeforms, Martians, and also for artificial intelligent agents. But it says something very fundamental about those who use it—it characterizes them as *normative beings*. Questions concerning AI are, eventually, questions about us as normative beings.

The success story of AI will continue. There will certainly be setbacks and frustrations, too. But the presence of AI in daily life will increase and the performance of systems will improve. One should therefore be very careful with predictions of the kind ›Machines will never be able to do X‹. They have been proven wrong too many times. Above all, it should be clear that the intelligence of artificial systems will grow steadily—at least if intelligence is understood to mean the ability to solve problems in the broadest sense. This should be seen as welcome news and by no means as a horror scenario. AI will make life easier for many people. This is not blind enthusiasm for technology, but a reasonable forecast based on experience with other technologies. Of course, there will also be problems and possibly certain groups of people will be worse off because of the development. It is the task of a far-sighted policy to mitigate these negative consequences.

A completely different question is what influence the further development of AI will have on our self-image. If we see ourselves as beings who can play chess or Go particularly well, then this self-image will suffer considerable damage or has already done so. AI systems are already much better at this and there will be more and more areas where their skills trump ours. Perhaps, however, we should not see ourselves primarily as intelligent beings, but rather as rational beings, and more precisely as beings for whom it is characteristic that they play the discursive game of giving and asking for reasons. We are beings who ask who ›we‹ are and that is not a question that can be answered with intelligence alone. Moreover, it is not a capacity in which one can outperform another.

¹⁹ Brandom 1994, 5.

It could be that one day we will find that this ›we‹ also includes other beings than humans—higher animals, inhabitants of other planets, or AI systems. If that should happen, that is, if a representative of one of these groups asks, ›Who are you?‹, then we should probably answer ›Someone like you‹ and welcome them into our midst. Today it is completely unclear whether this will ever happen.

References

- Brandom, Robert (1995): Knowledge and the Social Articulation of the Space of Reasons, in *Philosophy and Phenomenological Research* 55, 895–908.
- Brandom, Robert (1994): *Making It Explicit. Reasoning, Representing, and Discursive Commitment*, Cambridge, Mass.: Harvard University Press.
- Floridi, Luciano (2015): *The Ethics of Information*, Oxford: Oxford University Press.
- Frankish, Keith & Ramsey, William M. (2018): Introduction, in Frankish, Keith; Ramsey, William M. (eds.): *The Cambridge Handbook of Artificial Intelligence*, 3rd ed., Cambridge: Cambridge University Press, 1–11.
- Gottfredson, Linda S. (1997): Mainstream Science on Intelligence: An Editorial With 52 Signatories, History, and Bibliography, in *Intelligence* 24, 13–23.
- Knight, Kevin; Rich, Elaine; Nair, B. (2010): *Artificial Intelligence*, Noida: Tata McGraw Hill.
- Moor, James H. (ed.) (2003): *The Turing Test: The Elusive Standard of Artificial Intelligence*, Dordrecht: Kluwer.
- Russell, Stuart J. & Norvig, Peter (2010): *Artificial Intelligence. A Modern Approach*, 3rd ed., Upper Saddle River: Prentice Hall.
- Sellars, Wilfrid (1997): *Empiricism and the Philosophy of Mind*, Cambridge, Mass.: Harvard University Press.
- Turing, Alan M. (1950): Computing Machinery and Intelligence, in *Mind* LIX, 433–460. Reprinted in Copeland, B. Jack (ed.) (2004): *The Essential Turing*, Oxford: Oxford University Press, 441–464.

Michael H. Herzog

The Temporal Microstructure of Conscious Perception

1. Introduction

Intuitively, there is no problem with time. We feel that time passes by in a smooth, continuous fashion. A diver jumps from a cliff, and we see her trajectory at each single moment in time. Continuous motion in the external world elicits a continuous stream of conscious percepts. However, when dissecting the problem, happy intuition quickly ends up in puzzling questions, which are far from being answered even though the discussion has started more than 1500 years ago.

St. Augustine was driven to despair when thinking about time. Since the »now« is an infinitesimally¹ short moment of time, we can be conscious only at these short moments. However, how can we then perceive motion and change, which, by definition, need to rely on the past and, potentially, the future? St. Augustine proposed that each conscious percept is restricted to a moment and change, and motion percepts are derived from memory:² We are conscious of the current state of the world and may remember past states, which can give rise to the »conclusion« that things have changed. A classic example is that you may leave your home and see the neighbors' trash bin at the walkway. When you return to your home later, the trash bin may be at a different position, and you conclude that it has been moved. Whereas such inferences indeed play a role in cognition, they do not

¹ I am freeing myself from hermeneutic philosophy, sticking to an interpretation, based on my very restricted knowledge of classic texts, formulated in a modern science language; the idea is to mainly lay out the main problems, as I see them, within a historically disguised context; for example St. Augustine did not use the word »infinitesimal«, which here is used in its mathematical meaning, e.g., when time (t) is represented on the x-axis with real numbers.

² Following Dainton (2018), models of this type are nowadays called cinematic models. See Figure 1.

seem to explain our vivid percepts of motion. For this reason, Brentano and Husserl proposed that consciousness is restricted to a moment of time but that these moments contain true motion information. As Dainton pointed out, each time point on the physical clock time axis contains a second axis of mental time with exactly the information that is perceived during the physical moment.³ Of course, one may ask what makes this second axis different from memory. Finally, extensionalism proposes that consciousness lasts for a substantial amount of time rather than being restricted to short moments. Often, these conscious periods are identified with the specious present, which is experimentally estimated to vary between a few hundreds of milliseconds to 3 seconds.⁴ In all of these philosophical models, we are always conscious, which may not be surprising since philosophy usually relies on introspection, making it impossible to realize »gaps« in consciousness. Hence, the main question in the philosophy of the temporal aspects of consciousness is concerned with the duration of the chunks of consciousness (infinitesimally short or substantial, e.g., longer than half a second).

In psychology, for more than a century, the discussion has rather focused on whether consciousness is discrete instead of continuous, i.e., restricted to certain moments in time. During a discrete moment, by definition, no temporal information is available. Discrete perception is motivated by examples known at least since ancient Buddhist times. Consider the following sentence: »the mouse was next to the computer«. Reading this sentence, did you think about the cheese-loving animal or a computer mouse? Often, people think only about the latter because the later coming word »computer« disambiguates the former coming word »mouse«, i.e., well before a conscious percept was elicited. In a strictly continuous model, a thought about the animal would consciously be evoked before the word »computer« is heard.

If we combine the philosophical and psychological models, we arrive at 6 models (see Figure 1).

³ See Dainton 2018. These models are called retentional models. See Figure 1.

⁴ See Pöppel 1997 and Figure 1.

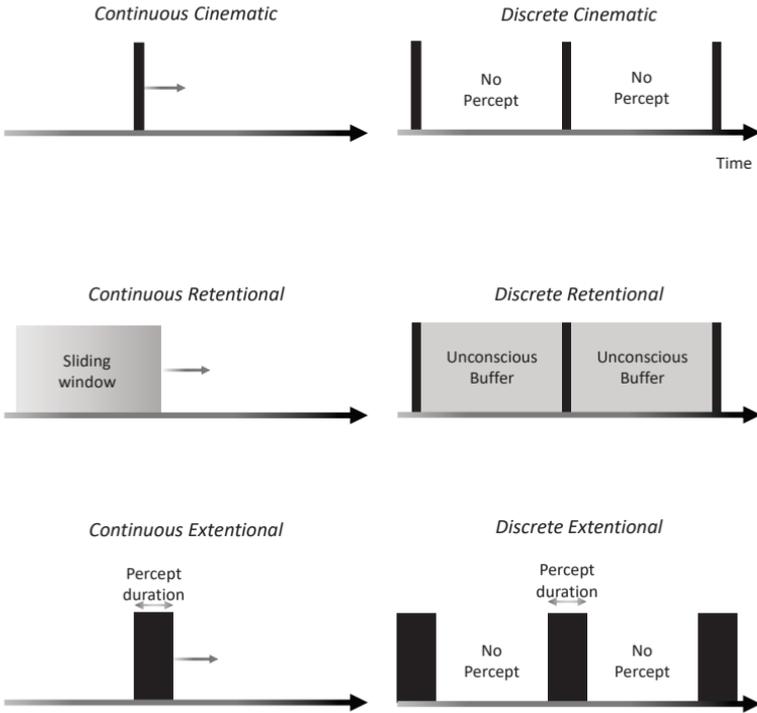


Figure 1.

Continuous models are presented on the left, discrete models on the right. The vertical bars indicate the duration of consciousness, which is (infinitesimally) small in cinematic and retentional models and extended in extentional models. In discrete models, conscious events are separated by periods of no consciousness. The horizontal line in the continuous retentional model indicates the time axis with the temporal information available at this moment in time.⁵

Effects similar to the mouse example, but experimentally better controlled, have been well known in vision research for a long time, such as apparent motion (see Figure 2). These temporal illusions have triggered an intense debate in both neuroscience and philosophy, with very little crosstalk between the two fields. In philosophy, the surplus

⁵ Adapted from Herzog et al. 2020.

content argument states that retentional models predict too many percepts. In apparent motion, for instance, one should first see a static disk, then motion, and finally the second static disk (see Figure 2). However, this is not the case. We *only* see motion. More complex models, including low-pass filtering, do not do any better.⁶ In neuroscience, it is hotly debated when a conscious percept is elicited. Brain time apologists propose that we perceive an object as soon as its representation is elicited in the brain. In other words, time is presented by time. Event time apologists, to the contrary, propose that perception can be delayed in order to account, for example, for latencies between audition and vision. Since auditory processing is faster than visual processing, one would perceive the sound of a spoken word even before one sees the lips moving, similar to a thunder preceding lightning. However, this is not the case. The debate in both philosophy and neuroscience centers on effects and illusions based on delays shorter than 100ms and focuses on the question whether neural delays in the brain can account for these effects.⁷

Recently, a series of postdictive effects have changed the discussion. In these effects, for example, features of elements have been shown to be mandatorily integrated for several hundreds of milliseconds without conscious access to the single elements. As in the example of apparent motion, consciousness cannot have occurred before the last element was presented; thus, consciousness is substantially delayed. These delays are too long to be explained in terms of simple neural latencies. Figure 3 shows the sequential metacontrast paradigm (SQM).⁸

⁶ See Herzog et al. 2020.

⁷ See Holcombe 2013, Arstila 2015 and Piper 2019.

⁸ See Drissi-Daoudi et al. 2019.

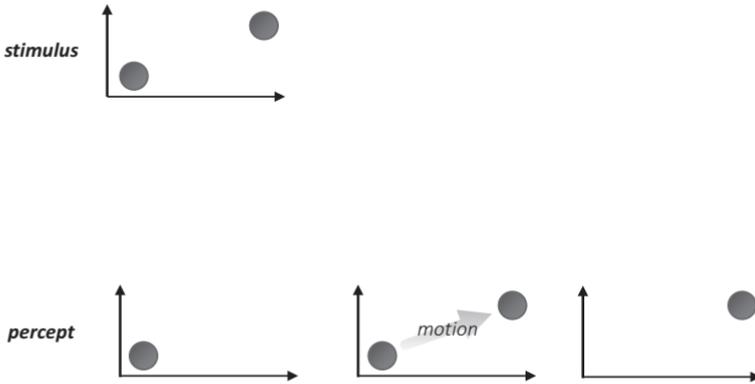


Figure 2.

Apparent motion. One disk is presented for a short duration, followed by a blank interval, and a second disk at a nearby location. We perceive motion from the first to the second disk, even though both disks are static. Importantly, we perceive the motion before the second disk is presented. However, motion direction can only be known after the second disk has appeared. A »too many percepts«-problem occurs for continuous models. In continuous models, one should see the first static disk during the time when it is presented (after some short neural delay), followed by a blank screen and afterwards the third static disk. Finally, one may perceive the illusory motion. However, humans have only one percept and not many.

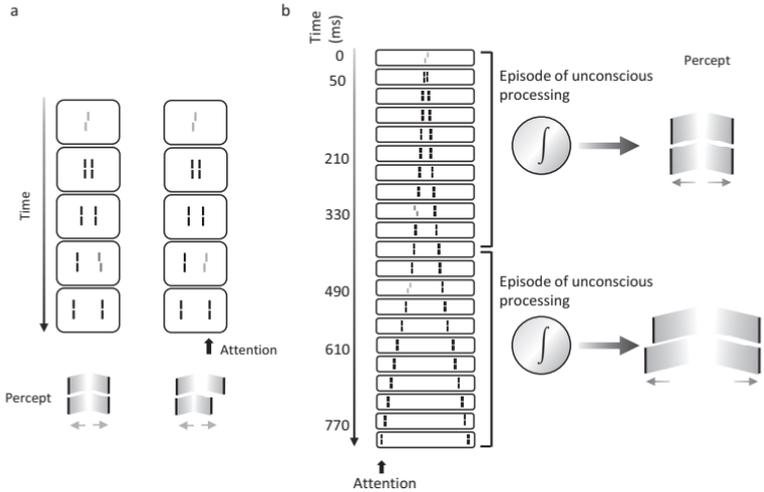


Figure 3.

An example of a long-lasting postdictive effect, the sequential metacontrast paradigm (SQM). A line with a horizontal offset, a so called vernier, is presented on a computer screen. The line is offset either to the left or right, randomly chosen. If presented alone, the vernier and its offset are clearly visible (not shown). **a**) When the vernier is followed by a sequence of lines without offset, the vernier is rendered unconscious. One perceives a stream of lines or, equivalently, two lines expanding from the center. In an experiment, the sequence of lines was presented either with or without the first vernier (the without offset condition is not shown here). Observers could not tell whether or not the vernier was present.⁹ This effect is called metacontrast masking and is known for a century. Next, observers are asked to attend to the left stream of lines and report the offset direction of the visible lines. Even though the lines are all not offset and the vernier is invisible, observers reported the offset direction of the unconscious vernier. Hence, even though invisible, the offset survives metacontrast masking and is integrated with the visible lines. When one of the lines is offset itself, the two offsets integrate when the offsets are separated by less than approximately 400ms (right streams). If vernier and line are offset into the same direction, the percept of the offset becomes stronger (right stream of right example). When they are offset in opposite directions (right stream, left example), no offset is perceived on average. Importantly, integration is mandatory, i.e.,

⁹ Cf. Otto 2006.

observers cannot report the first or second offset separately, even when asked explicitly. If, however, the second offset is presented later than 500ms after than the first one, both offsets can be reported (not shown). They perceive only one integrated offset. **b)** We used streams with more lines and offsets. In this example, the vernier offset and the opposite offset at 330ms integrated mandatorily but the offset at 490ms could be reported independently of the previous two, as we like to argue, because it is in a different period of discrete processing. Importantly, the first two offsets are 330ms apart from each other, the second and third offset only 160ms, hence, integration is not determined by temporal proximity *per se*.¹⁰

Simple neural delays cannot account for these results. Importantly, because these periods are so long, continuous models face a »too many percepts«-problem, which is similar to the surplus content argument.¹¹ In the SQM, for example, one would perceive the first offset consciously, then a series of straight lines, and finally the second offset. In continuous models, it is unclear when the integrated percept should occur or why integration occurs at all. Definitely, consciousness cannot occur before the onset of the second vernier (plus the basic neural latencies). However, as shown experimentally, observers have no conscious access to the individual verniers offsets (see Figure 3).

These results seem to favor discrete models, in which consciousness can be substantially delayed. However, traditional discrete models have considerable problems. For example, when we are conscious only at certain moments, it seems we would miss all the information in between the discrete moments because, by definition, a moment has no time. Moreover, how can we perceive motion at all when consciousness is short? Sometimes, it is proposed that motion occurs as change between the conscious moments. However, in this case, the discrete moments would need to be shorter than 3ms because we can perceive apparent motion on this time scale.¹²

¹⁰ Adapted from Herzog et al. 2020.

¹¹ See Dainton 2018 and Grush 2016.

¹² See Westheimer 1977.

2. The Many Components of the Problem

Physical Time

Thinking about time has changed physics strongly. Pre-Einsteinian physics has largely relied on a continuous flow of absolute time, which holds true for the entire universe¹³ usually indicated with the letter t on the x-axis. This has changed with the theory of relativity where time depends on a reference frame. First, hence, there is no absolute time. Second, in the space time of general theory of relativity, the past and the future exist just as the present. Existence is not restricted to a physical moment. Third, an open question is whether physical time is continuous with infinitesimal small moments of time, well represented by the real numbers. It may well be that time is also discrete. In any case, time comes in quanta of a length of about 10^{-44} s, hence, much shorter than any neural time and hence too »fast«, as we would like to argue, for (conscious) perception.

Hence, our intuition about physical time, derived from our intuition about mental time, is not supported by physics. St. Augustine's conscious moments are not linked to an »objective« physical time. Conscious moments are not linked to physical moments. Hence, there is no gold standard of time we could use. As we will see next, it is also unlikely that mental time is directly linked to neural time.

Neural Time

The timing of neural activity depends on the type of neurons. Classically, neurons can elicit an action potential within 1–2ms, followed by a refractory period of a few milliseconds. However, since perceptual and behavioral measures tap into the response of the entire system or at least several thousands of neurons, temporal resolution may be even higher. Hence, in principle, neural resolution of humans could be in the millisecond regime, which is much lower than most measures of perceptual resolution.

Moreover, perception cannot be immediate because of neural latencies, e.g., from the retina to the brain as well as within the brain. These latencies depend on the contrast of the objects and are roughly

¹³ Galileo has already had the notion of a relative time, however.

in the range of 40ms to 100ms from retina to early visual areas of the brain. These delays are inevitable. However, visual processing takes often much longer before a conscious percept is elicited.

Mental and Temporal Resolution

First, we need to distinguish between the (specious) present, temporal resolution, perception of time, and the timing of a conscious percept. These aspects, even though interwoven, are all logically independent.

Since we cannot link the duration of the moments of consciousness to physical and neural time, we may link them to the perceived present, which is often called the specious present. This time is often estimated to be 2–3s long.¹⁴ During this interval, we feel »now«. Empirical studies have shown that the estimation of the duration of time intervals is rather precise during this time. However, 2–3s are far too long for the minimal duration of consciousness because we can have numerous conscious events during this time.

A classic way to estimate the duration of consciousness is by assessing an observer's temporal resolution. For example, two bars are presented one after the other and observers are asked about their perception. For long durations, the two bars are perceived one after the other. When durations are short (e.g., 40ms), the bars are perceived as simultaneous, as if they were presented during the same moment of time. This makes it seem as if 40ms would be a good estimate for the duration of consciousness. One problem, however, is that different paradigms have led to very different estimates, ranging from 40ms to far beyond 100ms. An even more fundamental problem is that measures of temporal resolution are not necessarily measures of the duration of consciousness. A 40ms duration of simultaneity may simply reflect the temporal resolution of a simultaneity detector. For different tasks, such detectors have different resolutions. These measures are about the content of consciousness but not about the timing of consciousness *per se*. These two temporal aspects are logically independent.

It thus seems that we cannot measure the time of conscious perception directly. However, long lasting postdictive effects, such as the SQM, can provide upper and lower bounds. Since offsets are manda-

¹⁴ See Pöppel 1997.

torily integrated for about 400ms, this sets a lower bound. On the other hand, offsets beyond 490ms can be reported separately, thus setting an upper bound.

3. A Model

The provided examples of postdictive effects (computer mouse, SQM) show that conscious perception of a stimulus can occur much later after stimulus onset than expected from the basic neural delays. In particular, the relative timing of elements and their features is puzzling: we see a continuous stream of lines in the SQM. However, the combined offset we perceive enters into consciousness much later. Hence, the following question arises: what is perceived in the period before a conscious percept occurs?

We suggest: nothing. Conscious percepts are preceded by extended periods of unconscious processing during which we have no percept at all, i.e., consciousness occurs only at discrete moments of time. Each moment contains a »summary« of the entire event structure that happened during the unconscious processing. In this respect, our model is a discrete, retentional model with extended periods of unconscious processing, i.e., the second model in the second row of Figure 1.¹⁵

This model combines the advantages of both continuous and discrete models. Quasi-continuous, unconscious processing assures that information is continuously picked up from the environment and continuously processed. This raw input, however, is rather meaningless: For example, when an element is presented for the first time, it is unclear whether this element is a single static object that will stay confined to one location or will move in the next moment. Likewise, in the mouse example, the presentation of a first word leaves open many interpretations that depend on the subsequent words. We propose that humans have delayed consciousness to render only one interpretation of the stimulus visible. In this respect, slow consciousness, which contains a »summary« of the fast unconscious processing, makes sense. One may wonder whether unconscious processing has indeed processing capacities exceeding the ones of conscious processing by far. However, it is well established that unconscious processing

¹⁵ See Herzog et al. 2016 and Herzog et al. 2020.

is sufficient for carrying out complex cognitive tasks, such as understanding language or chess board situations. Likewise, as the SQM shows, fine-grained spatial and temporal features can be processed unconsciously with high capacity. Consciousness seems to come into play only when information is integrated across large parts of the visual field and semantic information becomes crucial, for example, when comparing two objects or searching for a person in a crowd.¹⁶

There are counter-intuitive aspects of our model. First, we intuitively feel that we perceive motion at each moment of time, as in the example of the diver above. The claim of the model is that we do not perceive the motion of the diver when it occurs. Instead, we perceive it much later and condensed. The situation is similar to the detection of other visual features. For instance, color is detected by dedicated color detectors, which for example indicate the saturation of a color by a number. Likewise, a higher luminance is coded by a larger number than a lower luminance, and a motion detector indicates how strong motion is by a number. In some way, the motion detector can be seen as mini-model of our model.¹⁷ For example, when a line is moving, the line is first detected at one location and then at a second location. Because of a delay in neural latency (see Figure 4), the neural signals to both elements may arrive at the same time at a downstream neuron, which in this case elicits an action potential. If the arrival of the action potentials of the two neurons is not simultaneous, there is no action potential in the downstream neuron M. A bank of neurons with different delays can thus deliver a set of detectors, which code for different velocities and motion directions. Importantly for our considerations, a neural response is only elicited after the second element is presented. Hence, the neural delay functions as a short-term memory. Such motion detectors can also explain why we are sensitive to apparent motion or, to turn it this way, why we perceive motion where there is none. For such detectors, it does not matter whether or not there is motion between the detectors (see Figure 4). In analogy to the motion detector, we can code entire event structures in a similar way. For example, when a line is presented for 30ms and followed by a square that is moving rightwards for 100ms, we perceive the entire event not during this epoch and not for 130ms but only for a short moment and much later, since it is coded by a corresponding event

¹⁶ See Mudri et al. 2014.

¹⁷ See Lee 2014 and Figure 4.

detector activated far after the event. Likewise, in the mouse example, information could be held in an unconscious memory before the entire information is rendered conscious.

Such event detectors can settle the memory vs. perception debate between cinematic and retentional models. For example, a motion detector makes up a retention without the necessity to propose a second time axis. As mentioned, there is memory through the neural latency in the detector. However, this memory is not consciously accessible and, thus, only a strong motion percept is elicited. Inferences, to the contrary, work by explicit memories where both elements are consciously perceived as in the example of the trash bin above.

In philosophical terms, the above considerations support the view that external world timing is not represented by neural time, much like color is not coded by colored neurons.

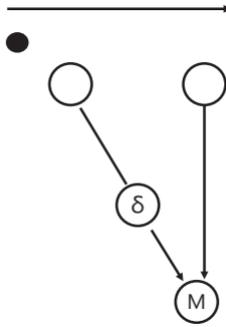


Figure 4.

A basic motion detector. When the dot moves from left to right, it first activates the left-most upper detector, which sends a signal to the downstream neuron M. The signal is delayed by a time δ . Then, the dot activates the right detector, which also sends a signal to M. When the two signals arrive at the same time at M, M also sends a signal indicating that there is left-right motion with a certain speed corresponding to δ .

4. Discussion

Dennett and Kingsbourne proposed that we cannot determine the microstructure of consciousness because we cannot discriminate

between two scenarios: whether consciousness at one very moment is immediately suppressed (Stalinian scenario) or whether consciousness is retroactively modified (Orwellian scenario).¹⁸ For example, in the SQM, one may well have a conscious percept of the first offset before the second offset is presented. This conscious percept does not enter memory and is suppressed when it is combined with the second offset. We cannot exclude the possibility of such fleeting consciousness. In this respect, our model only addresses consolidated consciousness, which is stable for at least a few seconds and could potentially be transferred into long-term memory.

One may argue that actions can be much faster than a few hundred milliseconds. While this is true, we argue that eliciting an action does not need consciousness as evident in reflexes and fast reactions in sports. We rather claim that actions are part of the event structures and are also perceived as such.

Further, one may argue that the existence of detectors coding for entire, complex events is highly unlikely, given the vast space of possible events. This may be true. Likely, neural coding is more complex than the outputs of simple detectors. However, this problem is not restricted to temporal events but holds true for static scenes as well. Humans have a strong capacity not only to distinguish many similar scenes, such as beaches, but can also memorize a large number of them. Thus, here is a general problem to solve.

One may also argue that the typical measure for neural responses are spikes per second, i.e., a measure extended in time. This is true. However, the timing of the spikes does not need to go hand in hand with the timing of the temporal events. However, as it is largely unknown what the neural code is, new insights may lead to different interpretations.

If perception is discrete, why do we not realize the gaps in between conscious events? This is simply because consciousness is, by definition, slower than the gaps. A system cannot go beyond its own resolution or, to turn it this way, we can only be conscious in and about the conscious events when we are conscious and not when we are unconscious. For this reason, conscious events follow each other smoothly in conscious time but not in physical time, which is consciously not accessible but likely unconsciously. In this respect, conscious time is shorter than physical time. However, we may still feel

¹⁸ See Dennett & Kingsbourne 1992.

an extended consciousness because unconscious duration detectors may determine exactly the time elapsed during unconscious processing and render this number conscious together with all the other information.

The surplus content argument was originally used by Dainton to argue against retentional models (Grush, 2016). However, as shown with the examples of the SQM and apparent motion, the surplus content argument argues not against retentions but against continuous consciousness including extensionalistic ones. Our discrete model avoids this problem because there is not a percept at each moment of time. If there were only two conscious percepts per second, there would definitely not too many percepts. Traditional discrete models have a similar problem because they are usually (implicitly) continuous: we are conscious at each moment of time, but percepts are constant during extended moments because a moment cannot contain any temporal information. We avoid this problem for the same reason as above: there are no continuous percepts and hence no constant percepts. The unconscious processing periods therefore solve the problems of both approaches in the same way.

There is a high price to pay: the model is at odds with how we feel our perception comes about. For example, in some instances, we argued that certain percepts are simply not what we believe they are, whereas, in other instances, we heavily relied on a certain percept being true. For example, we believe we see a continuous stream in the SQM, but our model claims this is an illusion. On the other hand, we claimed that there is no conscious percept of the first offset. Hence, one may prioritize the percepts this way or the other way around and arrive at different conclusions. An additional problem is that phenomenology becomes unreliable when we approach the temporal limits of conscious time. It is usually difficult to know what we perceive when time is short. This seems to be a general problem of the timing of consciousness. The future will show whether there are models that can marry our feelings about phenomenology with explanations for the puzzling results from temporal illusions.

Acknowledgments

I like to thank Lukas Vogelsang and Maëlan Menétray for comments on the manuscript. This work was supported by the Swiss National

Science Foundation grant ›Basics of visual processing: from elements to figures‹ (176153).

References

- Arstila, Valtteri (2015): Defense of the brain time view, in *Front. Psychol.* 6.
- Dainton, Barry (2018): Temporal consciousness, in *The Stanford Encyclopedia of Philosophy*.
- Dennett, Daniel & Kinsbourne, Marcel (1992): Time and the observer: the where and when of consciousness in the brain, in *Behav. Brain Sci.* 15, 183–201.
- Drissi-Daoudi, Leila, Doerig, Adrien, & Herzog, Michael (2019): Feature integration within discrete time windows, in *Nature Communications* 10(1), 1–8.
- Efron, Robert (1967): The duration of the present, in *Ann. N. Y. Acad. Sci* 138, 713–729.
- Grush, Rick (2016): On the temporal character of temporal experience, its scale non-invariance, and its small scale structure, Manuscript, doi:10.21224/P4WC73
- Herzog, Michael, Drissi-Daoudi, Leila & Doerig, Adrien (2020): All in Good Time: Long-Lasting Postdictive Effects Reveal Discrete Perception, in *TICS* 24, 826–837. <https://doi.org/10.1016/j.tics.2020.07.001>.
- Herzog, Michael, Kammer, Thomas & Scharnowski, Frank (2016): Time Slices: What Is the Duration of a Percept, in *PLoS Biology* 14(4), doi:10.1371/journal.pbio.1002433.
- Holcombe, Alex (2013): The temporal organization of perception, in Wagemans, J. (ed.), *Handbook of Perceptual Organization*, Oxford University Press, 820–840.
- Lee, Geoffrey (2014): Temporal experience and the temporal structure of experience, in *Philos. Impr.* 14, 1–21.
- Mudrik, Liad et al. (2014): Information integration without awareness, in *Trends Cogn. Sci.* 18, 488–496.
- Otto, Tom, Ögmen, Haluk & Herzog, Michael (2006): The flight path of the phoenix— The visible trace of invisible elements in human vision, in *Journal of Vision* 6, 1079–1086.
- Piper, Matthew (2019): Neurodynamics of time consciousness: an extensionalist explanation of apparent motion and the specious present via reentrant oscillatory multiplexing, in *Conscious. Cogn.* 73, 102751.
- Pöppel, Ernst (1997): A hierarchical model of temporal perception, in *Trends Cogn. Sci.* 1, 56–61.
- Westheimer, Gerald & McKee, Susan (1977): Perception of temporal order in adjacent visual stimuli, in *Vis. Res.* 17, 887–892.

Time

How to explain the experience of the passage of time?

1. Introduction

Philosophical questions regarding the subjective »experience of time« in general are puzzling.¹ It is not even certain that it is possible to experience time, since time is not a sensory object that can be experienced by one (or a combination of) sensory organ(s): time cannot be seen, smelled, heard, touched or tasted. While our experiences are temporal and it is possible to temporally experience objects, at least as being present, it is not clear that one can experience time itself.²

However, it seems unproblematic to claim that we experience some elementary experiences of temporality, such as the experiences of (i) duration, (ii) non-simultaneity, (iii) order, (iv) past-present or (v) change, according to Ernst Pöppel's classification.³ Still, even this classification and the definition of each of these elementary experiences are controversial. They have been subjects of vivid philosophical discussions for centuries. »Duration«, for instance, was intensely discussed by Augustine many centuries ago.⁴ »Simultaneity« and »order« also raised numerous debates, which have been rekindled by Daniel Dennett based on Benjamin Libet's famous experience on »backwards time referral«.⁵ Analyses on the so-called »specious present«⁶ continuously question the »past-present« relation. However, for the different reasons I will now focus on, I consider that the last experience (v) raises the most challenging difficulties: the experience

¹ For a helpful overview see e.g. Dainton 2001 and Le Poidevin 2019.

² See e.g. Husserl 1991.

³ See Pöppel 1978.

⁴ See Augustine 1961.

⁵ See Dennett 1991 and Libet 1993.

⁶ On the very notion of »specious present« see e.g. Le Poidevin 2019 and Arstila 2018.

of a temporal change, which includes the experience of the passage of time (E_{PT}) (if we trust Pöppel's previous classification).

What does the notion of »passage of time« mean? If we refer to our daily experiences and even to ordinary language (at least in English), we can admit that we spontaneously use many expressions that suggest that we happen to feel that time »passes,« at least more or less quickly. This is what expressions such »time flies,« »time went by too fast,« »time drags on,« »I found the time very long,« etc. testify. Without even referring to the idea of acceleration or deceleration, the ordinary notions of »flow« or »passage« as applied to »time« simply suggest the dynamic character of time. It remains to be seen to what extent those locutions are not mere metaphors.⁷

From a historical point of view, the so-called »retentionalists« and »extensionalists«⁸ models have already supported the view that we can experience *temporally extended* phenomena. Even if it is true, this is not sufficient to support the recent »passage of time« hypothesis. The new idea regarding this old view is not so much that our present experiences have a duration (that they are »extended«) but that we can subjectively experience the present as being *dynamic*, i.e., as »passing« (from the past and towards the future). The »passage of time« locution is indeed essentially used to refer to this feeling that what we experience as present just arises from the future and is shifting to become past. This is why this experience is not to be considered in the »past-present« but in the »temporal change« category, at least if I refer to the previous classification.

The vast majority of the E_{PT} examples to be found in recent literature are therefore examples of experiences of a temporal change based on the perceptual experiences of a motion. The most quoted example is the experience of the motion of the second-hand of a clock.⁹ Some prefer to refer to the perceptual experience of a falling

⁷ See e.g. Lakoff & Johnson 2003.

⁸ According to the »retentionalists,» the contents of our experiences are temporally extended but not the experiences themselves. By contrast, according to the »extensionalists,» the experiences are themselves temporally extended and structured by temporal parts (for this distinction, see Arstila 2018). As representatives of the retentionalist view, see e.g. Gallagher 2003 and Gallagher & Zahavi 2021. Gallagher & Zahavi are referring to Husserl 1991 and to Brentano 2009. Examples of extensionist positions can be found in Dainton 2000 or in Foster 1991 who refers to James 1890 and to Bergson 1910.

⁹ See e.g. Le Poidevin 2007, Dainton 2001 or Prosser 2016.

leaf (for reasons that I will later explain, see Section 4).¹⁰ The examples work as follows: when I watch the movement of the second-hand of a clock or the fall of a leaf, simultaneously with the experience of these motions, I experience in first person not several temporal instants t_1 , $t_2 \dots t_n$ but the dynamic passage from t_1 to t_2 to t_n : I feel that time is passing.

From an empirical point of view, as different as the clock or the leaf examples are from one another, they are all based on experiences of quick motions.¹¹ The motions that generate E_{PT} are indeed supposed to be relatively quick, for instance compared to the motion of the hour-hand of a clock or of a wilting leaf (Sattig's comparison) which, by contrast, do not provoke any feeling of the passage of time. Accordingly, E_{PT} would be somehow connected, in a way that I seek to define further, to direct non-inferential perceptual experiences of quick motion.

Now, the difficult point concerning the phenomenological view that we are experiencing temporal phenomena in a dynamic way—which provides us the feeling of the passage of time—is that it seems at odds with the metaphysical reality of time.¹² More exactly the »passage of time« question manifests a conflict between (i) a metaphysical thesis and (ii) a phenomenal one that seem to be contradictory: (i) time is not (really) passing, (ii) we (subjectively) experience the passage of time.

The purpose of this chapter is to show that not only are both claims true (as I will argue), they are also consistent. In order to clarify why the contradiction between (i) and (ii) is only apparent, I will review the current representationalist models¹³ that offer an explanatory model of E_{PT} . In Section 1, I will first focus on the first-order representationalist (FOR) strategy in order to show that it faces an

¹⁰ See Sattig 2018.

¹¹ There is no consensus, however, on the exact duration of the interval between the two beats of the hand-clock or two positions of the leaf. Some claim that it does not have to last more than a half-second (Dainton 2001), others no more than 2 seconds (Pöppel 1978). In any case, all claim that there is an empirical constraint.

¹² For a list of the very good arguments against the reality of time passing, see Prosser 2016.

¹³ Note that I will exclusively focus, in what follows, on representationalist models rather than on other influential empiricist models (such as the global workspace theory) because I am not entirely convinced that they can explain the phenomenality of our experience. More exactly, it would be a different agenda to show how the global workspace theory can explain the experience of time.

unsatisfactory alternative: either FOR accepts that the experience of time passing is a perceptual illusion or it accepts that it is not an experience at all (that it is a cognitive illusion). In Section 2, I will focus on the first option and argue that it is difficult to explain E_{PT} as a perceptual illusion, particularly if E_{PT} relies on the non-illusory perception of a movement. In Section 3, I will explain why E_{PT} cannot convincingly be considered a cognitive illusion, even if it is a powerful explanatory model. None of these objections are definitive. However, to avoid the reductivist and eliminativist implications of these first-order solutions, I will suggest theoretical alternatives. In particular, I will explain (Section 4) why I consider that higher-order representationalist (HOR) views on E_{PT} have some explanatory advantages over FOR. Eventually, I will focus (Section 5) on Giuliano Torrenzo 2017's view, who argues convincingly that E_{PT} should be seen as a *phenomenal modifier*.

2. E_{PT} as a Major Challenge for the Representational Theories of Consciousness

From a metaphysical point of view, there are good reasons to doubt that time is *really* passing. The *locus classicus* of the metaphysical discussion on the reality of the passage of time (and on the reality of time in general) is John McTaggart's distinction between the »A-theory« and the »B-theory« of time.¹⁴ According to the A-theory, only one time t is really present. All other times are either past or future, to some degrees (»distant past,« »near future,« etc.), compared to the real present. If you are an »A-theorist« (i.e. a »presentist«), you hold that time is continually *passing* from one present to another present and you can logically claim that there is indeed a »passage of time.« By contrast, if you are a »B-theorist« (i.e., an »eternalist«) you do not believe that there are any »A properties« (such as »being past« or »being present«). There are just »B properties« (such as »earlier,« »later«), which are relative to the context of your utterance. Then, a specific time, say t_1 , can be considered »present« but only relative to a specific context. Against that background, it does not make sense to claim that time is passing, for instance from the near future to the

¹⁴ See McTaggart 1908 and McTaggart 1909. Note that Taggart himself intended to show that the A-theory was self-contradictory.

present or from the present to the near past, since *time is not passing at all*.

Simon Prosser, among many others, has convincingly shown that there are good arguments in favour of the B-theory, which derive from progresses in physics (partly from the Growing Block Theory and from the Albert Einstein's 1905 Special Theory of Relativity, STR) as well as from empirical experiences (in particular on phi phenomena, see Section 2). Accordingly, I rely on the overwhelming majority of commentators to postulate, as a starting point, the plausibility of the B-Theory.

If it is true, one direct upshot of the B-theory is that time is not really passing. Hence, the passage of time is nothing real in the block universe and the so-called flow of time does not correspond to any physical fact. This is the (meta)physical claim on the passage of time.

The problem is that such a metaphysical theory seems to contradict directly the phenomenal claim that I introduced to begin with: from a first-person point of view we are frequently experiencing the passage of time, at least if we trust the numerous philosophers who believe that we do experience the passage of time¹⁵ as well as phenomenal reports.¹⁶

My first objective is to show that this apparent contradiction is indeed a challenging difficulty, at least for today's standard philosophical explanatory model of phenomenal consciousness: representationalism. In that respect, I will first investigate whether first-order representationalism (FOR) provides a satisfactory explanation of E_{PT} .

The FOR view, made popular by Gilbert Harman, Fred Dretske and Michael Tye in the 1990's,¹⁷ holds that any conscious experience, including its phenomenal character, can be entirely »exhausted« in its representational »content« (the view is also called the »content view«),¹⁸ provided that the (first order) representational content has some specific properties. For instance, according to Michael Tye's last

¹⁵ I already mention the so-called »retentionalist« and »extensionalist« philosophers. Prosser (2016) also provides a substantial list of philosophers who do think that there is an experience of the passage of time.

¹⁶ For an interesting survey on people subjective reports about the phenomenology of time passing, see Hoerl et al. 2021. According to the survey, a majority of the participants (although not all of them) reports that they already experienced the passage of time.

¹⁷ See Harman 1990, Tye 1995 and Dretske 1995.

¹⁸ See e.g. Brewer 2006.

model—called PANIC –, an experience has a phenomenal character in virtue of being about a representational content that is a Poised, Abstract, Non-conceptual and Intentional Content. The phenomenality of the experience would be entirely explained by the different properties of such a PANIC content.¹⁹ Accordingly, E_{PT} —as any other phenomenal experience—would be fully explained by some specific features of its representational content (i.e., that time is passing).

A first naive objection is to say that the experience of the passage of time has no representational content because (according to B-theory) there is no such thing as a worldly PT that could explain E_{PT} . This is of course not a definitive objection since FOR has more than one string to its bow. In particular, there is no reason for FOR to postulate that each of the properties of the representational contents are necessarily mundane properties: they can be non-mundane intentional properties. Consequently, E_{PT} could be entirely explained as an *illusory* experience, more exactly as an experience built on a *perceptual illusion*.

3. E_{PT} as a Perceptual Illusion

Accordingly, the most influential strategy, which is consistent with FOR and which is widespread nowadays, considers E_{PT} a perceptual illusion.²⁰ More exactly, the view holds that E_{PT} is grounded on the perceptual experience of a moving object (for instance, the second-hand of a clock), which is illusory. This perceptual experience would be illusory as it would provide the non-veridical representational content that there is an objective flow (that the second-hand is flowing) and then that there is a subjective flow (that the time itself is flowing) although there are neither an objective nor a subjective flow: the positions of the second-hand, just like the positions in time, are discrete.

Making the view plausible requires proving that it is at least possible to experience an objective or a subjective flow (or a passage) where there is none. The supporters of the view that E_{PT} is a perceptual illusion thus pay specific attention to empirical phenomena that can be used to prove the very possibility of an illusory perceptual experi-

¹⁹ Tye 2000.

²⁰ See e.g. Prosser 2007, Prosser 2016, Paul 2010, Dainton 2011 and Dainton 2012.

ence of a »passage.« In that respect, »phi phenomena« appear to be good candidates.

»Phi phenomenon« experiments have indeed shown that it is possible to create the illusion of an objective passage, for instance of the passage of a dot, in presenting nearby optical stimuli, which alternate with a rapid frequency. Interestingly, the stimuli that are provoked by the light flashes are static and their position does not change but the observers experience a dot as passing from a position to another.²¹ Even if the experiment is initially not meant to tell anything about the »passage of time,« it convincingly demonstrates how a dynamic experience (in that case: the experience of the passage of the dot) can be created by static stimuli. The experiment at least shows that our mind may create an illusion of an objective passage even if nothing is passing.

Another interesting example provides similar theoretical outcomes. It is famously known as »the waterfall illusion« (or the »motion after-effect«). This is a classical example of optical illusion, which was brought up to the current debates by Tim Crane,²² and which today raises numerous discussions among the philosophers of perception and of time.²³ The experience goes as follows: after watching the flow of the water in a waterfall, I look at a motionless scene, for instance at a rock close to the waterfall. I will then experience the rock as if it was flowing. Just as in the previous case, it is an illusion of flow.²⁴

Consequently, some experiments provide good reasons to think that it is at least possible that our consciousness fools us to the point that we believe that we perceive a passage where there is none. If we are to believe L.A. Paul, this idea is even supported by recent results from cognitive science:

²¹ See Wertheimer 1912 and Kolers 1976 for the empirical work. See e.g. Dennett 1991 or Paul 2010 for the philosophical investigations. Note that the study of »beta movements« presents similar theoretical conclusions (see also Wertheimer's »Experimentelle Studien«).

²² See Crane 1988.

²³ See e.g. Le Poidevin 2007, Mather 1998 and Arstila 2018.

²⁴ The »waterfall illusion« is also used to investigate the nature of the perceptual content of perception. For instance, it has been used to show that one can perceive contradictory states of affairs while knowing that there are no such states. Consequently, it has been used to support the view that the content of perception is non-conceptual (see Crane 1988).

The idea is that, just as the cognitive science suggests, the brain processes the series of inputs and produces a mental representation or experience as of O changing in some suitably animated or flowing way from being P into being Q. [...] Thus, according to the reductionist, there is no real flow or animation in changes that occur across time. Rather, a stage of one's brain creates the illusion of such flow, as the causal effect of prior stages on (this stage of) one's brain.²⁵

Then, it is at least formally possible to explain in the same way the experience of the passage of time as a perceptual illusion: we would be deceived the same way when we experience the passage of the dot and when we experience the passage of a second-hand (and then the passage of time). Consequently, it is theoretically possible that although it seems that time is passing, it does not pass at all.

Even if it is true, in order to build a complete explanatory model, one should also explain at least *how* the illusion of an objective (and then of a subjective) passage is generated. This is all the more important since the experience of the passage of time cannot be considered a localized and revisable error of judgement, but rather a *systematic* illusion, if we take for granted that (almost) everyone has already subjectively experienced the passage of time?²⁶

Here again, it may help to start from the explanation of the »phi motion« illusion. Prosser provides further clarifications on how the illusion of a passage could be explained based on the phenomenal properties of perceptual experiences.²⁷ Using David Velleman 2006's distinction between »enduring« and »perduring« properties,²⁸ Prosser postulates that an illusory experience of passage consists in misrepresenting a moving object (say a static dot) as enduring (to be present at each moment through change) instead of perduring (to have different temporal parts at different times). Hence, an illusory experience of passage would consist in attributing wrongly incompatible attributes (being present through temporal change) to a representational content.

By analogy, the subjective experience of the passage of time would consist in misrepresenting an object, say the second-hand of a clock, as »enduring« instead of »perduring.« Accordingly, the illusion of E_{PT} could be entirely explained within a FOR model based on

²⁵ Paul 2010, 352.

²⁶ See again Hoerl et al. 2021.

²⁷ See Prosser 2016: Chapter 6.

²⁸ Velleman 2006.

the analysis of the intentional content of the perceptual representation of an objective motion.

Such an explanation comes up against several objections, however. First, the explanation presupposes that the illusory experience rests on an illusory perception. But most of the perceptual experiences the experiences of the passage of time are based on are not themselves illusory.²⁹ Whereas in the phi motion case, the movement of the dot appears to be continuous (although it is not), the movement of the second-hand appears to be discontinuous. Strictly speaking, I do not have the illusion that the second-hand is flowing.

We can imagine, however, that an experience of time passing is sometimes based on a perceptual illusion. Let's consider for instance (Sattig's example, see Section 4), an experience of passage generated by the perceptual representation of a falling leaf. In this specific case, E_{PT} is indeed based on the illusory perception of a continuous movement (a flow). But, even in this case, how to explain E_{PT} by solely referring to the content of the perceptual experience of a motion? How could the perceptual content of an object enduring in time (the leaf) generate the content of a subjective temporal flow? I do not think FOR can answer this question.

As a first result, even if I agree that some of our E_{PT} are based on perceptual illusions, I believe that looking at the content of the perceptual experience and its properties alone (as FOR requires) does not explain why the perceptual content of a motion generates the illusion that the leaf, the second-hand and also time itself are flowing.

Interestingly, there is a radical alternative to FOR that avoids the delicate task of explaining how the conscious experience of some perceptual representations raises the illusory phenomenal reports that we experience time as passing. FOR can indeed radically postulate that the illusion of the passage of time is not phenomenal in nature. The passage of time would rest on a double illusion: the cognitive illusion that time passes and the illusion that I subjectively experience it as passing. In other words, the »passage of time« would be a mere cognitive illusion and there would be no E_{PT} at all.

²⁹ See also Hoerl 2014.

4. E_{PT} as a Cognitive Illusion

One may consider that E_{PT} is indeed an illusion, but a cognitive one.³⁰ According to this eliminativist view, we do not experience the passage of time at all. We just misinterpret our perceptual experiences of motion and wrongly infer from some perceptual content that we are experiencing time as passing. For instance, we wrongly infer from the perception of the motion of the second-hand the *belief* that we are indeed experiencing the passage of time. The illusion is to believe that we are even experiencing anything. To quote Christof Hoerl, it would just »seem« to us that we are experiencing the passage of time: there would be no phenomenology of the passage of time.

The rhetoric of this view is very close to the recent illusionist explanatory models on consciousness, which postulate that phenomenal experience (in general) is an illusion artificially generated by our phenomenal reports.³¹ According to Keith Frankish, for example, it just »seems« to us that we experience phenomenal experiences although the very idea of a phenomenal experience is illusory. Hoerl's position is of course much less radical in that it is localized. It is not aimed at experience in general but just at the quasi-phenomenal experience (to use the illusionist terminology) of the passage of time. One can however postulate that Hoerl's proposal presents some of the virtues and limitations of the illusionist position in general.³²

In particular, this eliminativist model presents the great advantage of being parsimonious and consistent with the naturalistic requirements. It does not postulate anything beyond perceptual experiences and non-phenomenal cognitive inferences. Accordingly, this explanatory model is perfectly consistent with FOR: it explains entirely the illusion of E_{PT} by the representational content of our beliefs inferred from perceptual representations. The sole task of supporters of such a view would be to explain the inferential errors at work, which seems less »hard« (to use Chalmers' terminology) as a task than explaining the spontaneous genesis of a phenomenal experience.

Even if I do not exclude that the view is correct, I think it is not entirely satisfactory to claim that E_{PT} is a cognitive illusion. First, even

³⁰ See e.g. Braddon-Mitchell 2014 and Hoerl 2014.

³¹ See e.g. Frankish 2016.

³² On the virtues and limitations of illusionism about phenomenal consciousness, see Gauvry 2023.

if the view is able to explain the so-called cognitive illusional inferences (and I guess it can),³³ I am not sure it would be enough for the view to be convincing. A convincing explanatory model of the E_{PT} illusion must not only explain the cognitive inferences that generate the belief (that we are experiencing the passage of time) but also the reasons why the illusion seems to be subjective: why it »seems« to us that we are *experiencing* time passing although we do not experience it. In other words, the explanatory model must both explain why we believe that time is passing and why we believe that we experience time as passing. I am not sure, however, that the advocates of the »cognitive illusion« hypothesis, who cannot refer, by definition, to the evidence of the experience itself, can offer such an explanation.

Second, even if the passage of time is just the intentional content of a false belief, it also has to be proven that this false belief is not itself phenomenal. As the literature made obvious, cognitive phenomenology is a plausible option and it is not unlikely that there is a »what it is like« to (wrongly) think that time is passing.³⁴ Even if E_{PT} is a cognitive illusion, there would still be a (cognitive) experience of the passage of time.

For all that, I think that the eliminativist strategy on E_{PT} is intriguing and that everything is indeed much simpler and less contradictory if there is no E_{PT} . I even believe that anyone without phenomenological intuition about the experience of the passage of time should favor this model. On the other hand, I also consider from a first person perspective that it is not entirely satisfactory to just deny the reality of the phenomenal experience of the passage of time. This is the reason why I will now seek to find out alternative views to the standard first-order content view that could account for the phenomenology of the passage of time.

³³ Hoerl, for instance, has investigated on the role of »episodic memory« to clarify the status of the wrong inferences. See e.g., Hoerl 2014.

³⁴ See e.g. Chudnoff 2015 and Bayne 2017. However, representatives of the idea of cognitive illusion would be entitled to answer, just like Frankish, that: »it is essential to [their] approach that the posited introspective [mis-]representations are not themselves phenomenally conscious ones« (Frankish 2016, 15).

5. A Higher-order Account of E_{PT} ?

As explained above, if the first-order representationalist theories are generally considered to have a strong explanatory power for our phenomenal experiences, I believe that they cannot explain the phenomenology of E_{PT} . At least, I have shown that it is difficult to offer a reductive reading of E_{PT} by referring to a single perceptual content. The difficulty is to explain within a FOR framework how the (non-illusory) representation of some properties necessarily generates the illusion that we are experiencing other properties. In particular, it is not clear how the sole content of a perceptual representation (say, that the second-hand is perduring through change) can exhaust another experience (E_{PT}), which seems to have its proper content (that the time is passing). Consequently, I think that the sole explanatory option within a FOR program is to present E_{PT} as a cognitive illusion. But, as I just suggested, this eliminativist proposal also presents some limitations from a phenomenal and explanatory perspective.

Fortunately, there are representationalist alternatives to FOR and to eliminativism. A first one is offered by the higher-order representationalist (HOR) models. According to HOR, a conscious experience cannot be exhausted into one but into *two* contents of representation that are not necessarily identical. Contrary to FOR, HOR postulates that an experience is subjectively conscious by virtue of a higher-order representation directed towards a first-order representation. Then, HOR views present the advantage to attribute a different content to first-order and higher-order representations which proves beneficial in explaining an illusion.

Against that background, it is possible to consider the representation of the passage of time a higher-order representation directed towards a first-order perceptual representation. More exactly, E_{PT} would be conscious by virtue of a third-order representation of the second-order representation of the passage of time based on the first-order perception of a movement. This is for instance Sattig 2019's proposal. I use his example as an illustration.³⁵ Let's imagine the perception of a leaf falling from a tree. Such perception is directed towards the properties of a perceptual content: I see the colour, the shape, the trajectory and the movement of the leaf. As previously mentioned, when I pay attention to its movement, I do not perceive a succession

³⁵ See Sattig 2019.

of different discontinuous positions but a continuous movement of fall. The perception of this continuity can be explained (Sattig's proposal) by the introduction of the concept of »objective replacement:« what I perceive is a leaf that is continuously replaced by itself. Now, it is clear that the sole notion of »objective replacement,« which is a property of the first-order perceptual content, is not sufficient to explain the non-objective »replacement« of time. The interesting idea of HOR, however, is that we do not have to assume that the perceptual representation of the leaf and the representation of time passing share the same content. As the second-order representation is not essentially linked to the first-order representational content of a moving object (the leaf), it can be a purely subjective representation of a »subjective replacement,« decorrelated from any specific object, which seems satisfactory to explain the passage of time. Accordingly, the representation of time passing would be a second-order representation of a (non-objective) replacement and E_{PT} would be constituted by a third-order representation of the second-order representation of this subjective replacement.

I think that this explanatory model is satisfactory for various reasons. The HOR view presents the advantage that it explains the singularity of the E_{PT} content without being eliminativist or reductionist: it does not even have to reconduct the content of E_{PT} to some first-order perceptual properties. One of the main virtues of the view is also that it can explain why it is theoretically possible that E_{PT} , although a (higher-order) experience, is illusory, i.e. why its content does not coincide with the perceptual content of the falling leaf. Finally, postulating that E_{PT} ultimately (but not reductively) rests on the representation of certain formal properties (such as »replacement«) of a first-order representational content is nothing else than to provide an explanation of the genesis of E_{PT} .

I am not convinced that the explanation is complete, however. In particular, the view does not entirely explain how the representation of the passage of time simply arises from the higher-order representation of a subjective replacement. In addition, it is not clear how systematic the explanation could be. Even if one considers that the model works quite well for describing the motion of the falling leaf, not all E_{PT} are based on the perceptual representation of such replacement. For example, as explained in Section 2, it is dubious that I perceive any continuous replacement when I perceive the second-hand of a clock.

More generally, one may wonder whether it is really possible to explain a phenomenal experience by referring only to the different properties of its (first, second and third) representational *contents* (whether these contents are the contents of perceptions or of thoughts). This is the reason why, as a last hypothesis, I will now focus on a non-standard view that claims that the properties of E_{PT} cannot be explained as properties of our representational contents but as properties of our intentional *attitudes*.³⁶

6. E_{PT} as a Phenomenal Modifier

The previous analyzes aimed to show the virtues but also some of the limits of the most common versions of different representationalist models on E_{PT} . I will now turn to an alternative proposal that considers E_{PT} to be a »phenomenal modifier«. This is Giuliano Torrenzo's position.

For the sake of his demonstration, Torrenzo himself is first referring to what he considers the most »sophisticated« version of representationalism (as a content view),³⁷ which explains the phenomenal character of E_{PT} by the »tensed properties« of some perceptual contents.³⁸ According to Torrenzo, however, even if our perceptual contents were constituted by such »tensed properties«, it is not clear that they could explain the *dynamic* features of E_{PT} . One should at least introduce a distinction between two meanings of »tense:« a *perspectival* one and a *dynamic* one. Although it seems possible to provide a representational account of the perspectival properties of an experience (for instance based on the semantic uses of some indexicals), it is very unlikely that the dynamic features of E_{PT} could be exhausted in a »tensed« content. Torrenzo even offers an interesting argument by comparing the perspectival representation of time and of space. Just as it is not because we have a perspectival representation of space, for example by using an indexical (e.g., »this« place), that we think that space is dynamic, it is not because we have a perspectival repre-

³⁶ Even if the next view is non-standard, it is closed to the adverbialist proposal. For a review and defense of adverbialist theory, see Kriegel 2011.

³⁷ See Torrenzo 2018.

³⁸ The influential debates on »indexicality« have indeed convincingly shown that our representations have a »tensed format«. See e.g., Perry 1979, Mellor 1998 and Torre 2009.

sensation of time (in using »today,« »tomorrow,« etc.) that we infer that time is somehow dynamic. Consequently, at best, sophisticated representationalism can grasp the perspectival dimension of E_{PT} in terms of »tensed properties« of a »tensed content,« but not the dynamic feeling that constitutes »time passing.«

Torrenço's conclusion is that no content view, however sophisticated, is capable of explaining the dynamic features of E_{PT} in terms of representational content. As an alternative model, he proposes that the dynamic features of E_{PT} are not to be found in the properties of its representational content but in the dynamic way our perceptual contents are represented. Although there are no dynamic contents« there would be »dynamic attitudes.«

Such a view is very close to the one presented under the label »attitudinalism« by Uriah Kriegel.³⁹ Kriegel indeed believes that what he calls the »temporal orientation« of an experience is explainable as an *attitude* towards a representational content. For instance, the »orientation towards the present« is a constitutive feature of our attitudes towards *perceptual* content. By contrast, the »orientation towards the past« is a constitutive feature of our attitudes towards the contents of our *memories*. Consequently, an act of perception distinguishes itself from an act of memory or imagination by virtue of its attitude whose definitional feature is precisely the temporal orientation.

One of the most interesting assumptions of this proposal is that our attitudes have an essential link with our representational contents. The next step is to believe that our attitudes condition and potentially modify our representational contents. This is precisely Torrenço's strategy. According to his 2017 proposal, the dynamic characters of E_{PT} are indeed the constitutive elements of some of our phenomenal attitudes that modify our perceptual contents and the way we experience them. This is what he calls the »phenomenal modifier view.«

In order to both explain some limitations of standard representationalism and to provide an illustration of a »phenomenal modification« of a representational content Torrenço first focuses on the interesting example of »blurred vision,« using Boghossian and Velleman's helpful description.⁴⁰ If we follow Boghossian and Velleman, in a case of blurred vision, the objects are not represented »as being

³⁹ See Kriegel 2015.

⁴⁰ See Boghossian & Velleman 1989.

blurry.« I experience them »blurrily« and I potentially believe that they are »blurry.« Accordingly »blurriness« is not a property of a representational content but a feature of an attitude that modifies the way I experience some representational contents.

The »blurriness« case can be used by analogy to explain the passage of time experience, at least a deceleration of the passage of time. Let's imagine that time has suddenly slowed down facing a dangerous animal, say a ferocious tiger.⁴¹ In this case, even though time has not really slowed down, this is the way I experience the situation. While some of the temporal and spatial perspectival properties of my representation of how the tiger moves—the fact that the tiger is here now—may contribute to explain some other features of my experience, they will certainly not explain my feeling that time slows down. Conversely, Torrenco suggests that it is possible to explain my representation of the tiger (for instance as being in slow motion) by my attitude of which E_{PT} is a constitutive character. In other words, perceptual contents do not generate E_{PT} ; on the contrary, E_{PT} modifies perceptual contents. Hence, when I experience the passage of time, the dynamic aspect of the felt passage of time cannot be explained as a property, not even as a tensed property, of any representational content but as a feature of my phenomenal attitudes towards some perceptual contents.

Let me mention another example to tentatively corroborate this explanation: boredom. Imagine that I am bored and that I find that time drags on or »seems long.«⁴² When I am deeply bored, no matter what representational contents I experience (interestingly, even if they become objectively more exciting), I will boringly experience them. The same could probably be true of other moods.⁴³ If I am in a happy mood, for instance, it is likely that I will feel like that time flies and it will also modify the contents of my representations. Consequently, it seems possible to describe those experiences (and the cor-

⁴¹ The example is quoted by Torrenco 2017.

⁴² One could of course wonder if the »boredom« example is an example of »time passing.«. Some people consider this is a case of interminable »duration.« Even if it is the case, it makes sense to say that the experience of extended duration can be considered, by contrast, as an illustration that we normally experience the time as passing, since what happens in boredom is precisely that time no longer passes.

⁴³ Note that moods are borderline cases from a representational point of view because their intentional character is controversial. See Searle 1983.

related experience of time passing) in terms of attitudes that phenomenally modify our representational contents.

My point is that this »phenomenal modifier view« presents all the advantages of the precedent views without their limitations. First, just as the eliminatist model, the view is parsimonious as it does not even pretend that E_{PT} has its own content. Strictly speaking, the experience of the passage of time would not be the experience »that times is passing« but the »time(ly) passing(ly)« representation of some perceptual content. Just as a hermit crab, the experience of the passage of time would appropriate the content of other intentional experiences to modify it.

Second, as the HOR proposal, the »phenomenal modifier view« is well-equipped to explain why the experience of the passage of time does not correspond to the reality of time. If E_{PT} does not have any content, the mismatch between its hypothetical content and the content of first-order representations is indeed dissolved.

It is true, however, that the »phenomenal modifier view« raises a new objection: if E_{PT} has no content of its own it is no longer possible to refer to any content to *explain* E_{PT} . One could then say that it becomes very difficult to offer an explanation of E_{PT} within the »phenomenal modifier« model.⁴⁴ According to such a model, one could at best use E_{PT} as an *explanans* to explain the modification of our perceptual contents but E_{PT} itself would not be explained.

As a partial answer to the objection, it can be argued that the model is compatible with a certain type of naturalistic explanation: evolutionary explanations. Torrenço quotes for instance Ian Phillips 2013's explanation of the evolutionary reasons why the brain generates E_{PT} :

Imagine that you are a caveman or—woman on the veldt. Scanning the horizon, you spot a saber-toothed tiger heading your way. Then suddenly the world around you seems to slow down and the tiger appears to be running more slowly. How is this helpful? The tiger is not actually running any more slowly. And the illusion of time being drawn out gives you no extra seconds in which to flee. We get a much more satisfactory explanation of what is going on if we consider how things look on a mental activity picture. Here the effect of the fear-based dopamine

⁴⁴ See Kriegel 2011.

spike is to speed mental activity. That, in and of itself, is an adaptive response.⁴⁵

Hence, there would be an explanation of the experience of time slowing down in terms of »adaptative response.« Interestingly, this adaptative response can be described as an experience that modifies, just as »the fear-based dopamine,« the perceptual content of the movement of the saber-tooth tiger in order to »speed our mental activity.« Accordingly, the experience of the slowing down of time, i.e. an experience of a temporal flow, could be explained, on the basis of adaptative explanations, as a phenomenal modifier. Thus, the »phenomenal modifier« model, coupled with an evolutionary explanation, would make it possible both to explain the illusory nature of the experience of passing time (in terms of modification of external contents) and its systematic nature (in evolutionary terms).

I do not believe, however, that this model is not problematic at all. In particular, I am not sure that the »phenomenal modifier view« is valid for explaining all the E_{PT} cases. However, I hold that it is today the most convincing explanatory model that is consistent with our phenomenal reports.

7. Conclusion

The ambition of this chapter was both to offer a review of the representationalist models that are available today to explain the experience of the passage of time and to propose a non-eliminativist model in order to explain the phenomenology of E_{PT} . From this review, I draw two conclusions.

The eliminativist model, which postulates that there is simply no E_{PT} at all and that E_{PT} is a cognitive illusion is a parsimonious model with great explanatory power. Consequently, I suggest to anyone who does not have a phenomenal intuition about E_{PT} and who does not believe that the eliminativist model contradicts her own experience to favor it without hesitation. On the other hand, I believe that the only model that is compatible with our phenomenal reports on the passage of time and that also has real explanatory power is the »phenomenal modifier« model.

⁴⁵ Phillips 2013, 246 (my emphasis).

References

- Addams, Robert (1834): An account of a peculiar optical phenomenon seen after having looked at a moving body, in *London and Edinburgh Philosophical Magazine and Journal of Science* 1834/5, 373–374.
- Armstrong, David (1968): *A Materialist Theory of the Mind*, London: Routledge & Kegan Paul.
- Arstila, Valterri (2018): Temporal Experiences without the Specious Present, in *Australasian Journal of Philosophy* 96/2, 287–302.
- Augustine (1961): *Confessions*, Harmondsworth: Penguin.
- Bergson, Henri (1910): *Time and Free Will: An Essay on the Immediate Data of Consciousness*, New York: Humanities Press.
- Boghossian, Paul & Velleman, J. David (1989): Colour as Secondary Quality, in *Mind* 389, 81–103.
- Braddon-Mitchell, David (2014): Against the illusion theory of temporal phenomenology, in Sato, T., Sugimoto, S. & Sakon, T. (eds.), *CAPE Studies in Applied Philosophy and Ethics Series 2*, 211–222.
- Brentano, Franz (2009): *Philosophical Investigations on Space, Time and the Continuum*, London: Routledge.
- Brewer, Bill (2006): Perception and Content, in *European Journal of Philosophy* 14/2, 165–181.
- Carruthers, Peter (2005): *Consciousness. Essays from a higher-order perspective*, Oxford: Oxford University Press.
- Crane, Tim (1988): The Waterfall Illusion, in *Analysis* 48/3, 142–147.
- Dainton, Barry (2000): *Stream of Consciousness. Unity and Continuity in Conscious Experience*, London: Routledge.
- Dainton, Barry (2001): *Time and Space*, Chesham: McGill-Queen's University Press.
- Dainton, Barry (2011): Time, Passage and Immediate Experience, in Callender, C. (ed.): *The Oxford Handbook of Philosophy of Time*, Oxford: Oxford University Press, 382–419.
- Dainton, Barry (2012): Time and Temporal Experience, in Bardon, A. (ed.): *The Future of the Philosophy of Time*, New York: Routledge, 123–148.
- Dainton, Barry (2018): Temporal Consciousness, in *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition). URL: <<https://plato.stanford.edu/archives/win2018/entries/consciousness-temporal/>>.
- Dennett, Daniel (1991): *Consciousness Explained*, Boston: Little, Brown & Company.
- Dretske, Fred (1995): *Naturalizing the Mind*, Cambridge, Mass.: MIT Press.
- Einstein, Alfred (1905): On the electrodynamics of moving bodies, in *Annalen der Physik* 17, 891–921.
- Foster, John (1991): *The Immaterial Self*, London: Routledge.
- Frankish, Keith (2016): Illusionism as a Theory of Consciousness, in *Journal of Consciousness Studies* 23, 11–39.
- Gallagher, Shaun (2003): Sync-ing in the Stream of Experience: Time-Consciousness in Broad, Husserl, and Dainton, in *Psyche* 9/10.

- Gallagher, Shaun & Zahavi, Dan (2021): *The Phenomenological Mind*, London: Routledge.
- Gauvry, Charlotte (2023): Phenomenal reports as a fiction pump. What if illusionism is partly right, in Gabriel, M., Gymnich & Münch, B. U. (Hg.): *Wirklichkeit/Fiktion. Spannungsfelder eines Verhältnisses*. Tübingen: Mohr (forthcoming).
- Harman, Gilbert (1990): The Intrinsic Quality of Experience, in *Philosophical Perspectives* 4, 31–52.
- Hoerl, Christof (2014): Do we (seem to) perceive passage?, in *Philosophical Explorations* 17, 188–202.
- Hoerl, Christof (2014): Time and the Domain of Consciousness, in *Annals of the New York Academy of Sciences* 1326, 90–96.
- Hoerl, Christof, Shardlow, Jack, Lee, Ruth; McCormack, Teresa, Burns, Patrick & Fernandez Alison S. (2021): Exploring peoples's beliefs about the experience of time, in *Synthese* 198, 10709–10731.
- Husserl, Edmund (1991): *On the Phenomenology of the Consciousness of Internal Time (1893–1918)*, Dordrecht: Kluwer.
- James, William (1890): *The Principles of Psychology*, New York: Public Library.
- Kolers, Paul & von Grünau, Michael (1976): Shape and Color in Apparent Motion, in *Vision Research* 16, 329–35.
- Kriegel, Uriah (2011): *The Sources of Intentionality*, New York: Oxford University Press.
- Kriegel, Uriah (2015): Experiencing the Present, in *Analysis* 76, 407–413.
- Lakoff, George & Johnson, Mark: *The Metaphors we live by*, Chicago: University of Chicago Press.
- Le Poidevin, Robin (2007): *The Images of Time*, Oxford: Oxford University Press.
- Le Poidevin, Robin (2019): The Experience and Perception of Time, in *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition). URL: <<https://plato.stanford.edu/archives/sum2019/entries/time-experience/>>.
- Libet, Benjamin (1993): The Neural Time. Factor in Perception, Volition and Free Will, in *Neurophysiology of Consciousness*, 367–384.
- Lycan, William (1996): *Consciousness and Experience*. Cambridge, Mass.: MIT Press.
- Mather, George, Verstraten, Frans A. J. & Anstis, Stuart (1998): The motion aftereffect: A modern perspective, in *Trends in Cognitive Sciences* 2/3, 111–117.
- McTaggart, John Ellis (1908): The Unreality of Time, in *Mind* 17, 457–474.
- McTaggart, John Ellis (1909): The Relation between Time and Eternity, in *Mind* 18, 343–362.
- Mellor, David Hugh (1998): *Real Time II*, London: Routledge.
- Paul, Laurie A. (2010): Temporal Experience, in *Journal of Philosophy* 107, 333–359.
- Perry, John (1979): The Problem of the Essential Indexical, in *Noûs* 13, 3–21.
- Phillips, Ian (2013): Perceiving the Passing of Time, in *Proceedings of the Aristotelian Society* 113, 225–246.
- Pöppel, Ernst (1978): Time Perception, in Held, R. et al. (eds.): *Handbook of Sensory Physiology*, Vol. VIII, Berlin: Springer.

- Prosser, Simon (2007): Could We Experience the Passage of Time?, in *Ratio* 20, 75–90.
- Prosser, Simon (2016): *Experiencing Time*. Oxford: Oxford University Press.
- Rosenthal, David (2002): Explaining Consciousness, in Chalmers, D. (ed.): *Philosophy of Mind. Classical and Contemporary Readings*, Oxford: Oxford University Press, 109–131.
- Sattig, Thomas (2019): The Sense of Temporal Flow. A Higher-Order Account, in *Philosophical Studies* 176, 3041–3059.
- Searle, John (1983): *Intentionality: An Essay in the Philosophy of Mind*, New York: Cambridge University Press.
- Torre, Stephan (2009): Tense, Timely Action and Self-Ascription, in *Philosophy and Phenomenological Research* 80, 112–132.
- Torrenzo, Giuliano (2017): Feeling the passing of time, in *Journal of Philosophy* 114, 165–188.
- Torrenzo, Giuliano (2018): Perspectival Tenses and Dynamic Tenses, in *Erkenntnis* 83, 1045–1061.
- Tye, Michael (1995): *Ten Problems of Consciousness*, Cambridge, Mass.: MIT Press.
- Tye, Michael (2000): *Consciousness, Color and Content*, Cambridge, Mass.: MIT Press.
- Velleman, J. David (2006): So it Goes, in *Amherst Lecture in Philosophy*, <http://www.amherstlecture.org/velleman2006/velleman2006_ALP.pdf>.
- Wertheimer, Max (1912): Experimentelle Studien über das Sehen von Bewegung, in *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* 61, 161–265.

Christoph Horn

Husserl's Phenomenology of the Inner Time Consciousness: Achievements and Limits

Making our subjective experience of time a research topic is an important challenge both for psychologists and philosophers. Nevertheless, the interdisciplinary conversation between the two widely separated academic subjects is not so easy. This is because there are some significant differences in the way inner time consciousness is made a topic of study on the part of psychology and philosophy, respectively. The differences concern both how the phenomenon is described and how it is related to the physical time of the external world. Indeed, the ambiguity of the concept of time—the distinction between the ›physical time‹ of natural processes and the ›internal time‹ of our experience of time—forms a central problem of any systematic discussion of the concept of time. It was the philosopher John M.E. McTaggart who brought this difference to the point with his well-known differentiation between the so-called ›A-series‹ (the subjective conception of time) and the so-called ›B-series‹ (the objective notion of time). In McTaggart's view, our double concept of time is irrevocably contradictory: For while the description of time from the perspective of the A-series is indexical and operates with the terms ›earlier than,‹ ›simultaneously,‹ and ›later than,‹ time from the perspective of the B-series appears from an objective distance and thus seems irrelevant for the truth value of events.

Now, in the philosophical tradition there are fundamentally different approaches to how the relation between objective external time and subjective internal time should be described. Two of them are of particular interest. First, one can advocate an empiricism or physicalism that makes external time the factual reference point for all analyses of the internal experience of time. Such a position has been held historically since Aristotle, who takes the real motion of the cosmos as his theoretical starting point; time is determined in Aristotle as ›the

measure or number of motion according to the earlier and later«. ¹ In the line of empiricism and physicalism, internal experiences of succession are understood as caused by external events; the internal is explained from the perspective of the external. On the other hand, there exists a type of view that interprets time in general as an activity of our consciousness, and in so doing also conceives of external time as subjectively constituted. This position is based on the subject philosophy of modern times since Descartes, but is often (not quite correctly, as we will see) already traced back to the later ancient philosopher Augustine. Augustine defines time as »expansion of the soul« (*distentio animi*). ² In a philosophy of subjectivity, the external is explained as constituted by the internal.

In what follows, we are concerned with Edmund Husserl's (1859–1938) conception of time, which is usually assigned to the second tradition. In fact, it is correct to claim that since his book *Ideas* ³ he advocated a version of transcendental philosophy that sees time, and the entire world of experience in general, as constituted by a ›pure Ego‹. But strictly speaking, Husserl's analyses of time are independent of this transcendental turn in his thought. His form of analysis, for which he chose the term ›phenomenology,‹ basically formulates rather an intermediate position or bridge between physicalism and subject philosophy. I therefore begin with a brief account of Husserl's concept of phenomenology (I), then move on to his reference to Augustine (II), and finally attempt a survey of his overall position with its strengths and weaknesses (III).

I.

Husserl's theory of inner time-consciousness is a core component of his phenomenology as a whole. ⁴ This is because Husserl is fundamentally concerned with the description of our experience from the inner perspective of our experience—in which the the stream of consciousness naturally plays a prominent role. By ›phenomenology‹ Husserl meant a philosophical method of unprejudiced, accurate

¹ Aristotle 1957, IV 219b1–2.

² Augustine 2016, XI.23.30 and XI.26.33.

³ *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie* (Husserl 1913).

⁴ Römer 2010, 18.

description of mental processes or experiences of consciousness. ›Unprejudiced‹ and ›factually adequate‹ imply following Husserl that no theories which are not legitimized by the phenomena should be included in the respective description and no unjustified assumptions of existence should be made. Especially important in the phenomenological description of the mental is the object-relatedness of numerous ›acts‹: their intentionality. The crucial question here is how we are related in our consciousness to a world meant to be ›objective‹. With regard to the phenomenon of time, some relevant phenomenological questions are these: How can we relate to one and the same object across time? How do we experience ourselves in time? How is an awareness of processes, i.e. temporally structured events, possible? Are we ourselves in time, or do we constitute time?

For Husserl, the starting point of phenomenology lies in the change of attitude of the researcher: he or she should refrain from accepting the external world of objects and events as well as the internal world of thoughts, ideas or emotions as given and should instead describe the way of their being given. Husserl calls this retreat from the contents (Husserl speaks of ›bracketing‹: *Einklammerung*) to the contemplation of their mode of appearance *epochê*, using an expression from ancient skepticism. Instead of being absorbed in the given phenomena, in *epochê* one considers the phenomena without preconceptions. Thus, according to Husserl, it is not the real existence of this or that object that is the topic of phenomenology, but only the intentional act directed towards it, in which the object appears. While empirical psychology presupposes the external world as real and thus explains phenomena of consciousness psychophysically, phenomenology leaves it at reproducing the claim to validity of an appearing phenomenon, that is, at describing the reference of an act of consciousness (*noesis*) to an object (*noema*). Husserl calls this ›phenomenological reduction‹. In a further step, the ›eidetic reduction‹, he wants to filter out certain invariant structures from the innumerable observations resulting from the *epochê*, which he characterizes as ›laws of essence‹ (*Wesensgesetze*).

From this sketch of the program of phenomenology, it is clear that Husserl is an anti-representationalist; he denies that the perception of spatiotemporal things is mediated by images or ideas. Rather, using the tools of phenomenology, the ›sense of direction‹ of perceptions can be described as directly object-related. Empirical objects and events (e.g., I see a car passing me) are accessible to me in immediate

presence (›self-givenness‹), not as a synthesis of singular impressions. Already in his early days Husserl—very much like Gottlob Frege at about the same time—developed a critique of contemporary psychologism. As both thinkers claim, logical laws are not reducible to empirical regularities of thought as identified by psychology.

The previous characterization of Husserl's approach referred to it as ›pure phenomenology‹. In his later ›transcendental phenomenology‹ (beginning with the *Ideas* in 1913) that builds on it, Husserl's perspective shifts from describing the givenness of contents of consciousness to the constitution of the contents of consciousness by a pure Ego. Husserl now believes that the phenomenological and the eidetic reduction are leading to the conclusion that both *noesis* and *noema* and their regularities must go back to a unified principle of constitution. This, for Husserl, is the transcendental Ego, a resting pole which he understands as a completely content-less, spontaneous, a priori I. Husserl thus belongs to the transcendental tradition in philosophy established by Descartes and rich in egological considerations in Kant and especially Fichte. His transcendental idealism is based on the idea that we must describe reality indirectly—on the basis of its constitution in subjectivity.

Hence, the crucial point of Husserl's analysis of time is the idea that one must exclude all assumptions about the real (consciousness-transcendent) existence of temporal processes. Thus, the phenomenology of the inner experience of time does not assume the real duration of real things, but merely addresses their »appearing duration.« In this way, the phenomenological approach is limited to the perceived time as something given by an ›original time consciousness‹. Husserl addressed the phenomenology of inner time consciousness in three phases of his work.⁵ First, there are his lectures from 1905–11, which Heidegger published in 1928 under the title *Vorlesungen zum inneren Zeitbewusstsein*. Second, there are the so-called *Bernau Manuscripts on Time Consciousness* from 1917–18 (published in 2001); Husserl seems to have attached particular importance to them. Finally, thirdly, we possess some hitherto unpublished manuscripts from 1930–34. That there exist unpublished manuscripts should, of course, make us cautious. On the other hand, the texts of the first two phases are basically in continuity with each other. However, the *Analyses on Passive*

⁵ Cf. Römer 2010, 5–7 and Seel 2010, 43–44.

Synthesis (1918–26) and the *Cartesian Meditations* (1931) are also important.⁶

Husserl's starting point is not based on the contrast of ›subjective‹ and ›objective‹ time; it does not consist in an attempt to interpret the A-series against the background of the B-series. Rather, Husserl believes that one must start from the subjective appearance of time in order to correctly grasp the time phenomenon as a whole (including the B-series). As seen in Husserl's successor positions (Heidegger, Sartre, Merleau-Ponty, and Ricoeur), his real achievement has been seen in understanding the objective time of Newtonian physics as a derived, theoretical construct by science. On the other hand, a critique also refers precisely to the fact that Husserl—ignoring Einstein's theory of relativity—stops at Newton's absolute world time.⁷

Husserl was interested in how we perceive physical or objective time on the one hand and how inner experiences of time take place on the other. Inner time is a rather neglected topic in the history of philosophy; for Husserl it is therefore the merit of Augustine to have drawn attention to inner time and its constituting role.⁸

II.

There is a historical role model for Husserl's phenomenology of inner time consciousness, namely Augustine, *Confessions* XI.14 – 28. The interesting point in Augustine's theory of time is that he was the first who described it from a subjective perspective. Augustine ›psychologizes‹ time. In doing so, he starts from various paradoxes to which he offers solutions throughout the text. The paradoxes developed by Augustine are:

- (a) How can a now-point which is constantly experienced by a subject be always another one in the very next moment? What exactly is it that passes when a now-point vanishes, and what is it that remains? How can one speak of the present as something existing, if it passes away almost immediately? So how can there be time and why isn't there eternity instead?

⁶ Orig. *Analysen zur passiven Synthesis* (1918–26) und *Cartesianische Meditationen* (1931).

⁷ Cf. Römer 2010, 54.

⁸ On this in more detail: Herrmann 1992.

- (b) How can a given now-point pass into the past, and how can future become the present of a given now-point? How can thereby something existing, the present, come from something non-existing, the future, and immediately pass again into something non-existing, the past?
- (c) If only the present now is available to us, how can we know about the past and the present?
- (d) Since we can obviously speak of so-and-so long periods of time, does this mean that the present in each case is not a now-point at all, but something extended? But how big is then the extended present?

Augustine devotes much space to the question of what we mean by a ›present period‹. Augustine's treatise impressively demonstrates the difficulty by describing a kind of interval nesting:

Behold, thus the present, which, as we thought, could alone be called long, is hardly extended to the duration of a day. But let us also still dissect this itself, since not even one day is present in its entirety. It is filled by twenty-four hours of day and night; for the first of these all others are future, for the last all others past, but for each of the intervening hours those before it past, those after it future. And even the one hour passes in fleeting moments; what has flown away from it is past, what is left of it is future. If any time could be conceived which could no longer be divided into any, even the smallest particles, it alone could be called the present; and yet even this particle of time passes so quickly from the future into the past that it cannot be extended even a moment beyond its duration. [...] ⁹

Since, in a thought-experiment, we can do an infinitesimal interval nesting, Augustine concludes that the time cannot consist of extended parts. What actually exists at the time is only the extensionless moment. But if time consists in a succession of pure now-points, how can it simultaneously be extended? On which basis can we speak of ›short‹ or ›long‹ periods of time, as we actually do? Obviously, the past and the future must somehow exist and be connected to the present. Augustine solves the problem by assuming an existence of past and future which is mediated by the present. At this point, one sees especially the ›psychological‹ character of Augustine's theory of time: past, present, and future exist respectively in the presenting con-

⁹ Augustine 2016, XI.20.

sciousness as *praesens de praeteritis, praesens de praesentibus, praesens de futuris*.¹⁰ Augustine concludes that time exists nowhere else than in the soul (*anima*), namely in the form of memory (*memoria*), current perception (*contuitus*) and anticipation (*expectatio*).

Time, however, is experienced as a continuum; we do not experience the transition from one phase of time to the next. Augustine takes this point into account by describing the unity of our mental experience as the background of the unity of the experience of time. He illustrates this with the example of a song (*canticum*):¹¹ I reproduce a song known to me in such a way that my memory retrieves the section with the corresponding stored memories and then reproduces them step by step. Then my *memoria* has present at each moment what has already been reproduced on the one hand and what is yet to be reproduced on the other. At this point, one can see Husserl's theory of retention and protention anticipated (more on this in a moment). Augustine thus traces our temporal experience of duration and succession, i.e., of a constant awareness of the past and the future, to the unifying action of the soul:

I want to recite a song which I know by heart; before I begin, my expectation is directed to the whole, but once I have begun, what I have already supplied to the past from the expectation extends within my memory. So this activity of mine is divided in its duration into memory, as far as I have said it, and into expectation, as far as I want to say it; present, on the other hand, is my attention, through which what was future passes to become past. Now the more this happens, the more the expectation decreases and the memory increases, until the whole expectation is exhausted, because the whole action has ended and passed into memory. And what happens with the whole song, happens also with its individual sections and in its individual syllables, the same also in a longer action, of which the song is perhaps only a part, the same in the whole life of man, the parts of which are all the individual actions of man, the same finally with the being of the whole human race, which is composed of the lifetimes of the individual men.¹²

Manifestly, Husserl's inspiration from Augustine's *Confessions* is quite far-reaching. Nevertheless, one cannot claim that Husserl took the core of his phenomenological analysis of time from Augustine. The

¹⁰ Augustine 2016, XI.26.

¹¹ Augustine 2016, XI.28.

¹² Augustine 2016, XI.28.

phenomenological method is highly original; it cannot be dated back to late antiquity. Moreover, a contrast would be exaggerated to characterize Augustine's view of time exclusively as a psychological one, while Aristotle and the Stoics would have understood time merely as a process in nature. Rather, it must be said that Augustine takes the perspective of the psychic in *Confessions* XI because he is under the impression of the Platonic theory of the world-soul and relates it to the genesis of time.¹³ The idea of a cosmic time, which is constituted by the world soul, implies precisely an objective, not a subjective description of the phenomenon of time.

III.

Four of the problems addressed by Husserl in his studies on the experience of time readily reveal the influence of Augustine's reading, namely:

- (a) What is the structure and properties of our consciousness of temporally ordered entities?
- (b) How do internal and external objects appear to us over time?
- (c) How do we arrive at a coherent process consciousness that presents us with a unified process and a stable object? For example, in a piece of music, how do we get from the perception of each individual note to the unity of the melody?
- (d) On which fundament does the unity of time consciousness arise? How does consciousness constitute its objects?

Husserl's considerations can be reconstructed as follows. What is revealed in the phenomenological attitude is, as we saw, the nature of the appearance of temporal objects. Here, physical, objective time is no longer accepted as a point of reference for explanations, but is described from the perspective of our experiencing consciousness. Even from this different point of view, some of the basic properties of objective time still remain, namely the properties two-dimensionality, infinity, successivity, and transitivity. But all assumptions of existence are left aside (Husserl speaks of an ›elimination‹ or ›*Ausklammerung*›).

¹³ Seen this way, not ›we,‹ i.e. the subjects of time perception, are constitutive for physical time, but rather the world-soul. Augustine, however, does not clearly profess this theory, because it seems to be dogmatically incorrect from a Christian perspective.

merung»). Further, he distinguishes between ›transcendent‹ and ›immanent‹ time objects; this distinction arises from the fact that we perceive temporal objects both outside or inside consciousness. External temporal objects are spatial things; according to Husserl's anti-representationalism, these are not mediated by internal images or conceptions. Rather, we are connected with these things in direct intentionality. Any theory of images is explicitly rejected by Husserl. For Husserl, inner objects are, e.g., tones and melodies; the perception of tones forms even Husserl's preferred example for the analysis of time consciousness.

Husserl now goes beyond Augustine in that he not only describes the objective and the subjective dimension of the experience of time, but also interprets both as constituted by the immanent time consciousness. In doing so, he does not understand the spatiotemporal things of the external world as phenomena of consciousness; rather, they appear as existing outside of consciousness. For Husserl, the pure Ego standing behind consciousness is that which generates the objects. Husserl assumes that consciousness is in permanent, immediate change, whereby it comes spontaneously (in the sense of a *genesis spontanea*), from a ›source point‹ (*Quellpunkt*), to the constitution of objects:

The ›source point,‹ with which the ›generation‹ of the lasting object begins, is a primal impression. This consciousness is in constant transformation: constantly the bodily tone-now (scil. consciously, ›in‹ consciousness) changes into something past, constantly a new tone-now replaces the one that has passed over into modification. But when the consciousness passes from the tone-now, the primal impression, into retention, then this retention itself is again a now, an actual being. While it is itself actual (but not actual tone), it is retention of been tone. [...] But every actual now of consciousness is subject to the law of modification. It changes into retention of retention, and that steadily. Accordingly, there is a steady continuum of retention in such a way that every later point is retention for every earlier one. And every retention is already continuum.¹⁴

¹⁴ Husserl 1928, 390. »Der ›Quellpunkt‹, mit dem die ›Erzeugung‹ des dauernden Objektes einsetzt, ist eine Urimpression. Dies Bewusstsein ist in beständiger Wandlung begriffen: stetig wandelt sich das leibhafte Ton-Jetzt (scil. bewußtseinsmäßig, ›im‹ Bewußtsein) in ein Gewesen, stetig löst ein immer neues Ton-Jetzt das in die Modifikation übergegangene ab. Wenn aber das Bewußtsein vom Ton-Jetzt, die Urimpression, in Retention übergeht, so ist diese Retention selbst wieder ein Jetzt, ein

Husserl's analysis is opposed to Brentano's theory of ›original association,‹ according to which now-points are linked by consciousness, more precisely by the imagination. Brentano assumes that it is imagination that keeps just past tones present and thus generates the unity of the tone sequence. For Husserl, on the other hand, ›primordial impressions,‹ i.e., objects that change over time in a similar way as we know from the phenomena of spatial perspectivity, rather appear within a processual continuum. Continuity results from the fact that the primordial impressions, the retentions and the protentions merge into each other. Such an ›original time field‹ is what Husserl sees as his task to describe. Thereby a peculiar perspectival change of the objects occurs:

While a movement is perceived, moment by moment an as-now apprehension takes place, therein the now current phase of the movement itself is constituted. But this now-perception is, as it were, the core to a comet's tail of retentions, related to the earlier now-points of the movement. But if perception no longer takes place, if we see no more movement, or—if it is a matter of a melody—if the melody is played and silence has occurred, then the last phase is not followed by a new phase of perception, but by a mere phase of fresh recollection, to this again such a phase, and so on.¹⁵

In the experience of time we experience, as Husserl explains following Augustine (and Brentano) using the example of listening to a melody, a ›passive synthesis‹ that generates a temporal extension. The presence of the past is thereby precisely not attributed to a ›reifying‹ memory, but to a ›retention‹. Husserl distinguishes an active memory

aktuell Daseiendes. Während sie selbst aktuell ist (aber nicht aktueller Ton), ist sie Retention von gewesenem Ton. [...] Jedes aktuelle Jetzt des Bewußtseins unterliegt aber dem Gesetz der Modifikation. Es wandelt sich in Retention von Retention, und das stetig. Es ergibt sich demnach ein stetiges Kontinuum der Retention derart, daß jeder spätere Punkt Retention ist für jeden früheren. Und jede Retention ist schon Kontinuum.«

¹⁵ Husserl 1928, 391. »Während eine Bewegung wahrgenommen wird, findet Moment für Moment ein Als-Jetzt-Erfassen statt, darin konstituiert sich die jetzt aktuelle Phase der Bewegung selbst. Aber diese Jetztauffassung ist gleichsam der Kern zu einem Kometenschweif von Retentionen, auf die früheren Jetztpunkte der Bewegung bezogen. Findet aber keine Wahrnehmung mehr statt, sehen wir keine Bewegung mehr, oder—wenn es sich um eine Melodie handelt—ist die Melodie abgespielt und Stille eingetreten, so schließt sich an die letzte Phase keine neue Phase der Wahrnehmung an, sondern eine bloße Phase frischer Erinnerung, an diese wiederum eine solche usw.«

of things past (secondary memory) from a lingering experience of what has just passed (such as the sound just heard) and calls this ›primary memory‹ or retention. Similarly, a presence of what is immediately forthcoming is to be explained by the notion of protention, which denotes a kind of ›anticipatory thinking‹ in which the experience of time continues into its immediate future. The retentive and protentive continuity and extension of the present cannot be traced back to the dimension of real time as a structure of experienced time, but requires a subjective constitutional analysis of the way in which time experience is ›temporalized‹ and thereby presents itself as temporal:

It still remains to be discussed in more detail what kind of modification we called retentive. One speaks of fading away, subsiding etc. of the contents of sensation, when actual perception passes over into retention. Now it is already clear after the previous explanations that the retentive ›contents‹ are not contents at all in the original sense. When a tone subsides, it is itself at first perceived with a special fullness (intensity), and this is followed by a rapid decrease of intensity. The tone is still there, is still felt, but in mere reverberation.¹⁶

Thus, the diminishing of the intensity of a sound does not mean its complete disappearance, but its ›shadowing‹ (*Abschattung*). For Husserl, the retention of a sound that has just been perceived is an entirely different phenomenon from that of ›secondary memory‹. Unlike Augustine, who explains the presence of the complete song in consciousness by means of *memoria*, it is the case, according to Husserl, that the present experience relates to retention as a comet relates to a comet's tail.

The performance of a constitution of the time experience is attributed to the transcendental I, which, however, cannot itself be in time in this function. The pure I does not appear in the immanent time of the time experience. How does it constitute time? First, it is important

¹⁶ Husserl 1928, 392. »Noch bleibt näher zu erörtern, welcher Art die Modifikation ist, die wir als retentive bezeichneten. Man spricht von Abklingen, Verblässen usw. der Empfindungsinhalte, wenn eigentliche Wahrnehmung in Retention übergeht. Nun ist es aber schon nach den bisherigen Ausführungen klar, daß die retentiven ›Inhalte‹ gar keine Inhalte im ursprünglichen Sinne sind. Wenn ein Ton abklingt, so ist er selbst zunächst mit besonderer Fülle (Intensität) empfunden, und daran schließt sich ein rasches Nachlassen der Intensität an. Der Ton ist noch da, ist noch empfunden, aber im bloßen Nachhall.«

for Husserl to emphasize that the possibility of memory does not presuppose that one can compare an image to its original:

It is fundamentally wrong to argue: How can I know in the now about a not-now, since I cannot compare the not-now, which is no more, with the now (namely the memory-image existing in the now)? As if it belonged to the essence of memory that an image existing in the now would be substituted for another thing similar to it and I could compare and would have to compare as with pictorial imagination. Memory or retention is not picture-consciousness, but something totally different.¹⁷

The important point is rather that time is constituted as an objective entity by reference on the part of memory:

We run through the melody in the imagination, we hear, as it were, first the first tone, then the second tone, and so on. In each case there is always one tone (or one tone phase) in the now point. The preceding ones, however, are not erased from the consciousness. With the perception of the now appearing, as it were now heard tone, the primary memory of the just as it were heard tones and the expectation (pro-tention) of the outstanding ones merge. The now point has for the consciousness a time yard (*Zeithof*), which takes place in the continuum of memory conceptions, and the entire memory of the melody consists in a continuum of such time yard continuums, or of perception continuums of the kind described.¹⁸

This requires the secondary, reproductive memory, i.e., recollection in which the former time field is reproduced. In this way, we arrive at

¹⁷ Husserl 1928, 394. »Es ist grundverkehrt zu argumentieren: Wie kann ich im Jetzt von einem Nicht-Jetzt wissen, da ich das Nicht-Jetzt, das ja nicht mehr ist, nicht vergleichen kann mit dem Jetzt (nämlich dem im Jetzt vorhandenen Erinnerungsbild)? Als ob zum Wesen der Erinnerung gehörte, daß ein im Jetzt vorhandenes Bild für eine andere, ihm ähnliche Sache supponiert würde und ich wie bei bildlicher Vorstellung vergleichen könnte und vergleichen müsste. Erinnerung bzw. Retention ist nicht Bildbewußtsein, sondern etwas total anderes.«

¹⁸ Husserl 1928, 395f. »Wir durchlaufen die Melodie in der Phantasie, wir hören gleichsam zuerst den ersten, dann den zweiten Ton usw. Jeweils ist immer ein Ton (bzw. eine Tonphase) im Jetztpunkt. Die vorangegangenen sind aber nicht aus dem Bewußtsein ausgelöscht. Mit der Auffassung des jetzt erscheinenden, gleichsam jetzt gehörten Tones verschmilzt die primäre Erinnerung an die soeben gleichsam gehörten Töne und die Erwartung (Protention) der ausstehenden. Der Jetztpunkt hat für das Bewußtsein einen Zeithof, der sich in der Kontinuität von Erinnerungsauffassungen vollzieht, und die gesamte Erinnerung der Melodie besteht in einem Kontinuum von solchen Zeithofkontinuen, bzw. von Auffassungskontinuen der beschriebenen Art.«

an objective sequence of several nows. This means that we constitute the objective time by the possibility to refer to certain points in time in the memory.

People have often expressed fundamental criticism of Husserl's conception of our understanding of time. One of the main problems of the phenomenology of inner time consciousness lies in the question whether this phenomenology can be developed at all without hidden borrowings from objective time. After all, Husserl's descriptions are themselves processual, thus seeming to presuppose a course while they are supposed to explain it. To this, one could possibly reply that Husserl describes consciousness as a whole as thoroughly temporal.¹⁹ Another criticism arises from the fact that Husserl uses metaphors such as stream, river, spring, tail, yard, etc., for his descriptions, which tend to reduce conceptual precision due to their strong imagery. Further, it has often been criticized that Husserl possibly focuses on the concept of now in his analysis of time perception in the sense of a ›metaphysics of presence‹.

Three further objections against Husserl's conception have been advanced by Wolfgang Kersting: (i) There is a reduction of the subject of time to the direct perception of objects; thereby reduction of the subject of time to a theory of perception. (ii) Husserl's analysis of single points of temporal perception fails to make intelligible the unity of perceived objects; but we hear a melody, not the succession of tones. (iii) Subjectivity is not properly described from the perspective of perceiving a sequence of now-points.²⁰ But as a philosophical contribution to the analysis of inner time consciousness, Husserl's analyses nevertheless remain valuable. I think that the point that deserves to be maintained is his ›phenomenological reduction‹: the description of mental phenomena free from any ontological and epistemological commitments.

References

Aristotle (1957): *Physics*, Volume I: Books 1–4, Cambridge, Mass.: Harvard University Press.

¹⁹ Cf. de Warren 2009, 40.

²⁰ Kersting 1992, 85–86. Similarly, Seel 2010, 76–88, criticizes Husserl's attempt to derive the course of time from the consciousness of an absolute observing subject.

- Augustine (2016): *Confessions*, Volume II: Books 9–13, Cambridge, Mass.: Harvard University Press.
- Beyer, Christian (2011): Husserl's Conception of Consciousness, in: Cramer, Konrad/Beyer, Christian (eds.): *Edmund Husserl 1859–2009. Contributions on the occasion of the 150th anniversary of the philosopher's birth*, Berlin: de Gruyter, 43–54.
- von Herrmann, Friedrich-Wilhelm (1992): *Augustinus und die phänomenologische Frage nach der Zeit*, Frankfurt a.M.: Klostermann.
- Husserl, Edmund (1976): *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*. Erstes Buch: Allgemeine Einführung in die reine Phänomenologie. 1. Halbband, ed. by Karl Schuhmann, The Hague: Nijhoff.
- Husserl, Edmund (1969): *Zur Phänomenologie des inneren Zeitbewusstseins (1893–1917)*, ed. by Rudolf Boehm. Reprint of the 2nd verb. Edition, The Hague: Nijhoff.
- Husserl, Edmund (1973): *Zur Phänomenologie der Intersubjektivität. Texte aus dem Nachlass*. Dritter Teil: 1929–1935, ed. by Iso Kern, The Hague: Nijhoff.
- Husserl, Edmund (1980): *Vorlesungen zur Phänomenologie des inneren Zeitbewusstseins*. 2. Aufl., Tübingen: Niemeyer.
- Husserl, Edmund (1984): *Logische Untersuchungen*. Zweiter Band: Untersuchungen zur Phänomenologie und Theorie der Erkenntnis, ed. by Ursula Panzer, The Hague: Nijhoff.
- Husserl, Edmund (1999): *Erfahrung und Urteil*. 7th ed., Hamburg: Meiner.
- Kersting, Wolfgang (1992): Selbstbewußtsein, Zeitbewußtsein und Zeitwahrnehmung, in Forum für Philosophie Bad Homburg (ed.), *Zeiterfahrung und Personalität*. Frankfurt a.M.: Suhrkamp, 57–88.
- Seel, Gerhard (2010): Why Time Cannot Be a Form of Inner Anschauung. Critical Remarks on Kant's and Husserl's Explanations of Time Consciousness, in Lohmar, D. & Yamaguchi, I. (eds.): *On Time: New Contributions to the Husserlian Phenomenology of Time*, Dordrecht : Springer.
- Römer, Inga (2010): *Das Zeitdenken bei Husserl, Heidegger und Ricoeur*, Dordrecht: Springer.
- de Warren, Nicolas (2009): *Husserl and the Promise of Time. Subjectivity in Transcendental Phenomenology*, Cambridge: Cambridge University Press.
- Zahavi, Dan (2002): Husserl and the Problem of Pre-Reflexive Self-Consciousness, in: Hüni, Heinrich & Trawny, Peter (eds.), *The Appearing World*. Berlin: Duncker & Humblot, 697–724.
- Zahavi, Dan (2004): Time and Consciousness in the Bernau Manuscripts, in *Husserl Studies* 20, 99–118.
- Zahavi, Dan (2005): *Subjectivity and Selfhood*, Cambridge, Mass.: MIT Press.

Wittgenstein on the Language of Time and the Solipsism of the Present

1. Introduction

In his *Philosophical Remarks*, Ludwig Wittgenstein presents us with an intriguing comparison:

If I compare the facts of immediate experience with the pictures on the screen and the facts of physics with pictures in the film strip, on the film strip there is a present picture and past and future pictures. But on the screen, there is only the present.¹

Wittgenstein illustrates with this comparison how it might be tempting to distinguish between two radically different modes of time. There seems to be a substantive difference between the temporal succession of events in the world, neatly following one after the other along a timeline like a string of pearls, and the specific temporal mode of representation of these events in the mind. For, even though events represented in the mind also possess an analogous order, comprising not only present experiences but also recalled past events and anticipated future events, this order originates with, and is sustained by, an experiencing subject in the present. Hence it seems a sharp distinction must be drawn between an absolute order of time in which worldly events unfold, and a relative order of time, in which representations of present, past and future events are constructed by persons living in the perennial, timeless now.

As with some of the other philosophical temptations Wittgenstein tries to combat, this temptation to draw a substantive distinction between two radically different orders of events was one Wittgenstein

¹ Wittgenstein PR, V, 51. Wittgenstein's writings and other classical texts are cited using the usual abbreviations. In the case of *Philosophical Remarks* Roman numerals refer to chapters, Arabic numerals to sections.

himself was prone to. This is reflected by a distinction Wittgenstein appears to have had endorsed briefly during his so-called »middle« or »transition« period. More precisely, it occurs only in a few scattered remarks recorded by George Edward Moore, Friedrich Waismann, and Alice Ambrose, respectively, all of which have probably been made between 1929 and 1932. In these remarks Wittgenstein distinguishes between two different concepts of time, memory time and physical (or information) time. This short-lived distinction is somewhat obscure. On the one hand, Wittgenstein simply does not elaborate much on it. On the other hand, the passages in which he does so are in tension with one another, and it is difficult to tell whether these tensions are to be attributed to conflicting statements Wittgenstein made or to conflicting recordings of otherwise consistent statements.²

Regardless of how exactly the distinction is to be understood, it is closely connected to another one Wittgenstein used to draw—and which he set out to overcome in his *Remarks*—the distinction between a primary (phenomenological) and a secondary (ordinary or physical) language.³ In my view, one of the reasons why Wittgenstein abandoned both distinctions was the realisation that the ontological dualism implicit in them would result in a problematic position with which he struggled in many of his works: *solipsism*. *Solipsism*, in particular the *Solipsism of the Present Moment*, claims only the experience of the present moment deserves to be called »real.«

In this contribution, I will argue that Wittgenstein wanted to overcome this dualism while insisting on the characteristic differences between the concept of a physical time and a concept of time with regard to memories and representations in general.⁴ To support this claim, I will first consider the broader historical and philosophical context. The temptation of solipsism is particularly present in Russell's theory of sense-data and the protocol sentence controversy in the Vienna Circle—and Wittgenstein was aware of both. I will then

² The distinction occurs in Wittgenstein PO, 110, Wittgenstein WWC, 53, and Wittgenstein WL, 15. For a comparative analysis of these passages cf. Schulte 2006. For an account of the development of Wittgenstein's conception of time in general cf. Hintikka 1996 and Rizzo 2016.

³ Cf. Wittgenstein PR, I, 1.

⁴ Just as many other philosophical projects with naturalist but anti-reductionist tendencies, Wittgenstein's methodological approach can be broadly understood as exhibiting ontologically monist and epistemologically dualist characteristics. Cf. Sturma 2001, 175.

reconstruct Wittgenstein's argument against solipsism in the *Remarks*, which I take to be targeted at an underlying philosophical confusion that can be avoided by paying careful attention to the language of time.

2. In Search of an Epistemic Foundation

Appearances are deceptive. This is as basic an epistemic fact as it is a platitude about everyday life. But if we take seriously Descartes' advice never to trust those who deceived us only once, then it seems all the more necessary to investigate the reliability of our perceptual faculties given that they are one of our main sources of information about the world. And so, in good Cartesian fashion, Russell in his influential 1912 book *The Problems of Philosophy* employs the method of systematic doubt to acquire knowledge about the basic composition and structure of reality.⁵

Interestingly, in his search for an epistemic foundation for our access to the world, Russell is not only concerned about the fallibility of our perceptual faculties, i.e., the simple fact that representations can be true or false. In the history of philosophy, this fact alone gave rise to the search for a criterion of truth enabling us to distinguish with certainty between true and false representations. The *cataleptic representation* devised by the Stoics and Descartes' *clear and distinct representation* are well-known examples.⁶ But however troubling the potential indistinguishability of true and false representations may be, Russell also draws attention to the perspectival character of representations. This is due to the seemingly innocuous observation that the same objects may appear differently to different persons without anyone being epistemically at fault.

Consider, for example, an ordinary table. Depending on the lighting conditions and the angle from which it is seen, and the intensity with which we knock on it, the resulting experiences delivered by the senses—the table's particular colour, shape, touch, and sound—may vary significantly, not only among different persons but also with

⁵ Cf. Russell PP, ch. 1 and 2, on which the following reconstruction is largely based.

⁶ For a detailed analysis of the role of criteria of truth in anti-sceptical philosophical projects cf. Gabriel 2020, in particular ch. 9. Another important example are the *protocol sentences*, which I will address below.

regard to a single person. Since the very same table cannot possess the conflicting properties attributed to it in different experiences, a distinction between the table itself and how it appears to different observers must somehow be drawn. In doing so, Russell distinguishes between »sense-data« for that which is »immediately known in sensation,« and the »physical object« for the table itself.⁷ For Russell, there is no reason to privilege a particular way the table may appear to one of the observers under different conditions—all perspectives are considered equally valid and can lay equal claim to be considered the truth. Russell summarises his view in the following way:

It has appeared that, if we take any common object of the sort that is supposed to be known by the senses, what the senses *immediately* tell us is not the truth about the object as it is apart from us, but only the truth about certain sense-data which, so far as we can see, depend upon the relations between us and the object. Thus what we directly see and feel is merely ›appearance,‹ which we believe to be a sign of some ›reality‹ behind.⁸

This statement implies that the real properties of the table are never directly revealed to us in our experiences, but only indirectly inferred from them. This becomes clear by considering the problem of fallibility. It is important to note that the fallibility of representations is not per se worrisome. Rather, it is a reminder of the fact that each and every representation may be true or false. Even if a large portion of our beliefs about the world would prove to be false, this would not in itself threaten the idea of our having access to the reality of the world in principle. It would simply mean that we would be wrong about a lot more things than we thought. Making a virtue of necessity, one may even consider the fallibility of representations a sign of quality because the possibility of misrepresenting events in the world at the same time implies the possibility of representing them accurately. Regardless of how often our epistemic arrows may miss their worldly target, at least there is the genuine possibility of hitting it. So, the mere fallibility of representations is no cause for epistemic despair but rather reason to think of our perceptual faculties as conducive to fin-

⁷ Russell PP, 4.

⁸ Russell PP, 6.

ding the truth. To undermine our access to the world in a profound way requires further arguments.⁹

Now, according to Russell's view, we never have direct access to physical objects. Experiences may be of physical objects, but the *contents* of experiences are sense-data. As he expressly states, what we »directly see and feel is merely ›appearance‹,« so that even in cases where perceptual contents correctly represent the corresponding properties of physical objects, our epistemic access remains confined to the subjective sphere of sense-data. This opens up the possibility of radical scepticism. The ontological gap between appearance and reality may be so deep as to be unbridgeable since the connection between sense-data and physical objects is only contingent. Thus, it comes as no surprise that Russell is not only concerned with the question whether appearances accurately reflect their objects but also whether they represent anything at all. It may very well be, so the argument goes, that there is nothing outside our epistemic life of sense-data:

Thus, when we are trying to show that there must be objects independent of our own sense-data, we cannot appeal to the testimony of other people, since this testimony itself consists of sense-data, and does not reveal other people's experiences unless our own sense-data are signs of things existing independently of us. We must therefore, if possible, find, in our own purely private experiences, characteristics which show, or tend to show, that there are in the world things other than ourselves and our private experiences.¹⁰

Russell concedes right from the outset the impossibility of proving the existence of physical objects in addition to sense-data. Although our private experiences, our thoughts and feelings, enjoy the status of epistemic certainty, they can only serve as a relatively secure inference basis for what they are experiences of because there is not necessarily any connection between subjective experience and the object that is being experienced. Hence the distinction between sense-data and physical objects is consistent with a sceptical scenario in which there is only a world of sense-data. To fend off such sceptical manoeuvres, Russell appeals to common sense arguments and simplicity considerations. We are entitled to what he calls »instinctive beliefs«,¹¹ i.e.,

⁹ Cf. Gabriel 2020, ch. 5. Paradigmatically, Descartes' *dream argument* and his *evil demon argument* are designed to achieve precisely that.

¹⁰ Russell PP, 10.

¹¹ Russell PP, 11.

reasonable assumptions of explanatory value arrived at without irrefutable evidence. For example, the simplest explanation for the fact that I have the visual impression of a table in front of me, or of a cat passing by, is that there really is a table in front of me and there really is a cat passing by. Unless there is hard evidence to the contrary, we have no reason to doubt such common sense hypotheses. Thus, although sense-data cannot be invoked in a proof of the existence of physical objects, the certainty with which we can know them makes our experiences a »solid basis from which to begin our pursuit of knowledge.«¹²

A very similar rationale motivated the members of the Vienna Circle in their search for an epistemic foundation. Taking a linguistic turn, they debated vigorously whether it was possible to identify a certain class of sentences that registered pure facts and were thereby particularly suitable to serve as a basis for scientific knowledge. This class of sentences became known as »protocol sentences« or »observation sentences.« As Moritz Schlick, one of the debate's main participants, puts it:

»Protocol sentences,« as the term indicates, were originally thought to be those sentences that state the *facts* with absolute simplicity, and without shaping, modifying or adding anything, which are the subject matter of all science and which precede every claim about the world and all knowledge.¹³

Otto Neurath, another important member of the Vienna Circle, gives the following example of a protocol sentence: »In the room, at 3:15 AM, there was a table perceived by Otto.«¹⁴ And Rudolf Carnap—the third central figure in the debate besides Schlick and Neurath—defines protocol sentences as descriptions of »the immediate contents of experience or phenomena, that is of the most simple, discernible facts.«¹⁵ It should be noted that these are just some of the different kinds of examples discussed in the Vienna Circle and that the protocol sentence controversy was far more complex and multi-faceted than is

¹² Russell PP, 8.

¹³ Schlick, Moritz: *Über das Fundament der Erkenntnis* (1934), quoted in Stöltzner/Uebel 2006, 431. All translations of quotes from this volume were done by me.

¹⁴ Neurath, Otto: *Protocol sentences* (1932), quoted in Stöltzner/Uebel 2006, 403.

¹⁵ Carnap, Rudolf: *Die physikalische Sprache als Universalsprache der Wissenschaft* (1932), quoted in Stöltzner/Uebel 2006, 322.

often assumed. Carnap alone, for example, developed four different conceptions of protocol sentences.¹⁶

Now, Schlick has laid out clearly the semantics of protocol sentences. He starts by distinguishing them from what he calls »affirmations« (»Konstatierungen«), specified as a class of sentences reporting first-person perceptions such as »Here now such and such,« »Here now yellow is bordering blue,« and »Here now pain.«¹⁷ In general, affirmations »express a fact registered by one's own ›perception‹ or ›experience‹ in the present.«¹⁸ The modifier »in the present« is particularly important because it is precisely this feature that distinguishes them from protocol sentences. The latter type of sentences, such as »In the room, at 3:15 AM, there was a table perceived by Otto,« represents a fallible hypothesis for Schlick because recording it involves memory. Since memory is a notoriously fallible source of information and prone to error, protocol sentences can be subjected to doubt. Affirmations of the present, on the other hand, cannot sensibly be doubted. The deeper reason for this is that, whereas understanding and verifying a sentence are separate operations regarding most empirical sentences, in the case of affirmations they coincide: »by grasping their sense I simultaneously grasp their truth.«¹⁹ Thus, Schlick maintains it would be as nonsensical to concede the possibility of error for affirmations like »Here now pain« as it would be for tautologies such as »It is either raining or not.« In this regard, affirmations are also more like analytic sentences but with the added bonus of possessing empirical import providing us with knowledge about the world.

However, and despite his conception of affirmations as indubitable expressions of first-person experiences, Schlick has no intention to make use of them to build an epistemic foundation. On the contrary, Schlick has no illusions about the epistemological value of his affirmations in particular and of observation sentences in general:

Hopefully, it has become clear that what is essential here is the mark of the present, which is congenial to the observation sentences, and to

¹⁶ For more historical context of the philosophical topics of the Vienna Circle cf. the comprehensive introduction in the volume by Stöltzner/Uebel 2006. For the protocol sentence controversy in particular cf. LVI-LX.

¹⁷ Schlick in Stöltzner/Uebel 2006, 442.

¹⁸ Schlick in Stöltzner/Uebel 2006, 442.

¹⁹ Schlick in Stöltzner/Uebel 2006, 451.

which they owe their value as well as their uselessness: their value of absolute validity and their uselessness as a lasting foundation.²⁰

Why does Schlick take this surprising stance? Once again, the decisive point here is connected to the aspect of time. According to Schlick, there are, strictly speaking, no genuine reports of affirmations. Affirmations are reports of first-person experiences *in the present*, which require the use of indexicals such as »here« and »now.« For Schlick, the ostensive gestures accompanying the use of indexicals (e.g., pointing to something) are part of the identity conditions of affirmations. But as soon as the affirmation is recorded, the indexicals have lost their meaning because now they are used outside the context of a first-person experience, to which they owed their particular character in the first place: »A real affirmation can never be written down, for as soon as I record the deictic words ›here‹ and ›now‹, they lose their meaning.«²¹

3. Wittgenstein on the Language of Time and the Language of Solipsism

Throughout his philosophical career Wittgenstein battled the idea of solipsism in its various guises.²² As already indicated in the introduction, the intensity with which he was preoccupied with this issue has to do with the fact that he himself was not entirely immune to the appeal of this philosophical idea. As Moore recalls:

As regards Solipsism and Idealism he said that he himself had been often tempted to say ›All that is real is the experience of the present moment‹ or ›All that is certain is the experience of the present moment; and that any one who is at all tempted to hold Idealism or Solipsism knows the temptation to say ›The only reality is the present experience‹ or ›the only reality is *my* present experience‹.²³

²⁰ Schlick in Stöltzner/Uebel 2006, 451.

²¹ Schlick in Stöltzner/Uebel 2006, 452. For systematic considerations concerning indexicals and their relation to the concept of time cf. Sturma 2001, particularly 180–184.

²² For a systematic account of Wittgenstein's arguments against solipsism cf. Hacker 1986, ch. VIII.

²³ Moore 1955, 15.

Against the backdrop of the historical and philosophical context provided in the previous section the motivation and allure of solipsism should now be more transparent. Sceptical manoeuvres undermined our confidence in the reliability of our perceptual faculties and exposed the fragility of our epistemic access to the world. The increasing pressure of the sceptical onslaught provoked a search for an epistemic foundation to address this problem. This search basically assumed the form of a rearguard action. Forced to concede that there is no epistemic certainty regarding the sphere of physical objects and other persons besides oneself, Russell retreated to the sphere of sense-data, in which knowledge at least of one's own private experiences was said to be certain. However, as the protocol sentence controversy made clear, not even the sphere of private experiences turned out to be an epistemic safe haven. Since memory is a fallible source of information, the recollection of past experiences is just as much subject to doubt as are representations of the external world. Hence the sphere in which knowledge was possible had to be limited once more, this time to the *present* experiences of a single person. As a last resort, one might indeed be tempted to say that all that is certain and real are one's own present experiences—hence the solipsism of the present moment.

Of course, the most famous argument targeted at solipsism is the so-called *private language argument* in the *Philosophical Investigations*. But the fight against solipsism figures prominently in many phases of Wittgenstein's philosophical development, making an appearance at least as early as 1916 in his *Notebooks*.²⁴ From then onwards it resurfaces frequently as a philosophical challenge Wittgenstein takes on again and again. However, his treatment of solipsism—and his method of doing philosophy in general—differs significantly from the approaches of other philosophers.

Russell treats solipsism as an empirical statement put forward by a radical sceptic, a statement capable of being true or false. He concedes the possibility that life is nothing but a dream, that the external world comprising physical objects and other persons may be nothing but the figment of a sole person's imagination. Russell's response to this sceptical scenario is simply to point out that there are no good reasons to believe this statement to be true. Therefore, according to Russell, we are entitled to discard it, building our world view bottom up with private experiences as epistemic foundation.

²⁴ Cf. Wittgenstein NB, 82 (Entry from 2 September 1916).

In a similar manner, Schlick considers observation sentences »absolutely fixed points,«²⁵ although he has a different conception of the role they can play with respect to science. He does not believe they can be used to form an epistemic foundation due to their fallibility. Only affirmations—i.e., reports of first-person present experiences—possess the mark of infallibility. But since this mark derives from the fact that affirmations are reports of immediate experiences of the here and now, affirmations lose this special character as soon as they are written down because then the indexicals used have already assumed a different meaning. Schlick, therefore, assigns to private experiences and the sentences that record them the opposite role than Russell does. For Schlick, observation sentences in general are not the starting point for scientific inquiries but rather their end point. Science makes predictions that are to be verified by experience. Observation sentences may not be suitable for an epistemic foundation, but they are an »absolute end«²⁶ for scientific inquiries.

Wittgenstein tackles the problem of solipsism from quite a different angle. In his *Notes for Lectures on »Private Experience« and »Sense Data«*—drafted in the mid-1930s—he already addressed the problem of solipsism at some length. Wittgenstein is well aware of the primary motivation of introducing private experiences of sense-data, which is one of the routes leading to solipsism: the possibility for physical objects to appear differently to different observers.²⁷ But while he does not deny this possibility, he cautions us to resist the temptation of solipsism and wonders instead whether it is possible »to imagine a philosophy that would be the diametrical opposite of solipsism.«²⁸ And even though Wittgenstein points out several problematic consequences of the notion of sense-data and of private experiences in general, he mainly focuses on questioning the under-

²⁵ Schlick in Stöltzner/Uebel 2006, 448.

²⁶ Schlick in Stöltzner/Uebel 2006, 448. Schlick's example is the following scientific prediction: If you look through a telescope at a certain time, you will see a spot of light in the crosshairs. This prediction can then be verified by an affirmation (i.e., by simply looking through the telescope at the specified time), which concludes the scientific task. Cf. Schlick in Stöltzner/Uebel 2006, 448.

²⁷ Cf. Wittgenstein NL, 316.

²⁸ Wittgenstein NL, 282. The analogy invoked in Wittgenstein's famous description, in the *Investigations*, of philosophy's aim as showing »the fly the way out of the fly-bottle« (Wittgenstein PI, § 309) already appears in the *Notes for Lectures*: »The solipsist flutters and flutters in the flyglass, strikes against the walls, flutters further. How can he be brought to rest?« (Wittgenstein NL, 300).

lying philosophical misconception in which these notions are rooted. His diagnosis for the defect of solipsism is the mistaken belief that retreating to one's own sense-data and experiences is tantamount to leaving the public sphere behind entirely. It is to think of one's own mind as a private room to which no one else can have access, and that the language used to describe one's inner episodes and mental states can fully be understood only by oneself.

Against this picture Wittgenstein puts the utmost emphasis on what will become one of the major themes running through his writings, namely the public and social character of language use. That is why, in the *Notes for Lectures*, instead of denying the existence or usefulness of sense-data, Wittgenstein draws attention to the fact that even the language of the most private entities imaginable possess an interpersonal dimension:

What we call a description of my sense datum, of what's seen, independent of what is the case in the physical world, is still a description for the other person.²⁹

Observations like these are characteristic of the tectonic shift Wittgenstein wants to set in motion in our conception of language. In an attempt to clarify the workings of our language we ought to pay attention to the language games involving expressions and sentences themselves, rather than to the mental states of the persons playing the various language games:

Why should I say the ›expression‹ derives its meaning from the feeling behind it—and not from the circumstances of the language game in which it is used?³⁰

Now, with this shift away from private episodes to public language use when it comes to investigating the determination and source of semantic content the method of addressing philosophical problems changes as well. The consideration of the allegedly monologic and isolated mental life of a single person gives way to the examination of the rules governing the application of expressions shared by speech communities. Rules for the use of expressions in their entirety compose what Wittgenstein calls *grammar*, and they constitute and regulate the language games we play. Grammar forms an essential part of the structure of linguistic practices because it determines the limits of

²⁹ Wittgenstein NL, 308.

³⁰ Wittgenstein NL, 313.

meaningful discourse. In combatting grammatical confusions, philosophy may therefore act as a »custodian of grammar,« which »can in fact grasp the essence of the world, only not in the propositions of language, but in rules for this language which exclude nonsensical combinations of signs.«³¹

As a consequence, philosophical theses and problems are approached by testing whether the conditions under which they are framed and articulated are consistent or conflict with the conditions set by grammar. Of course, this does neither entail that every philosophical problem is a grammatical problem nor that grammatical investigations exhaust the philosophical toolkit for solving them. At least, however, it is a legitimate starting point for getting a grip on the particular nature of a given problem in the first place.

With these methodological considerations in mind we can now turn to Wittgenstein's argument against the solipsism of the present moment. It can be found in a short and dense chapter in his *Remarks*, which is concerned more generally with the concept of time. First, he describes the muddled situation holding the solipsist captive by comparing it to the relation between a film strip and the film projected on the screen:

If I compare the facts of immediate experience with the pictures on the screen and the facts of physics with pictures in the film strip, on the film strip there is a present picture and past and future pictures. But on the screen, there is only the present.³²

As spelled out in the previous section, the solipsist seeks epistemic comfort in his private experiences, which Wittgenstein equates to a metaphorical film screen. There may or may not be an external world outside the film theatre, so the solipsist reasons, but that there is a film playing on the private screen is certain, if anything is. Wittgenstein tries to expose the solipsist's position as inconsistent by pointing out a confusion in the concept of time involved in taking it. A few paragraphs after introducing the film metaphor Wittgenstein presents a response to the solipsist including an argument against it, as well as a diagnosis of its root cause. The following passage contains his response in a nutshell:

³¹ Wittgenstein PR, V, 54. For Wittgenstein's conception of grammar cf. Forster 2004.

³² Wittgenstein PR, V, 51.

We are tempted to say: only the experience of the present moment has reality. And then the first reply must be: As opposed to what? [...] The proposition that only the present experience has reality appears to contain the last consequence of solipsism. [...] If someone says, only the *present experience* has reality, then the word ›present‹ must be redundant here, as the word ›I‹ in other contexts. For it cannot mean *present* as opposed to past and future. Something else must be meant by the word, something that isn't *in space*, but is itself a space. That is to say, not something bordering on something else (from which it could therefore be limited off).³³

There is a lot to unpack here. In keeping with his philosophical method characterised above, Wittgenstein tries to pin down what exactly the solipsist is claiming in saying that only the present moment or the present experience is real since the solipsist's use of the term seems to deviate from how it would ordinarily be used. The usual meaning of ›present‹ derives a substantive part of its semantic content *in contrast* to other expressions of time, particularly expressions such as ›past‹ and ›future.‹ Roughly, part of the grammar of ›past,‹ ›present,‹ and ›future‹ seems to be such that these expressions share a range of objects to which they can be, in principle, meaningfully applied. That is to say, if it is meaningful to talk of present moments and experiences, then it is as meaningful to talk of past and future moments and experiences.³⁴

However, the solipsist is apparently not using the expression ›present‹ according to its ordinary meaning. Rather than merely privileging present experiences and singling them out from all possible kinds of experiences, including past and future ones, the solipsist wants to make a stronger claim. As Wittgenstein notes in his *Cambridge Lectures*: ›A person who says the present experience alone is real is not stating an empirical fact, comparable to the fact that Mr. S. always wears a brown suit.‹³⁵ At first sight, the property attributions in both cases seem similar. Just as Mr. S. is said to possess the property of always wearing a brown suit, the solipsist claims real experiences possess the property of always being present experiences. But the property attributions are actually categorically different. In the case of Mr. S. there is the genuine possibility for him to wear suits of a different colour; it just so happens that Mr. S.—for whatever reason

³³ Wittgenstein PR, V, 54.

³⁴ Cf. Rizzo 2016, 138.

³⁵ Wittgenstein WL, 25.

—always wears brown ones. The solipsist, on the other hand, does not admit the possibility of real experiences possessing the properties of past or future. Past and future experiences do not even qualify as candidates for being real experiences. There is simply no contrast involved among real experiences because there is only one category: present experiences.³⁶

Wittgenstein illustrates his point metaphorically. The solipsist would take the meaning of the expression »present« to be a space itself, as opposed to something being *in* a space. The usual meanings of »present,« »past,« and »future« jointly constitute a semantic space by mutually determining their contents. The solipsist, however, obliterates precisely this contrast from which the respective expressions of time derived their particular content in the first place, leaving the expression »present« without the semantic foil against which it is thrown into relief. As a result, »present« loses its usual meaning. That is also why Wittgenstein observes that in the claim »Only the present experience has reality,« the expression »present« must be redundant, for the solipsist conceives of »present« as an analytic component of »experience.« In the solipsist's mouth, the meaning of »present« has transformed from a contingent into a necessary property of real experiences, because it is impossible for them not to possess it. Therefore, the solipsist's claim »Only the present experience is real« cannot be an empirical statement, since empirical statements are contingent. They always allow for the possibility for things to be other than stated.

The semantic principle Wittgenstein implicitly appeals to in this passage of the *Remarks* also plays a decisive role in some of his other arguments against solipsism in particular and private language in general. The semantic principle in question is expressed by what I have called elsewhere the *contingency requirement*: for statements to be meaningful they must be contingent in that they express states of affairs that could be otherwise.³⁷ He already made use of this principle in the *Tractatus* and it resurfaces in the *Remarks*.³⁸ It formulates a basic requirement for meaningful discourse, violation of which results in nonsense. Tautologies form an important exception. Although they do possess the property of always being true, they do not

³⁶ Cf. Hacker 1986, 231.

³⁷ Bartmann 2021, 53.

³⁸ Cf. Wittgenstein TLP, 5.634: »Whatever we can describe at all could be other than it is.« Cf. Wittgenstein PR, V, 54: »Language can only say those things that we can also imagine otherwise.«

involve the contrast between true and false simply because they cannot be false. However, tautologies are »senseless« rather than nonsensical for they attribute necessary or logical properties to the objects they refer to.

Now, Wittgenstein's contingency requirement explains why he considers it just as misleading to assert the claim »Only the present experience has reality« as well as to deny it:

And the person who objects to the assertion that the present alone is real with »Surely the past and future are just as real« somehow does not meet the point. Both statements mean nothing.³⁹

The reason why the assertion or denial of the solipsist's claim is futile is that the semantic conditions required for meaningfully asserting or denying the claim are in conflict with the solipsist's main premise. For the claim »Only the present experience has reality« to be meaningful, it would have to be intelligible for »present« either to apply or not to apply to »experience.« But this is being ruled out by the solipsist. Real experiences are thought of as necessarily possessing the property of being present, thereby excluding the possibility for real experiences not to possess that property. In doing so, the solipsist deprives the expression »present« of its usual context of use and transfers it into a new one in which it would have to retain its previous meaning in order to express the solipsist's intention. The solipsist wants to state something along the lines of: »Real worldly events may have taken place in the past. This cannot be known. What can be known is that, when it comes to experiences, present experiences are the only real ones.« What the solipsist does not realise is that the usual context of use establishes the semantic content of »present,« so that transferring it to another context in which the constitutive contrast to »past« and »future« is erased must result either in a change or even a loss of meaning. But then the entire meaning of »Only the present experience has reality« has been changed or lost, against the solipsist's intention. It is as if the solipsist took a piece of a puzzle and put into an entirely different puzzle, in which it may superficially seem to fit in, due to matching colour and form, but would represent a different content because it is now part of an entirely different picture.

In sum, the solipsist commits a semantic transmission error. Wittgenstein illustrates this error again with his film metaphor:

³⁹ Wittgenstein WL, 25. Cf. also Wittgenstein PR, V, 56, where he says denying the solipsist's claim would be »just as wrong as to maintain it.«

The present we are talking about here [i.e., the solipsist's conception of ›present‹] is not the frame in the film reel that is in front of the projector's lens at precisely this moment, as opposed to the frames before and after it, which have already been there or are yet to come; but the picture on the screen which would illegitimately be called present, since ›present‹ would not be used here to distinguish it from past and future. And so it is a meaningless epithet.⁴⁰

However, the verdict of meaninglessness is not Wittgenstein's last word on the matter. Even though he considers the solipsist's position utterly misguided, he does not dismiss it altogether. He seems to be acknowledging a grain of truth in solipsism worth pursuing, because there is something peculiar about experience with respect to time after all, something the solipsist struggles to articulate but fails. The solipsist's mistake consists in employing our ordinary concept of time within a language game to which it does not belong because it is played according to substantially different rules. Wittgenstein's conclusion of the solipsist's transmission error, however, is not to demand uniformity in applying our ordinary concept of time across all language games, but rather to recognise and emphasise the characteristic differences of the meaning of »time« when considering experiences or physical events. The lesson Wittgenstein draws from the solipsist's confusion is therefore to acknowledge two different concepts of time involved in talking about physical events and immediate experiences:

Perhaps this whole difficulty stems from taking the time concept from time in physics and applying it to the course of immediate experience. It's a confusion of the time of the film strip with the time of the picture it projects. For ›time‹ has one meaning when we regard memory as the source of time, and another when we regard it as a picture preserved from a past event.⁴¹

For the sake of argument, Wittgenstein characterises one sense of »memory« as a picture of a past physical event, and the faculty of memory in general as a kind of storage unit in which the recorded events are archived, even though he quickly points out the potential philosophical pitfalls of this simile. However, the purpose of this characterisation is to contrast the concept of physical time with the concept of time with regard to memories in particular, and representations in general—Wittgenstein treats the former as a species of the latter.

⁴⁰ Wittgenstein PR, V, 54.

⁴¹ Wittgenstein PR, V, 49.

Now, the series of physical events corresponds to the film strip in the metaphor. And if our memory functions properly, then the temporal relations of the series of memories reflect the temporal relations of the series of physical events they recorded. The temporal relations of sooner and later, before and after thus constitute the concept of physical time. Wittgenstein then draws attention to quite a different concept of time, which is involved when memory is understood as »the source of time.« This sense of »time« applies to the projected picture in the metaphor, the film playing on screen. The projected picture represents the particular subjective perspective of an experiencing person.⁴²

There are indeed substantive differences between both concepts of time. Whereas the temporal relations of worldly events can be fixed independently, past, present, and future experiences can only be distinguished by reference to the perspective of someone to whom they belong. Of course, it would be nonsensical to say it were possible to experience past and future experiences. Even if it were somehow possible to relive a past experience, such that the present experience was qualitatively identical to the past one, it would still be an experience in the present. But this should not lead one to conclude, as the solipsist mistakenly does, that only present experiences are real. Instead, Wittgenstein recommends keeping apart the different concepts of time lest blending them together creates profound philosophical confusions:

Both ways of talking are in order, and are equally legitimate, but cannot be mixed together. [...]

Thus we cannot use the concept of time, i.e. the syntactical rules that hold for the names of physical objects, in the world of representation, that is, not where we adopt a radically different way of speaking.⁴³

⁴² Wittgenstein's distinction is reminiscent of McTaggart's famous distinction between the »A series« and the »B series« (McTaggart 1908, 458), although I cannot explore the extent of the similarities here. For an account of personal identity that takes into consideration subjective as well as objective time and that maintains the unity of reality without sacrificing the irreducibility of the subjective perspective of persons, cf. Sturma 2008, particularly 576–579.

⁴³ Wittgenstein PR, V, 49, transl. am. Wittgenstein's recommendation regarding the concept of time is a particular instance of his general strategy not to import terms of our ordinary language into the language of experience: »The worst philosophical errors always arise when we try to apply our ordinary—physical—language in the area of the immediately given. [...] All our forms of speech are taken from ordinary, physical language and cannot be used in epistemology or phenomenology without casting a distorting light on their objects« (Wittgenstein PR, VI, 57).

Finally, one may wonder whether Wittgenstein's epistemically dualist approach of distinguishing sharply between those two different concepts of time also has ontologically dualist consequences. Although Wittgenstein is always extremely cautious when it comes to ontological questions, his answer in this case seems to be clear:

Memory and reality must be in one space. I could also say: representations and reality are in one space.⁴⁴

As the film metaphor already suggests, the two concepts of time, different though they may be, are compatible with a basically monist view of reality. The projected picture on the screen and the film strip in the film reel exhibit characteristic differences, but it does not follow that there are actually two different films playing in the theatre—one film behind the other, if you will. Once more, this conception would lead right back to solipsism, and the film would start all over again.

References

- Bartmann, Marius (2021): *Wittgenstein's Metametaphysics and the Realism-Idealism Debate*, Cham: Palgrave Macmillan.
- Forster, Michael (2004): *Wittgenstein on the Arbitrariness of Grammar*, Princeton: Princeton University Press.
- Gabriel, Markus (2020): *The Limits of Epistemology*, Newark: Polity Press. [First published in German as *An den Grenzen der Erkenntnistheorie*, Freiburg: Alber 2008.]
- Hacker, Peter Michael Stephan (1986): *Insight and Illusion. Themes in the Philosophy of Wittgenstein. Revised Edition*, Oxford: Clarendon Press.
- Hintikka, Jaakko (1996): Wittgenstein on Being and Time, in id., Ludwig Wittgenstein: *Half-Truths and One-and-a-Half-Truths*, Dordrecht: Springer, 241–274.
- McTaggart, John McTaggart Ellis (1908): The Unreality of Time, in *Mind* 17, 457–474.
- Moore, George Edward (1955): Wittgenstein's Lectures in 1930–33, in *Mind* 64, 1–27.
- Rizzo, Giorgio (2016): Wittgenstein on Time: From the Living Present to the Clock Time, in Santioianni, Flavia (ed.), *The Concept of Time in Early Twentieth-Century Philosophy. A Philosophical Thematic Atlas*, Dordrecht: Springer, 137–148.
- Russell, Bertrand (2001): *The Problems of Philosophy* [1912], with an introduction by J. Skorupski, Oxford: Oxford University Press. [Abbr. as PP]

⁴⁴ Wittgenstein PR, III, 38, transl. am.

- Schulte, Joachim (2006): Wittgenstein on Time (1929–1932), in Stadler, Friedrich / Stöltzner, Michael (eds.), *Time and History. Zeit und Geschichte*, Frankfurt am Main: Ontos Verlag, 557–567.
- Stöltzner, Michael / Uebel, Thomas (eds.) (2006): *Wiener Kreis. Texte zur wissenschaftlichen Weltauffassung von Rudolf Carnap, Otto Neurath, Moritz Schlick, Philipp Frank, Hans Hahn, Karl Menger, Edgar Zilsel und Gustav Bergmann*, Hamburg: Meiner.
- Sturma, Dieter (2001): Bewußtsein und Sprache der Zeit. Ein Beitrag zur Reduktionismusproblematik in der Philosophie des Geistes, in Rentsch, Thomas (ed.): *Sprache Erkenntnis, Verstehen. Grundfragen der theoretischen Philosophie der Gegenwart*, Dresden: Thelem, 174–192.
- Sturma, Dieter (2008): Selbstreferenz, Zeit und Identität. Grundzüge einer naturalistischen Theorie personaler Identität, in *Deutsche Zeitschrift für Philosophie* 56, 569–583.
- Wittgenstein, Ludwig (2001): *Tractatus Logico-Philosophicus*. Translated by D.F. Pears and B.F. McGuinness, London/New York: Routledge Classics. [Abbr. as TLP]
- Wittgenstein, Ludwig (2001): *Wittgenstein's Lectures: Cambridge 1932–1935*. From the Notes of A. Ambrose and M. Mcdonald. Edited by A. Ambrose, New York: Prometheus Books. [Abbr. as WL]
- Wittgenstein, Ludwig (1993): *Philosophical Occasions: 1912–1951*. Edited by J. Klagge and A. Nordmann, Indianapolis/Cambridge: Hackett. [Abbr. as PO]
- Wittgenstein, Ludwig (1979): *Notebooks 1914–1916*. Edited by G.H. von Wright and G.E.M. Anscombe and translated by G.E.M. Anscombe, Oxford: Basil Blackwell. [Abbr. as NB]
- Wittgenstein, Ludwig (1979): *Wittgenstein and the Vienna Circle. Conversations recorded by F. Waismann*. Edited by B.F. McGuinness and translated by J. Schulte and B.F. McGuinness, Oxford: Blackwell. [Abbr. as WWC]
- Wittgenstein, Ludwig (1975): *Philosophical Remarks*. Edited by R. Rhees and translated by R. Hargreaves and R. White, Oxford: Basil Blackwell. [Abbr. as PR]
- Wittgenstein, Ludwig (1968): Notes for Lectures on »Private Experience« and »Sense Data«, in *The Philosophical Review* 77, 275–320. [Abbr. as NL]
- Wittgenstein, Ludwig (²1958): *Philosophical Investigations*. Translated by G.E.M. Anscombe, Oxford: Blackwell. [Abbr. as PI]

Mind & Time

The Language of Time and Self-Consciousness

1. Naturalism and the Life of Persons

When we look for distinctive features of the human life-form, we encounter language, self-consciousness, and time. It is these features that we must refer to in answering the question *What does it mean to lead the life of a person?* This question is usually answered in the narrower contexts of philosophy of mind and ethics. In contrast to the conventional approach, this question will be answered, in the following chapter, on the basis of an extended, i.e., non-eliminative, naturalism which seeks access to the essential features of the life of persons *via* their language, behavior, and action.

The eliminative varieties of naturalism have attracted a great deal of attention, especially in methodological debates of the twentieth century—consider, for example, Logical Empiricism, W. V. Quine, or Daniel Dennett – extended naturalism can look back on a long history. It spans from Anaximander and Aristotle to Spinoza, Hume, Wilfrid Sellars, Stuart Hampshire, Peter F. Strawson, and John McDowell. Like any version of naturalism, it is characterized by its insistence on the ontological unity of reality and the rejection of supernaturalistic explanations. It differs from the eliminativist versions, however, in its use of an enlarged concept of what there is, in its methodological pluralism, and in the importance it attaches to epistemological and linguistic clarifications, and, most of all, it does not exclude the question *What does it mean to lead the life of a person?*¹

Due to the importance that extended naturalism gives to conceptual analysis and epistemological evaluation, it comes close, in different ways, to both classical and more recent positions in the philosophy of mind. Unlike classical positions, it does not take a subjective, phenomenological, or normative approach. Self-consciousness

¹ Cf. Sturma 2019, 147–151.

and time are addressed—following the general naturalistic strategy—from the outside. This approach goes hand in hand with the rejection of the method of introspection, which plays a significant role, especially in traditional philosophy of mind and ethics. Naturalistic approaches, by and large, are characterized by the strong belief that the experiences of a person are not introspectively accessible—neither from the outside nor from the inside. There is no private language in the sense that a person could refer to her mental data or impressions. As with the experience of other persons, the ways of dealing with our own experience are linguistically determined.²

In contrast to eliminative approaches, however, extended naturalism implies that descriptions of the world which do not take into account the conduct of persons *as* persons, are based on unjustified eliminations and oversimplifications. According to an extended naturalism, persons are subject to causes that the natural sciences at least partially reveal, but they are also susceptible to reasons. The actions of persons show the ontological distinction to bring about changes in the world of events by virtue of their self-awareness and ability to differentiate and act for reasons. This unique capability manifests itself in the peculiar grammar of the word »I«³ and the language of time in which the elusive phenomena of self-consciousness and time leave identifiable traces—as can be impressively demonstrated by the expressions »I« and »now.« In the end, it turns out that the language of time holds the key to answering the question of what it means to lead the life of a person.

2. Self-Reference and Linguistic Behavior

In a general sense, linguistic behavior is about thought and action.⁴ Extended naturalism, in accounting for the human life-form, can rely

² Cf. Strawson 1959, 100, 106: »One can ascribe states of consciousness to oneself only if one can ascribe them to others. One can ascribe them to others only if one can identify other subjects of experience. And one cannot identify others if one can identify them *only* as subjects of experience, possessors of states of consciousness. [...] There is no sense in the idea of ascribing states of consciousness to oneself, or at all, unless the ascriber already knows how to ascribe at least some states of consciousness to others.«

³ See Wittgenstein 1964, 65–78.

⁴ Cf. Hampshire 1959, Austin 1975.

on linguistic behavior as its empirical foundation. Linguistic behavior is a specific way to change the course of events. It emerges from a systematic interrelation between mind, world, reference, and self-reference. The perspectives of the first-person, second-person, and third-person, respectively the standpoints of the reflecting and acting person, and external observers, register the inferential interaction between language, self-awareness, attitudes, behavior, and action in which shapes of the inside of the outside of the human life-form are revealed.

Thoughts are about something. They are expressed in a rule-governed manner by intentional and propositional attitudes. In line with this, the content of thoughts is constituted along syntactic and semantic rules.⁵ These rules are different from the workings of neural micro-mechanisms, which eliminativist positions take as their empirical reference point. Thus, the question arises as to when and where what is physical becomes something psychological and logical. When answering this question, it is particularly important to clarify how to understand the relationship between the various levels of neural micro-mechanisms, on the one hand, and the language of thoughts, on the other. But it has to be admitted that we are far away from understanding the dependencies between changes of behavior and their corresponding neural micro-mechanisms.

Thoughts can be questioned and examined within the space of reasons. They are about something and have content—all the same, they are subjective in the sense that the thinking person has in an epistemic sense privileged access to her thoughts. But it is the lesson of analytic conceptual analysis that privileged access does not mean that thoughts are exclusively private or confined to an inner world. When analyzing thoughts or states of consciousness, it is important to clarify the semantic foundations of the mentalistic vocabulary that is used to address the abilities and characteristics of persons. Ludwig Wittgenstein pointed out that the expressions of the mentalistic vocabulary are not meant as representations or reifications, but rather as

⁵ For Sellars thoughts and rules are inextricably linked to social practice. Accordingly, rules or rule-governed behaviour are to be regarded as essential manifestations of the human life-form: »When God created Adam, he whispered in his ear, 'In all contexts of action you will recognize rules, if only the rule to grope for rules to recognize. When you cease to recognize rules, you will walk on four feet.'» (Sellars 1980, 138).

different linguistic instruments.⁶ Facts are not mirrored in the language but are expressed in a specific syntactic and semantic form.

In the case of empirical knowledge, we do not stand in a mysterious kind of contact with the world outside. Wilfrid Sellars was very influential in unmasking the fallacies of the myth of the given. He demonstrated that mental data are not unconditionally presented to us, but rather depend on linguistic rules.⁷ According to Sellars, our everyday language already contains far-reaching assumptions and conjectures about the attitudes and states of our consciousness. This includes assigning the observable behavior of persons to an interpretation that contains hypotheses about attitudes of consciousness that cannot be observed. The semantic framework of everyday language ultimately determines the meaning of predicates that people ascribe to other people and to themselves. Sellars elaborates that the understanding of the attitudes and self-relations of persons are carried out in the logical space of reasons as the inferential system of determined predicates. In the case of conceiving the state of mind of other persons, our everyday language thus does not simply depict something, but rather makes assumptions about something hidden to us under the conditions of referential uncertainty. The consequence is that statements about mental states are more or less to be treated as theoretical hypotheses.

The mentalistic vocabulary develops by way of conjectures, confirmations, predictions and refusals, and persons must always deal with epistemic uncertainty in their attitudes and actions—this applies to the subjective perspective of a person, to communicative situations and, equally, to the intersubjective perspective of external observers. The respective scope of action is neither completely transparent to the individual person nor to external observers. These limitations, however, are no reason for a fundamental skepticism about propositional attitudes or epistemic and practical self-relations. It is plausible to assume that, with acute observation and good reasoning, we are often right.

⁶ See Wittgenstein 1964, 67: »The word »I« does not mean the same as »L.W.« even if I am L.W., nor does it mean the same as the expression »the person who is now speaking«. But that doesn't mean: that »L.W.« and »I« mean different things. All it means is that these words are different instruments in our language. Think of words as instruments characterized by their use [...].«

⁷ See Sellars 1997.

If the context is known, only a small number of possible explanations for the behavior in question have to be taken into account. Some types of behavior, such as routines, are easier for the observing person to recognize than for the observed persons themselves. External observers can predict the behavior from the outside with high probability because observable postures and gestures hint at how a person will behave—even if it appears that she has not yet made up her mind.

The observation of behavior, expression, and communication allows conclusions to be drawn about what may be going on from the first-person perspective. To be able to talk about own experiences and experiences of other persons, one does not have to rely on allegedly private information. Sellars has developed an argument for how linguistic behavior leads to forms of self-reference without the presupposition of contained private experiences. With the help of a thought experiment, he outlines minimal features of the logic of ordinary language about experiences. In his thought experiment, a Genius Jones proposes that individuals do not only exhibit intelligent behavior when they speak publicly.⁸ Rather, overt verbal episodes, i.e., intersubjectively accessible linguistic behavior, are to be understood as the culmination of a process that begins with inner episodes. Linguistic behavior is the expression of thoughts that cannot be observed as such. They should be regarded as unobserved and theoretically assumed inner episodes that provide information about observable behavior of persons.

What is revealed in linguistic behavior allows for more or less reliable inferences to be drawn about not observable thoughts, attitudes, and intentions. The ability to have thoughts is formed in the process of acquiring linguistic behavior, and only then does self-reference unfold. What started as a language of hypotheses and conjectures gradually takes on a reporting role, which in the end incorporates self-referential expressions.⁹ Self-referential expressions, though, are themselves sources of fallacies—at least in Western thought. This is especially true of artificial nouns such as »I« or »self.« They suggest

⁸ See Sellars, 1997, 98–109.

⁹ See Sellars 1997, 106, 107: »For once our fictitious ancestor, Jones, has developed the theory that overt verbal behavior is the expression of thoughts, and taught his compatriots to make use of the theory in interpreting each other's behavior, it is but a short step to the use of this language in self-description. (...) *What began as a language with a purely theoretical use has gained a reporting role.*«

that something can be addressed directly as I or self. In fact, these expressions are merely based on fallacious reifications. In linguistic behavior, we can justifiably refer to our attitudes, intentions, and actions and the ones of other persons but not to an ego or self.¹⁰

3. The Peculiar Grammar of the Expression »I«

Descartes is considered the founder of the theory of self-consciousness. However, he never uses the concept of self-consciousness explicitly and draws on a reflection of Augustine for his famous *existo* and *cogito* argument. Contrary to popular belief, the legacy of Descartes' analysis of self-reference does not consist in revealing introspective certainties about ourselves, but rather in the insight that the language we use to understand ourselves in our various epistemic situations reveals something about the constitution of our consciousness: Whenever I express the sentence *I exist* or conceive in my thinking, it cannot be false. The sentence *I do not exist* is self-refuting as long as it is expressed in the present.¹¹ Even if I am a finite being with a limited life span, at the moment of the utterance, the sentence *I exist* expresses an indubitable certainty.

In the famous passage from the Second Meditation, Descartes does not apply a method of conceptual analysis, but he nevertheless has a clear understanding of the semantic working and logic of self-reference, and in retrospect we can learn from him not that introspection tells us something about mind and self-consciousness, as is often claimed, but rather that language does.

The phenomenon of self-consciousness unfolds at the edges of the great divide between mind and world, and it does so with problematic reifications—like the »I« or »self.« These reifications are simplistic reactions to the complicated features of self-consciousness. Not long after Descartes introduced the *cogito* into Western philosophy, numerous attempts have been made to escape from the existential isolation that comes with these reifications.

All indexical expressions, like »I« and »now,« implicitly contain a wider context of predication and reflect a variety of structures and perspectives of human experience. A special case is the logic of the

¹⁰ Cf. Sturma 2005, 77–83.

¹¹ Cf. Descartes 1964, 16–23; Ayer 1956, 44–52.

expression »I.« It refers to the reflecting or speaking person, but points to nothing. This distinguishes it from the indicators »this« or »here,« which, as it were, select objects from the field of perception. The analytical theory of self-consciousness has taken its starting point from this referential feature to work out the structures of self-consciousness on the basis of the peculiar grammar of the word »I.« It has been demonstrated that »I« is neither a reifying noun nor a proper name nor a designation. Its special position is based on its indexical functions in speech situations. Furthermore, the use of »I« is infallible and not correctable. The speaking person assumes straightforwardly and beyond doubt that the mental state in which she finds herself is her own and not somebody else's.¹² In self-referential expressions of my existence, the pronoun »I« has a constitutive role, but its indicating function is not to be confused with an »I« or »self« as an ontological or metaphysical concept.¹³

Although the analytic philosophy of self-consciousness is primarily oriented towards simple deictic functions of the expression »I,« complicated structures of self-reference turn out as a result of its analyses. When the expression »I« is used in the explanation of the phenomenon of self-consciousness and ceases to have the function of a mere indicator, problems that have accompanied the philosophy of self-consciousness since its beginnings return. It appears that even in analytic philosophy, interpretation, and explanation of the phenomenon of self-consciousness are bound to get trapped in a semantic reification, a vicious circle, or an infinite regress.¹⁴

Without entering into the intricate problems of the philosophical theory of self-consciousness, it can be stated that the personal and reflexive pronouns used to express self-consciousness have a correlative structure. The sentence »I have self-consciousness« does not stand for the phenomenon of self-awareness because it assumes a kind of ownership that separates the person from her experiences. Rather, the philosophical interpretation of self-consciousness must work with a more complex structure.

According to Immanuel Kant, the self-consciousness of a person is the starting point of her self-relations over time and is therefore present—in whatever form—both in her reflection and in the content

¹² There are psychopathological disorders in which this certainty vanishes.

¹³ Cf. Geach 1957, 117–123; Cf. Austin 1975, 59–66.

¹⁴ Cf. Castañeda 1966 and 1999, 61–142.

of her mental states. It would therefore be wrong to interpret self-consciousness as ownership or as existential isolation. Rather, the self-conscious person inhabits a perspective by which she experiences herself in a natural and social environment. In this indirect way, self-consciousness already contains the spatiotemporal location and practical attitudes of the particular person.¹⁵

In the light of the contextualization of self-consciousness, the sentence »I exist« takes on a new character. It fulfills the function of an opening to the empirical content of reflection and consciousness. Therefore, regardless of its analytic elements, Kant can transform the cartesian *existo* and *cogito* into the empirical sentence »I exist thinking,«¹⁶ which indicates an activity in space and time and not, as it were, a free-floating and empirically unbound state. The expression »I exist thinking« captures the complicated relationship between the determining and the determined elements in self-consciousness as well as the unavoidable intertwining of existence and thinking in the life of persons. It expresses not only an indubitable self-certainty, but also the spatiotemporal position of the thinking and acting person. Contrary to what Descartes suggests, she experiences her self-certainty not in momentary evidence, but in specific attitudes over time and in relation to what she is not.

Appropriate sentences to express self-consciousness would be: »I am aware of *myself*« or »I am aware that I am in the state *p*.« Being conscious of myself, I have no consciousness of a »self,« but rather of having a certain experience in a specific situation, which stands in an identifiable relation to an event in the spatiotemporal world. As long as she lives, a person takes a specific standpoint in the world with identifiable spatiotemporal coordinates. For this reason, she is intimately familiar with herself in self-consciousness. She does not have to identify herself and cannot be deceived about being self-aware at this very moment. We can consider everything, which exists right now, as contingent from which we could abstract. But we can never take a contingent attitude towards ourselves and our in time moving present. The secret of the peculiar grammar of self-reference is the fact that there is no contingency in our self-awareness.

¹⁵ Cf. Sturma 2018, 143–146.

¹⁶ Kant 1999, 456 (B 428).

4. Indexicality and Non-Indexicality

The peculiar grammar of the expression »I« plays a special role in the system of indexicals and non-indexical expressions. Although persons cannot adopt a contingent attitude toward themselves, they are able to express facts without explicit self-reference and indexicality. Self-reference and indexicality do not belong to a closed domain. They are elements of a linguistic practice that provides access to both intersubjective and objective facts. This practice becomes most evident in the relationship between indexicality and non-indexicality, and, more generally, in the language of time.

Indexicals are the basis of spatiotemporal orientation. They provide information not only about the internal structure of temporal relations, but also about the conditions of experiential states as a whole. In this respect, indexicality turns out to be constitutive of both the epistemic *and* the social attitudes of persons because they reciprocally imply the subjectivity of experience. Persons know that for the indexicals they use from the perspective of first-person singular, there are corresponding expressions in the perspectives of the second-person and third-person singular and plural. Indexicals mark the physical and temporal position of a person. Subjectively, she experiences her personal identity over time as a continuous demarcation between her mind and the world,¹⁷ and existence initially means only being present at a certain time in an objective place.¹⁸

In their use, indexicals depend on situations of perception and predication that are identifiable in their spatiotemporal coordinates. Indexicals—like explicit self-reference—say something about the process of experiences and its semantic content. Since transformations from indexicality into non-indexicality are epistemic extensions in which the experiential perspective is methodically excluded, indexical reference does not appear on the syntactic surface of non-indexical sentences. Transformations are only possible in the case of complementary relations between subjective and objective localizations. Substitutions of subjective expressions in objective expressions cannot be carried out in a semantically appropriate way if they are depen-

¹⁷ Cf. Castañeda 1999, 180–203. [The Self and the I-Guises].

¹⁸ Cf. Tugendhat 1975, 28: »[A]lthough time and space are equally essential for the identification of the thing, the existence of the thing is essentially temporal; the limits within which it exists are dates, not places. For an extended thing *to be* thus means to be present *somewhere* in space during a *certain* time.«

dent on the experiential perspective of persons. The sentence *I am aware of being in a hopeless situation* differs in its meaning from the sentence *A person is aware of being in a hopeless situation*. The latter sentence can refer to both me and another person, but the emotional state changes dramatically in the case of the first-person perspective.

With the indexical representation of the experiential perspective, asymmetries arise which no longer allow transformations without loss of semantic content. Perspectives are essential for every process of experience. Animals also perceive their environment from a spatiotemporally identifiable position that is, as it were, the anchor of their perceptual field. Different from other life-forms, the perceptual field of persons is formed by their linguistic behavior.

The distinctive perspective of human experience is revealed in the structures of indexicality. Human experience unfolds in a space of semantic options shaped by spatial and temporal perspectives. They are constitutive of communication and interpersonal understanding. In states of understanding and self-understanding, persons orient themselves by means of indexical references to the corresponding communicative or experiential context.¹⁹

Regardless of the unavoidable experiential perspective and the non-contingency of self-reference, a person is able to formulate meaningful sentences that are not dependent on self-referential contexts for their meaning and validity. By way of her knowledge and productive imagination, she can move in dimensions that have no relation to possible experience and construct meaningful sentences beyond the limits of comprehensible temporal understanding—for example: »Four billion years ago there was no life on earth.« Strangely enough, such sentences are quite difficult to formulate without any time coordinates. One could say, at the expense of temporal precision, that there was no life on earth before the formation of the asteroid belt. However, it is obvious that a sentence like this is semantically dependent on the assumption of a human perspective on the cosmos. These difficulties of »de-indexicalization« suggest that non-indexical sentences must be experientially anchored in some form. In the case of the sentence »There was a period of time in which there was no life on earth,« it must be presupposed that we are in our solar system and that we are talking about *our* past.

¹⁹ Cf. Gale 1964, 103/104; Bar-Hillel 1954, 368: »We see that effective communication by means of indexical sentences requires that the recipient should know the pragmatic context of the production of the indexical sentence-tokens.«

Language is permeated by a system of deictic expressions, by which persons can refer to one and the same object with different pronouns at different times and places. This system is the manifestation of the irreducible interrelation of subjective localizations and objective localizations in space and time.²⁰ The internal connection between self-consciousness and time is particularly evident in the functions of personnel pronouns. Persons using the expression »I« are already aware that they are addressed by other persons from the second-person and third-person perspective in the same way as they themselves address other persons in the second-person and third-person perspective. In the case of the use of the expression »now,« persons know that the present event is addressed at other times with »earlier« or »later,« respectively.

In communicative situations, indexical functions are mainly carried out by personal, spatial and temporal pronouns. Nevertheless, expressions like »I,« »here,« and »now« do not identify anything. They only indicate a position in space and time or in the communicative situation. They share the epistemic and linguistic context-dependencies, but their respective indexical functions show remarkable differences. Spatial indicators are less unambiguous compared to temporal and personal indicators, and the expression »now« has a recognizably more restricted function than other temporal or spatial indicators.

Like the expression »I,« the expression »now« is subject to linguistic constraints that leave little room for modified uses. Whom I address with »you« or another personal pronoun is just as open to individual choice as the demonstrative identification of objects by »there,« »this« or the like. This is true to a limited extent even for the expression »here.« However, this flexibility is missing when using the expressions »I« and »now« in direct speech. This parallelism of »I« and »now« indicates the special relation between self-consciousness and time.

Within the system of indexical references, the expressions »I« and »now« occupy a special position because they are fundamental to any form of predication. But perceptions and predications only happen with indexicality *and* reference to identifiable objects or events. Persons refer to the same object through a system of indicators at different times and places by using corresponding expressions such as »now,«

²⁰ Cf. Strawson 1959, 15–58.

»then,« »here,« »there,« »you,« »they,« or »we.« This is only possible, when the subjective indexicalizations find support in objective localizations.

Personal and temporal indicators carry specific time relations. It is the linguistic function of the expressions »I« and »now« is to express immediacy. But considerable difficulties arise in trying to decipher this simple process in everyday language. While suggested by the usages of »I« and »now,« it is not possible to express immediacy precisely due to the discursive constitution of language. The sentence *the expression »I« refers to the one who is speaking now* has no meaning in the strict sense: »Now« is not now anymore, and »I« merely indicates a position in the speech act.

The relation between the expressions »I« and »the one who speaks now« remains unclear—at least in the sense that they do not share the same time segment. Such considerations may seem sophisticated, but they nevertheless reveal a special feature of human language: Language talks about immediacy by transcending it. In semantic terms, the immediate present is an abstraction. Self-consciousness can only unfold in time.

A person who says »now« finds herself already at different instants of time. Someone who understands the language of time is able to change the course of things in a very specific way. In this sense, the use of the expressions »I,« »now« and »here« clearly shows extensions of reality. How far the extension of the present goes depends on the state of attention and the normative potential of the respective practical attitudes.

5. The Language of Time

The ontological status of the human life-form is expressed not only in the linguistic manifestations of self-reference, indexicality, and non-indexicality, but also in the semantics of time. The language of time is primarily defined by tensed language, temporal adverbs, and prepositions. A time-dependent proposition makes a temporal statement for a particular moment in time, by which the statement is true only then and at no other time. It is a semantic oddity that temporal determinations, as such, are not without reference, but do not refer to something in the sense of events or objects in space and time. The unclear ontological status of temporal expressions is by no means

mysterious. While it must be admitted that a concise conceptual understanding runs into a variety of difficulties, we can deal reliably with time in everyday language.²¹

Persons experience time in two ways: as a passage from the future through the present into the past, and as a succession of earlier, simultaneous, and later events. For this distinction, John McTaggart coined the expressions *A-series* and *B-series*,²² which are still applied in the philosophy of the time, and it is to this day one of the greatest challenges of the theory of time to understand the reality of time and the asymmetric relation between the *A-series* and the *B-series*.²³

The time relations of the *A-series* are undoubtedly dependent on personal attitudes. When a person regrets something because she made decisions and performed actions in the past that turned out to be wrong, she makes, so to speak, moves through the space of reasons *within* her subjective framework of epistemic and practical attitudes. But this does not mean, even from the personal standpoint, that the *A-series* is independent of the *B-series* or that the *A-series* has priority

²¹ Cf. Augustine 2016, 238: »[Q]uid est ergo tempus? si nemo ex me quaerat, scio; si quaerenti explicare velim, nescio [...].« Augustine does not claim that time must be considered an eternal mystery, as is often implied. He merely draws attention to the fact that difficulties arise in explaining the phenomenon of time.«

²² Cf. McTaggart 1927, 9–10: »Positions in time, as time appears to us *prima facie*, are distinguished in two ways. Each position is Earlier than some and Later than some of the other positions. To constitute such a series there is required a transitive asymmetrical relation, and a collection of terms such that, of any two of them, either the first is in this relation to the second, or the second is in this relation to the first. [...] In the second place, each position is either Past, Present, or Future. The distinctions of the former class are permanent, while those of the latter are not. If *M* is ever earlier than *N*, it is always earlier. But an event, which is now present, was future, and will be past. [...] For the sake of brevity I shall give the name of the *A* series to that series of positions which runs from the far past through the near past to the present, and then from the present to the near future and the far future, or conversely. The series of positions which runs from earlier to later, or conversely, I shall call the *B* series.« Cf. Prosser 2016, 2–21.

²³ For McTaggart, the asymmetrical relationship between *A-series* and *B-series* has been a clear indication for the unreality of time; cf. McTaggart 1927, 10: »Since distinctions of the first class [*B-series*] are permanent, it might be thought more objective, and more essential to the nature of time than those of the second class [*A-series*]. I believe, however, that this would be a mistake, and that the distinction of past, present and future is as *essential* to time as the distinction of earlier and later, while in a certain sense, as we shall see, it may be regarded as more *fundamental* than the distinction of earlier and later. And it is because the distinctions of past, present and future seem to me to be essential for time, that I regard time as unreal.«

over the *B*-series. As we have seen, the *A*-series opens with the subjective perspective of a person. This fact is linguistically reflected in the special meaning of indexical expressions, in which the respective perceptual and predication situations of persons are manifested. But this opening takes place under the ontological conditions of the *B*-series. Persons establish their point of view in the world of events they share with other persons and living beings.

The question can be raised whether a *B*-series can exist at all without persons or other life-forms that experience temporal sequences. Even animals traverse the *A*-series in the *B*-series, though, without showing explicit self-consciousness or a language of time. It can be considered whether the local and temporal expressions in play can be replaced by elements of the *B*-series and its spatial equivalents, such as »closer than« or »further than.« Such replacements would result in even more artificial expressions, which are no longer of any use for the analysis of the indexical structures of human consciousness. The implicit indexical dependence of time-independent expressions must not be overestimated. It is merely an indication of the special epistemic situation of persons and can by no means be the reason for the assumption of an independence of the *A*-series from the *B*-series.

The scope of the language of time is demonstrated in the large amount of timeless and time-independent expressions whose meaning does not depend on explicit references to subjective perspectives. Whereas time-dependent propositions articulate temporal relations for a particular moment in time, and are true only in this respect, time-independent propositions are formulated, as it were, from nowhere. These intersubjective and objective components of the language of time play a significant role in our everyday communication and understanding.

It is the distinguishing quality of human language that it is capable of expressing situation-independent and time-independent propositions like *x is F*. The meaning of sentences with this formal structure is based on the elimination of temporal contexts: An object is in a timeless manner said to possess the property *F*. The transformation carried out in this process has been interpreted as an indication of the factual and methodological primacy of the *B*-series over the *A*-series. This conclusion is overhasty. Statements of the type *x is F* are indications of specific properties of propositional attitudes, which are unrecognizable in the transitions from time-dependent to timeless

statements. The loss of explicit linguistic representation is by no means a blind spot. To ignore this fact leads to the fallacy of mistaking linguistic peculiarities for ontological facts.

Time-dependent assertions have to be transferred into the dimension of *B*-relations to be able to abstract from subjective standpoints and to proceed to timeless assertions. Timeless descriptions are sentences in Wittgenstein's »book of the world« which records all physical properties and events of the world but leave untouched the internal structure of the experiential perspective of the one who can read this book at least in parts.²⁴ The first-person perspective is, in no significant sense, part of the world the book describes, and this is as true for the »omniscient person« as for any other. But this does not take us away from the fact that the thought experiment of the book of the world can only be conceived on the basis of the epistemic situation of the human life-form. The transition to timeless sentences takes place with the help of abstractions from the specific context of the speech act and is semantically supported by the propositional content of the respective statement. However, the abstraction is justified only if the context of the speech act is negligible with regard to its propositional content. The pages from the »book of the world« do not have the function of saying anything about the experiential situation of self-awareness. Their constructive profit excludes that the »book of the world« can stand in for reality as a whole. Besides, it is unclear what place such a book could have in the world at all, especially since it cannot contain itself by design.

Different ways of observing and exploring runs neither into ontological nor epistemological difficulties. But one must take the possibility into account that the language of the *A*-series, irrespective of its constitutive function in propositional attitudes, presupposes the objectivity of the *B*-series. The epistemological function of the *A*-series does not support the view that it could be sufficient for adequately constituting human cognition. Such an assumption would amount to the fallacy of tacitly inferring from an epistemological argument for the irreducibility of the *A*-series in human conscious-

²⁴ Siehe Wittgenstein 1965, 6: »Suppose one of you were an omniscient person and therefore knew all the movements of all bodies in the world dead or alive and that he also knew all the states of mind of all human beings that ever lived, and suppose this man wrote all he knew in a big book, then this book would contain the whole description of the world.«

ness the claim that time-consciousness finds its foundation only in itself.

The epistemic situation of the human life-form does not stand for itself but depends on ontological structures that are accessible for cognitive consciousness. The relations of the *B*-series always, though largely unnoticed, have an influence on the semantical and syntactical characteristics of the space of reasons and linguistic behavior. From an epistemic point of view, a person leads her life where the *A*-series and the *B*-series converge. Her mental activities move through the *A*-series, and, by virtue of her epistemic and practical capabilities, she is in position to intervene in the world of events whose relations are reproduced in the *B*-series.

The presence of a person is a segment of the path from future to past, along which her later action *b* as event *B* follows to her earlier action *a* as event *A* for all times. Persons lead their life in the perspective of the *A*-series and are able to react practically to objective events, which proceed in the sequence *earlier—at the same time—later*. The convergence of the *A*-series and the *B*-series indicates that the life of persons is subject to different orders and regularities in one and the same ontological space.

The naturalistic argument of the unity of reality is not opposed to presuppositions of the irreducibility of the subjective perspective in the life of persons. My present wanders, as long as I live, through the *B*-series and constantly creates new dimensions of past and future. When I pass away, my *A*-series collapses. With the *B*-series, we enter a world in which experiential perspectives emerge and fade away. Let us forget, time begins when someone relates processes to each other, such as the movement of the sun with the hand of a sundial.

6. Time and Self-Consciousness in Action

Persons live in a world that is structured by the relations of succession and simultaneity, and they perceive them through the temporal relations of future, present, and past. The *A*-series represents the *B*-series in the mode of self-referential consciousness. This way, structures of the world relate to structures of conscious finitude in the sense of being aware of its own limited existence. Persons, as manifestations of conscious finitude, live in a world of irreversible change. They are at the origin of the *B*-series, which becomes the ontological refe-

rence for the *A*-series. The passage from the future to the present and the past is nothing other than the conscious expression of irreversible changes. Accordingly, temporal divisions have a cognitive function regarding the identification of events, objects, and individuals that undergo irreversible change.

The relation between the *A*-series and the *B*-series determines the conduct of the life of persons. This phenomenon is evident in the case of time-independent attitudes. Persons have to live their lives over time, and their actions require them to be able to interpret particular temporal positions in their evaluation and conduct. They must gain an understanding of how they can and should deal with their past and future in the respective present. Such complicated attitudes can only be implemented if reasons for action are not completely situation-dependent. Only then can an adequate practical understanding that the present is only an elusive moment in the life of persons be acquired.

In normative terms, time-independent attitudes are relevant because they are the source of reasons for action and prudential considerations that are valid not only for a specific moment in time. A good reason for action cannot be found solely in the perspective of the immediate present. Reasons, which are not directly dependent on a specific situation, provide a vision of how a situation or an event may take shape in the future and will present itself in retrospect. The expansion of the present through time-independent evaluations generates reasons that are not only applicable to the present; they are expressions of practical continuities in time and over time.

Because people live their lives in the present, a bias towards the present can hardly be avoided.²⁵ An epistemic as well as practical response to the bias toward the present is the extension of the present. Although a person always acts only in the present, it would be a dramatic restriction of the potential of personal life if she were to limit her reflections and actions to the present. The extension of the present demonstrates the fact that persons have the capacity to adopt a reflective attitude towards the present moment and to conceive it from the outset as one among other temporal positions.

Unlike spatiotemporal objects, persons can deliberately and intelligibly relate to their past and future. Not only do they emerge and pass away in time, as everything that exists, but they can actively

²⁵ Cf. Parfit 1986, 158–193.

establish epistemic and practical relations to time and over time. Time consciousness is the orientation toward past and future in the present, and it opens up the possibility of alternatives and thereby enriches one's present experience with an awareness of other possible worlds. As persons, in the course of the temporal transitions of their lives, transcend immediacy and contingency, at least in part, through actions over time, they break the chain of external constraints. While this does not mean that they live in a self-determined world, they can at least initiate actions as events that can be traced back to reasons. In the extended present, the past and the future do not appear as a sea of chance, but as a range of options in which certain actions were realized or refrained from or are to be realized or refrained from on the basis of reasons.

Persons need by no means judge reasons for action only according to their proximity to the present.²⁶ Rather, it belongs to the inherent possibilities of their conduct not to have to reject reasons for past action because they contradict immediate needs or desires. They are able to exceed the present following the arrow of time. Leading the life *as* a person can only come about through reasons for action, which, in their internal structure, are not wholly dependent on the motivational situation of the present moment. The meaning and impact of reasons for action unfolds because they apply to more than one moment in time, and in this sense, they represent values that are time-independent.

Since the immediate present does not necessarily override time-independent reasons for action, persons are able to establish continuities over time. At most, the degree to which continuities extend across time may be debatable. Therefore, it has to be asked whether the popular maxim *carpe diem*, which clearly expresses a bias towards the present, is useful from a practical point of view.

The maxim *carpe diem* implies the possibility of continuities over time, only to warn explicitly against it: while it is an option for action, it deprives people of the benefits of the present. If the warning is supposed to suggest that every moment is to be considered only as if it could be the last one, it falls short in terms of leading the life of a *person*, i.e., living self-aware over time. Persons are at every stage of

²⁶ Cf. Rawls 1971, 420: »We are to see our life as one whole, the activities of one rational subject spread out in time. Mere temporal position, or distance from the present, is not a reason for favoring one moment over another.«

their lives more than an ensemble of centerless and random attitudes and intentions.

From the perspective of extended presence, the existentially short-sighted idea of permanently existing as the last moment vanishes, for even the last moment of a life can still contain visions of possible worlds. As the finitude of the individual person is only empirically, but not ideally the last word, persons should act as if the next moment could never be the last. Otherwise, the absurd situation would arise in which people would permanently have to experience the unpleasant surprise of being still alive. Such a case is perhaps conceivable as an extreme experience which can occur in or after life-threatening events, but in everyday situations it only generates absurdities and is not at all suitable for expressing the authentic view of a self-aware person.

Due to the arrow of time, it is the future which has a practical priority over the present and the past. We must operate on the assumption of an asymmetrical relationship between the past and the future in the present. Since the arrow of time already points into the future, persons cannot form an idea of the present without an image of the future, no matter how narrowly defined. At any point in her life, it is open to a person to think far beyond the probable time of her death and to feel committed to a world which she will no longer inhabit as a physically present person. Without denying that persons exist only for a very limited time and under conditions of external constraints, finitude and continuity can be reconciled with a conception of self-determination as self-extension that implies that there is enough time to make the realization of authentic continuities over time worthwhile. In the lives of persons, it is not contingency and blind chance that coincide, but finitude and continuity.

References

- Augustine (2016): *Confessions*, Volume II: Books 9–13, Cambridge, Mass.: Harvard University Press.
- Austin, J. L. (1975): *How To Do Things With Words*, Cambridge, Mass.: Harvard University Press.
- Ayer, Alfred (1956): *The Problem of Knowledge*, Harmondsworth: Penguin.
- Bar-Hillel, Yehoshua (1954): Indexical Expressions, in *Mind* 63.
- Castañeda, Hector-Neri (1966): ›He‹ On the Logic of Self-Consciousness, in *Ratio* 8, 130–157.

- Castañeda, Hector-Neri (1999): *The Phenomeno-Logic of the I*. Essays on Self-Consciousness, Bloomington: Indiana University Press.
- Descartes, René (1964): *Meditationes de Prima Philosophia*, in *CŒuvres de Descartes* vol. VII, ed. by Charles Adam & Paul Tannery, Paris: Vrin.
- Gale, Richard M. (1964): Is It Now Now? in *Mind* 73.
- Geach, Peter (1957): *Mental Acts. Their Content and Their Objects*, London: Routledge and Kegan Paul.
- Hampshire, Stuart (1959): *Thought and Action*, London: Chatto and Windus.
- Kant, Immanuel (1999): *Critique of Pure Reason*, Translated and edited by Paul Guyer and Allen W. Wood. Cambridge University Press.
- McTaggart, John McTaggart Ellis (1927): *The Nature of Existence, Volume II*, Cambridge: Cambridge University Press.
- Parfit, Derek (1986): *Reasons and Persons*, Oxford: Oxford University Press.
- Prosser, Simon (2016): *Experiencing Time*, Oxford: Oxford University Press.
- Rawls, John (1971): *A Theory of Justice*, Cambridge, Mass.: Harvard University Press.
- Sellars, Wilfrid (1997): *Empiricism and the Philosophy of Mind*, Cambridge, Mass.: Harvard University Press.
- Strawson, P. F. (1959): *Individuals. An Essay in Descriptive Metaphysics*, London: Methuen.
- Sturma, Dieter (2007): Person as Subject, in *Journal of Consciousness Studies* 14, No. 5–6, 77–100.
- Sturma, Dieter (2016): Self-Consciousness, Personal Identity, and the Challenge of Neuroscience, in Miguel García-Valdecasas, José Ignacio Murillo, Nathaniel F. Barrett (Hg.), *Biology and Subjectivity: Philosophical Contributions to Non-reductive Neuroscience*, Cham: Springer, 13–24.
- Sturma, Dieter (2018): The Practice of Self-Consciousness: Kant on Nature, Freedom, and Morality, in Eric Watkins (ed.), *Kant on Persons and Agency*, Cambridge: Cambridge University Press, 138–152.
- Sturma, Dieter (2019): Persons. A Thick Description of the Human Life Form, in Jörg Noller (ed.), *Was sind und wie existieren Personen?*, Leiden: mentis, 147–165.
- Tugendhat, Ernst (1975): Existence in Space and Time, in *Neue Hefte für Philosophie* 8, 14–33.
- Wittgenstein, Ludwig (1958): *Philosophical Investigations*, Oxford: Blackwell.
- Wittgenstein, Ludwig (1964) *The Blue and Brown Books: Preliminary Studies for the ›Philosophical Investigations,‹* Oxford: Blackwell.

Temporal Cognition, Impulsivity and Intertemporal Choice

1. Introduction

Temporal cognition refers to the processes relating temporal characteristics of the physical world to the subjective experience of time.¹ Time is an essential and pervasive dimension in virtually all aspects of cognition and behaviour. Much of our thought, experience and behaviour is anchored in time, ranging from the recollection of past memories via present experiences such as pain or boredom, to plans and expectations about the future.² Neurophysiological studies have failed so far to identify a single brain region solely dedicated to our capacity for temporal cognition, in notable contrast with the functional specialisation of areas processing auditory or visual information. Instead, a number of brain structures are thought to process time-related information, such as hippocampus, striatum, insula, fronto-parietal cortex and cerebellum, and there are numerous theoretical models of how these areas and connectivity among them may represent time.³

One domain of cognition that is almost by definition highly sensitive to aspects of time is decision making. Decision making refers to a set of perceptual and cognitive processes that unfold over time and result in the choice of one response or course of action over another (or several others). Decisions are usually evaluated with regards to the (subjective) consequences for the decision maker or others. Various scientific disciplines have contributed to our understanding of decision making, most prominently psychology, economics, neuroscience and philosophy. Within these interdisciplinary decision

¹ Matthews & Meck 2016.

² Droit-Volet & Meck 2007.

³ Wittmann 2013.

sciences, the cognitive and neural processes underlying decision making are frequently investigated in situations of risk or uncertainty, i.e., when the probability of obtaining an expected outcome is known (but not 0 or 1) or poorly specified, respectively.⁴

The temporal dimension inherent in the decision-making process has also been studied intensively.⁵ Specifically, not only the decision-making process itself but also the consequences of one's decisions always unfold over time, thereby placing decision-making squarely in the domain of temporal cognition. Notably, depending on the specific situation, the consequences of one's decisions may be experienced almost immediately, within a few hundred milliseconds, such as in the decision where to look next,⁶ or much later, after years and decades, as in the decision which profession to take up or which insurance scheme to invest in.⁷

One type of decision task that is particularly relevant in the context of temporal cognition is intertemporal choice.⁸ Intertemporal choice tasks model situations, abundant in everyday life, in which people have to decide between rewards of different magnitudes which are available at different times in the future. One example is the decision whether to spend money on an immediately rewarding activity, such as a night out, or to save it towards a future, and potentially more rewarding activity, such as an overseas holiday. In a typical intertemporal choice task, participants choose between smaller rewards that are available sooner and larger rewards that are available later. In general, people prefer rewards that are available earlier over rewards of the same size but available at a later time, thereby devaluing future rewards. This cross-species phenomenon is known as temporal discounting, or delay discounting.⁹

A number of theoretical explanations have been suggested to explain temporal discounting, including reward- or value-based accounts, cognitive accounts (focussing on representations, cognitive control, and prospection), personality-based accounts and (perceived)

⁴ Pleskac et al. 2015.

⁵ Kalenscher & Pennartz 2008 and Read & Scholten 2017.

⁶ Noorani & Carpenter 2016.

⁷ Malkoc & Zauberaman 2019 and Sutter 2014.

⁸ Frederick et al. 2002.

⁹ Ainslie 1975 and Frederick et al. 2002.

time-based accounts.¹⁰ In this chapter, we focus on time-based accounts of temporal discounting, given that time perception itself and its effects on intertemporal choice have recently gained considerably in interest.¹¹ In doing so, we also explore the relationship between intertemporal choice and impulsivity, as impulsivity has been hypothesised to reflect altered perception of time¹² and given that the widely reported finding of increased temporal discounting in individuals with higher levels of impulsivity¹³ may be mediated by individual differences in temporal processing.¹⁴

We will first introduce intertemporal choice tasks as a prototypical paradigm of temporal cognition. We will then review studies of the effects of temporal processing manipulations on discounting behaviour, such as designs involving episodic future thinking or framing of temporal information. Next, we will discuss explanations of the temporal discounting phenomenon that are based on temporal cognition rather than other aspects of cognition (such as risk preferences). Finally, we will summarise and discuss evidence on the interrelationships between temporal processing, intertemporal choice and impulsivity.

2. Fundamentals of Intertemporal Choice

In its most canonical form, intertemporal choice refers to the choice between a smaller reward that is to be received sooner and an alternative, larger reward to be received at a greater delay. Thus, both decision options have two attributes: reward size and delay. Examples include the Marshmallow Test («Would you rather have one marshmallow now or two in 15 minutes' time?»)¹⁵ and the so-called consumption-savings problem («How much of my income or wealth should I consume now and how much should I save and consume more at a later date?»).¹⁶

¹⁰ B. K. Kim & Zauberman 2009, Peters & Büchel 2011 and Zauberman & Urminsky 2016.

¹¹ Zauberman & Urminsky 2016.

¹² Paasche et al. 2019.

¹³ Keidel et al. 2021.

¹⁴ Paasche et al. 2019.

¹⁵ Mischel 2014, Mischel et al. 1989, Shoda et al. 1990 and Watts et al. 2018.

¹⁶ Choi et al. 2006, Hall 1988 and Modigliani & Brumberg 1954.

For most decision-makers, the two attributes of decision options in an intertemporal choice task induce a trade-off: larger rewards are typically preferred over smaller rewards and rewards received sooner are typically preferred over rewards received later, *ceteris paribus*. Research on intertemporal choice is concerned with the modelling of this trade-off and the factors that influence it.¹⁷

A central issue in negotiating the trade-off between reward size and delay is the (non-)commensurability of the two attributes—how to compare an increase in one unit of reward to a decrease in one unit of delay? Typically, a decision-maker's preferences for reward size and delay are inferred from their choices. The latter can be used to infer a ranking of different options that were available to the decision-maker, which is usually referred to as preferences. Under certain conditions, this ranking can be represented by a so-called utility function, a mapping of reward size and delay into utility.¹⁸ Thus, a utility function transforms the individual attributes of decision options—reward size and delay—into a common unit. The assumption is that in a given decision situation, a decision-maker would choose the option with the higher utility.

A utility function allows comparing a unit of reward, say 1 Euro, received at some delay, with a reward of similar size received immediately. The conversion of a future reward into its immediately-received equivalent, using a utility function, is called discounting and the value of a future reward discounted to the present is referred to as present value. Presenting a decision-maker with a series of decisions involving choices between rewards received immediately and rewards of different sizes received at varying delays allows inference of the so-called discount rate, the rate at which the decision-maker discounts future rewards per unit of time. A higher discount rate implies steeper discounting, that is, stronger devaluation of future rewards. Discount rates are often considered to be a measure of impulsivity.¹⁹

Differences in temporal discounting can manifest not only in the discount rate, the degree of discounting per unit of delay, but also in the functional form of the discount (utility) function. Different functions, such as exponential, hyperbolic and quasi-hyperbolic functions,

¹⁷ Frederick et al. 2002.

¹⁸ Mas-Colell et al. 1995.

¹⁹ Keidel et al. 2021.

typically differ in the weight they assign to different delays.²⁰ An important discount function is the exponential discount function. As its name indicates, future rewards are discounted as an exponential function of time: the utility (or discounted value) today of a future reward is given by $R/(1+r)^t$, where R is the value of the future reward, r is the discount rate per unit of time and t is the delay. This function assumes that the utility of delaying receipt of a reward is compounded over time: utility of delaying does not only accrue from the initial reward but also from additional utility gained in intermediate periods. This is reflected in the exponential form of the discount factor (Fig. 1).

Empirical research has demonstrated that human intertemporal choices are better described by a hyperbolic discount function, where the utility today of a future reward is given by $R/(1+kt)$.²¹ As the term indicates, this discount function assumes that utility only grows linearly over time. Compared with exponential discounting, a hyperbolic function discounts rewards more steeply at shorter delays and less steeply at longer delays. It can lead to so-called »preference reversals«: A person who prefers \$20 received today to \$30 received in three months may also prefer \$30 received in one year and three months to \$20 received in one year's time. Such behaviour is inconsistent with exponential discounting and is sometimes considered to be »irrational.«²²

²⁰ Frederick et al. 2002.

²¹ Kable & Glimcher 2007/2010.

²² Frederick et al. 2002.

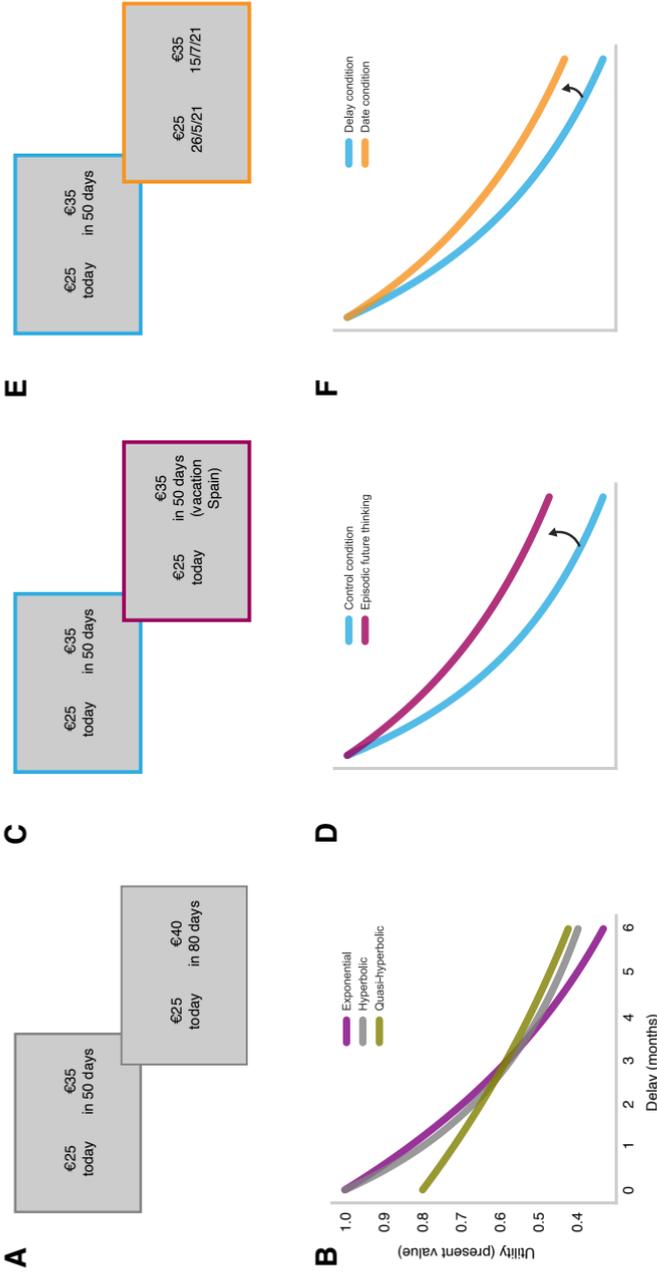


Figure 1: The Intertemporal Choice Task

Legend: (A) Examples of choices in an intertemporal choice task. (B) Different discount functions, mapping delay until a \$1 reward is received into utility (present value). (C) Examples of choices in an intertemporal choice task investigating the role of episodic future thinking in temporal discounting. (D) Illustration of the effect of episodic future thinking condition on the discount function. (E) Examples of choices investigating the role of different ways of presenting delays in an intertemporal choice task. (F) Illustration of the effect of date vs. delay condition on discount function.

The discount functions discussed so far take as their arguments objective quantities, in particular, reward size and delay, and transform them directly into utility units, which in turn can be mapped into behaviour. Other approaches convert objective quantities into subjective values before mapping them into utilities. One motivation for the latter approach is to take properties of the perceptual system, e.g., number perception, into account when mapping stimulus properties (e.g., reward size, delay) into behaviour. A prominent example is the application of the Weber-Fechner law to interval timing in scalar expectancy theory,²³ which we discuss in more detail below. Other models take into account behavioural heuristics.²⁴ At this point, it is still an open question which factors determine the functional form of an individual's discount function.

Utility (or discount) functions have also been useful for identifying and characterising brain processes associated with intertemporal choice.²⁵ The parameters of these functions, in particular, the discount rate, can be estimated based on the choices a person makes, for example, in an intertemporal choice task. Subsequently, these functions can be used to compute the utility of individual decision options that a person is presented with. Brain imaging techniques such as functional magnetic resonance imaging (fMRI) have been used to locate brain areas and networks in which activation is correlated with the utility of decision options during intertemporal choice.²⁶

This work has not only identified a network of brain regions associated with intertemporal choice but also structural and functional neural properties associated with individual differences in intertem-

²³ Takahashi 2005 and Zauberman et al. 2009.

²⁴ Lempert & Phelps 2016.

²⁵ For recent reviews, see Frost & McNaughton 2017, Peters & Büchel 2011 and Schüller et al. 2019.

²⁶ For meta-analysis, see Bartra et al. 2013.

poral choice.²⁷ The evidence so far suggests that certain structures and functional patterns in prefrontal, striatal, and subcortical brain regions associated with valuation, cognitive control, memory and future-oriented thinking, can, to some degree, account for individual differences in temporal discounting. It has been suggested that some of these structural and functional differences may serve as neural biomarkers of temporal discounting, that is, an objective biological measure of temporal discounting.²⁸

3. Time-related Experimental Effects

The previous sections have highlighted the importance of temporal processes in decision making, in particular in intertemporal choice, both in theory and experimental design. This section will add to this notion by focusing on two influential experimental effects on intertemporal choice behaviour.

Generally, experimental manipulations in intertemporal choice paradigms can help to elucidate the psychological and neural mechanisms of intertemporal choice that lead to non-rational, hyperbolic discounting behaviour. Within the context of temporal cognition, research has aimed primarily at two different processes, namely the perception of outcome values at different points of time and the perception of time itself.²⁹ On the one hand, subjective valuation of objectively smaller, sooner and larger, later outcomes can be differentially influenced by visceral or emotional processes³⁰ or by cognitive representations of outcomes.³¹ On the other hand, though possibly related, people are insufficiently sensitive to future time, and manipulations of time sensitivity can modulate intertemporal choice behaviour.³² Specifically, Zauberman et al. showed that participants' subjective estimates of prospective time duration (marks on a line ranging from »very short« on the left end to »very far« on the right end) were not reflected by a linear function of objective time but by

²⁷ Keidel et al. 2021.

²⁸ Cho et al. 2013, Guo et al. 2017, Ikuta et al. 2018.

²⁹ Ballard & Knutson 2009 and B. K. Kim & Zauberman 2009.

³⁰ Loewenstein 1996.

³¹ H. Kim et al. 2013 and Malkoc & Zauberman 2006.

³² Ebert & Prelec 2007, B. K. Kim & Zauberman 2009, Zauberman et al. 2009.

its logarithmic transformation.³³ Interestingly, this is consistent with both hyperbolic discounting in intertemporal choice tasks and the Weber-Fechner law.³⁴

Based on these important general findings, two specific experimental effects on intertemporal choice shall serve as examples to illustrate how manipulations of time-dependent valuation of outcomes or subjective time perception can influence intertemporal choice behaviour and, ultimately, reduce temporal discounting: episodic future thinking (EFT) effects³⁵ and the date/delay effect³⁶ (Fig. 1).

EFT is defined as the ability to simulate experiences of one's own future.³⁷ It plays a beneficial role in different cognitive and emotional functions including decision making, emotion regulation and spatial navigation.³⁸ In the context of intertemporal choice, EFT has been used to modulate discounting behaviour. In typical study designs, participants are first instructed to imagine and describe (positive) personal events at different points of time in the future. Subsequently, they complete a temporal discounting task in which the larger, later reward of each trial is cued by the self-generated event referring to that delay.³⁹ Although concrete induction methods of EFT (e.g., instructions) and control conditions (e.g., standard task without cues or cues referring to present/recent instead of future events) vary between studies, EFT has been shown to reduce temporal discounting relatively consistently and with moderate effect sizes.⁴⁰

While the effectiveness of EFT in reducing temporal discounting is well documented, the underlying psychological mechanisms remain to be fully understood. Noël et al. and Rung and Madden summarize several possible and interconnected explanations.⁴¹

First, EFT effects can be interpreted in terms of construal-level theory,⁴² according to which near future events are mentally represented as lower-level construals (more concrete) whereas events fur-

³³ Zauberman et al. 2009.

³⁴ Takahashi 2005, also see Dehaene 2003, who found that firing patterns of >number neurons< follow the Weber-Fechner law.

³⁵ Peters & Büchel 2010.

³⁶ LeBoeuf 2006; Read et al. 2005.

³⁷ Atance & O'Neill 2001.

³⁸ Schacter et al. 2017.

³⁹ Peters & Büchel 2010.

⁴⁰ See Rung & Madden 2018 and Scholten et al. 2019 for reviews.

⁴¹ Noël et al. 2017 and Rung & Madden 2018/2019.

⁴² Trope & Liberman 2003.

ther into the future are represented as higher-level construals (more abstract). While this explains the so-called present bias, which puts a 'premium' on rewards received immediately,⁴³ in EFT, future rewards are rendered relatively more concrete and thus gain a comparable construal-level as immediate rewards, which facilitates option comparison and reduces temporal discounting.⁴⁴ This explanation is supported by the idea that better visualization abilities moderate EFT⁴⁵ and might also be linked to altered time perception (see below). Second, EFT might evoke an expanded temporal horizon,⁴⁶ which means that the time perspective in considering consequences of one's behaviour is shifted towards the future, possibly moderated by working memory capacity.⁴⁷ However, a recent study did not find extension of temporal horizon to be a mediator of EFT.⁴⁸ Third, given that humans' time sensitivity is often inadequate and strongly susceptible to manipulation,⁴⁹ time perception rather than time horizon may play an important role in EFT. For instance, it has been suggested that EFT may lead to enhanced attention to episodic events rather than time.⁵⁰ Thus, it could elicit lower sensitivity for duration until receipt of a larger, later reward, ultimately leading to less temporal discounting. Fourth, EFT manipulations that require the imagination of positive future events may enhance positive mood and thus indirectly strengthen cognitive control and working memory mechanisms.⁵¹ In this case, the possibly moderating role of positive valence of EFT serves as important evidence.⁵² Fifth, it has been argued that demand characteristics may drive EFT effects.⁵³ However, recent studies provided evidence against this notion.⁵⁴

In sum, several non-exclusive mechanisms of EFT have been discussed and it is of course conceivable that several of these might combine to contribute to EFT effects. For instance, EFT, mediated by

⁴³ Trope & Liberman 2003.

⁴⁴ H. Kim et al. 2013, Rung & Madden 2018 and Yi et al. 2017.

⁴⁵ Peters & Büchel 2010.

⁴⁶ Lin & Epstein 2014 and Snider et al. 2016.

⁴⁷ Lin & Epstein 2014.

⁴⁸ Rung & Madden 2019.

⁴⁹ Ebert & Prelec 2007.

⁵⁰ Radu et al. 2011, Rung & Madden 2019.

⁵¹ Noël et al. 2017.

⁵² Liu et al. 2013.

⁵³ Rung & Madden 2018.

⁵⁴ Rung & Madden 2019 and Stein et al. 2018.

working memory mechanisms, may lead to more concrete positive future representations, which at the same time shift attention away from duration and enhance positive mood, resulting in less discounting behaviour. Future studies are required to separate these contributions in order to elucidate the psychological mechanisms of EFT effects in temporal discounting.

Another robust effect in the experimental intertemporal choice literature directly refers to time framing, namely, the date/delay effect.⁵⁵ In studies examining this effect, participants complete a temporal discounting task in which time until receipt of rewards is specified either as units of delay (e.g., 0 days vs. 23 days) or as dates equivalent in delay from the date of assessment (e.g., 1st of March 2021 (= today) vs. 24th of March 2021). The date condition has repeatedly been found to lead to a lower degree of temporal discounting than the delay condition, with moderate effect size,⁵⁶ both in studies using between-⁵⁷ and within-subject designs.⁵⁸ Interestingly, Read et al. also found this manipulation to reduce hyperbolic discounting,⁵⁹ i.e., decision making became more consistent with ›rational‹ behaviour specifying delays in terms of specific dates.⁶⁰

Like EFT effects, the date/delay effect has often been replicated but the underlying mechanisms are not entirely clear, and several explanations have been offered.⁶¹ First, like in EFT effects, construal-level theory has been drawn upon to suggest that the more concrete date framing might lead to a lower-level construal than delay framing (especially in larger, later rewards). This could make construal levels of choice options more similar, facilitate option comparison and thereby reduce temporal discounting.⁶² Second, while delays and rewards are presented as continuous numeric variables, thus enabling a direct computational strategy (i.e., estimating exchange rates), presenting delays in terms of dates requires additional cognitive operations. For example, dates would either need to be transformed into delays to

⁵⁵ LeBoeuf 2006 and Read et al. 2005.

⁵⁶ See Rung & Madden 2018 and Scholten et al. 2019 for reviews.

⁵⁷ E.g. Read et al. 2005.

⁵⁸ E.g. Dshemuchadse et al. 2013, though see Lempert et al. 2016.

⁵⁹ Read et al. 2005.

⁶⁰ See DeHart & Odum 2015, however, who found that hyperbolic/hyperboloid models provided a better description of choices.

⁶¹ E.g. Rung & Madden 2018. For additional approaches see LeBoeuf, 2006 and Read et al. 2005.

⁶² H. Kim et al. 2013.

perform computations or people might generally change their cognitive strategy, for example, by deciding primarily based on reward magnitude, which is considered more important.⁶³ Third, and relatedly, the date condition may lead to attentional shifts away from durations towards the reward attributes and their difference,⁶⁴ which increases the relative importance of reward values and diminishes the role of delay for choices (lowered time sensitivity). Fourth, estimates of time duration may be reduced with dates in comparison to delays because they are more concrete.⁶⁵ Indeed, both Zauberman et al. and Jiang and Dai showed that dates led to lower subjective estimates of objective time than delays.⁶⁶ Fifth, risk perception, which is partly confounded with temporal discounting,⁶⁷ might also be altered by dates in comparison to delays: Uncertainty of receiving rewards might be higher with delays than with concrete dates,⁶⁸ possibly as a consequence of differences in time perception.⁶⁹ Sixth, subadditivity (i.e., subdivision of elements into units, which are evaluated individually instead of in their entirety) increases discounting behaviour⁷⁰ and might also contribute to the effect because only delays (and not dates) can directly be subdivided into units.⁷¹ Seventh, a date could be emotionally valued with stronger positive valence than a delay, as it focuses less on duration and is more concrete, eliciting positive mood and thereby causing less discounting (LeBoeuf, 2006; Lempert et al., 2016).

Again, some of these explanations could be integrated. As argued above, the date condition is more concrete than the delay condition, i.e., may reflect a lower-level construal. This in turn might lead to differential perceptions of durations and risk, general shifts in attention towards the rewards, or changes in strategy, causing higher valuation of rewards and less time sensitivity. Additionally, a more positive (or rather less negative) valence of duration and better mood might be elicited when dates are specified. As a result, people engage in less

⁶³ Read et al. 2005.

⁶⁴ Dshemuchadse et al. 2013 and Read et al., 2005.

⁶⁵ Jiang & Dai 2021, Lempert et al. 2016, Read et al. 2005 and Zauberman et al. 2009.

⁶⁶ Zauberman et al. 2009 and Jiang & Dai 2021.

⁶⁷ Lopez-Guzman et al. 2018.

⁶⁸ DeHart & Odum 2015, Jiang & Dai 2021.

⁶⁹ Jiang & Dai 2021.

⁷⁰ Read 2001.

⁷¹ DeHart & Odum 2015, however, see LeBoeuf 2006: Experiment 5 and Read et al. 2005: Experiment 1.

temporal discounting. Future studies will need to examine whether such an integrative explanation or only specific parts of it apply.

It should be noted that both the EFT and the date/delay effect may have direct implications on our daily lives. For example, they could be used in persuasion strategies in politics (e.g., conveying a better picture and creating better acceptance of energy-efficient measures by inducing EFT or speaking of specific dates) or offers of investments and saving schemes (e.g., engaging in EFT of a wealthy future, referring to specific dates on which profits are returned).⁷² Importantly, given that temporal discounting is increased in a number of mental disorders,⁷³ these manipulations could possibly be integrated into psychotherapy settings in order to reduce impulsive choice behaviours. For instance, EFT has been shown to be an effective intervention in reducing both obesity⁷⁴ and cigarette smoking.⁷⁵

In conclusion, EFT effects and the date/delay effect represent two important experimental effects in the intertemporal choice literature that illustrate how aspects of time influence temporal discounting behaviour. These manipulations not only improve our understanding of the cognitive and neural mechanisms underlying intertemporal choice but may also be used to optimize strategies in daily life and therapeutic contexts.

4. Intertemporal Choice, Temporal Processing and Impulsivity

Intertemporal choice and the role of temporal processes in temporal discounting has not only been studied using experimental or framing manipulations. There also exists a large literature on individual differences in intertemporal choice.⁷⁶ In this section, we turn to a particular trait correlate of temporal discounting, namely impulsivity. Specifically, we will first turn to time-relevant aspects of impulsivity and will then explore the role that sensitivity to temporal information might play in the often-reported relationship between intertemporal choice and impulsivity.

⁷² LeBoeuf 2006 and Read et al. 2005.

⁷³ Amlung et al. 2019.

⁷⁴ Sze et al. 2017.

⁷⁵ Stein et al. 2016/2018.

⁷⁶ For review, see Keidel et al. 2021.

Impulsivity is a broad and rather poorly specified term that has been defined most generally as »a predisposition toward rapid, unplanned reactions to internal or external stimuli without regard to the negative consequences of these reactions to the impulsive individual or to others«. ⁷⁷ Theoretical criticisms of the very construct of impulsivity aside, ⁷⁸ impulsivity is widely studied in psychology, psychiatry and neuroscience. Impulsive symptoms, traits and behaviours can be assessed with clinical diagnostic instruments, psychometric personality inventories and experimental tasks, respectively. ⁷⁹ However, correlations between these different levels of measurement are typically low, ⁸⁰ likely reflecting the heterogeneity of the impulsivity concept. ⁸¹

Theoretical models of impulsivity generally accept the heterogeneity of the construct, but differ with regards to the number and nature of dimensions they postulate. For example, in one of the historically most influential conceptualisations and assessments of impulsivity, Barratt distinguished between attentional impulsiveness (such as difficulties in focussing on a task at hand and suppressing distracting thoughts), motor impulsiveness (such as acting on the spur of the moment or having an inconsistent lifestyle) and non-planning impulsiveness (such as a low tendency to think ahead carefully or engage in challenging mental tasks), resulting in the Barratt Impulsiveness Scale. ⁸² Another widely adopted psychometric model and questionnaire is the UPPS Impulsive Behavior Scale by Whiteside and Lynam. ⁸³ The original UPPS contains four dimensions, namely, negative urgency (the tendency to act rashly during intense negative affect), lack of premeditation (the tendency to act without thinking), lack of perseverance (the tendency to be distracted and give up on boring or difficult tasks) and sensation seeking (the tendency to seek new and exciting experiences). Lynam et al. later added the positive urgency dimension (the tendency to act rashly during intense positive affect), thereby creating the UPPS-P Impulsive Behavior Scale which

⁷⁷ Moeller et al. 2001: 1784.

⁷⁸ Strickland & Johnson 2021.

⁷⁹ MacKillop et al. 2016 and Sharma et al. 2014.

⁸⁰ Aichert et al. 2012 and Cyders & Coskunpinar 2011.

⁸¹ Evenden 1999.

⁸² Patton et al., 1995.

⁸³ Whiteside & Lynam 2001.

has since been used and validated in numerous studies.⁸⁴ It should be noted that other models and measurements of impulsivity have also been proposed.⁸⁵

Findings from intertemporal choice tasks are often interpreted within an impulsivity framework in mind,⁸⁶ and temporal discounting is sometimes explicitly regarded as a behavioural measure of impulsivity in both humans and non-human animals.⁸⁷ Evidence for this interpretation comes *inter alia* from large-scale studies of healthy individuals, showing that individual differences in self-report, psychometric impulsivity are consistently albeit weakly associated with temporal discounting.⁸⁸ Further support comes from studies of patients with clinical disorders that are known to involve impulsive symptoms, such as drug abuse or attention deficit hyperactivity disorder (ADHD). Numerous studies have provided evidence that those disorders are reliably characterised by increased temporal discounting.⁸⁹

Of particular relevance to this chapter is the observed association between intertemporal choice and impulsivity, albeit of small magnitude,⁹⁰ which has been explained theoretically with reference to models of time processing.⁹¹ Specifically, both temporal discounting and impulsive actions may be interpreted as the choice of immediate, smaller gratification (with negative consequences) over delayed, larger gratification.⁹² Indeed, the ability to delay gratification in time has been considered to be a hallmark feature of self-control,⁹³ and impulsive behaviour is often conceived of as the opposite of self-control.⁹⁴

To illuminate the role of time processing in impulsivity and in the relationship between impulsivity and temporal discounting, theoretical models of how objective time flow is subjectively perceived should first be considered. Several such models exist.⁹⁵ An influential, early model is that by Treisman, who proposed an ›internal clock

⁸⁴ Lynam et al. 2006.

⁸⁵ Dickman 1990 and Eysenck et al. 1985.

⁸⁶ Peters & Büchel 2011.

⁸⁷ Hamilton et al. 2015 and Scholten et al. 2019.

⁸⁸ E.g. Keidel et al. 2022 and MacKillop et al. 2016.

⁸⁹ Amlung et al. 2019.

⁹⁰ Keidel et al. 2021.

⁹¹ Paasche et al. 2019 and Wittmann & Paulus 2008.

⁹² Paasche et al. 2019.

⁹³ Mischel et al. 1989.

⁹⁴ Rachlin & Green 1972.

⁹⁵ For review see Matthews & Meck 2016.

mechanism based on the existence of pacemaker and accumulator units.⁹⁶ The pacemaker produces a series of pulses travelling along a pathway at a constant rate. Their number is recorded by a counter, or accumulator, with the number of pulses corresponding to temporal duration. An additional component is the attentional gate, with pulses counted only when attention is oriented towards the passing of time. Importantly, time sensitivity is not rigid but instead is subject to external and internal influences;⁹⁷ anecdotally, it is well known to us that time seems to pass very quickly when having fun but seems to extend endlessly in a state of boredom. These state influences on temporal processing can be explained by two mechanisms.⁹⁸ First, the production rate of the pacemaker has been postulated to increase with heightened arousal, leading to a greater number of pulses produced and thus the experience of a longer duration of time. Second, increased attention to time (as may happen in boredom, but which is unlikely to happen when one is engaged in a fun, attention consuming task) is similarly postulated to lead to a greater number of produced pulses and, thereby, a longer perceived duration.

Of relevance to the present discussion, it has been proposed that impulsive behaviour may result from altered time processing, specifically from an overestimation of the flow of time.⁹⁹ Thus, for a highly impulsive individual, a given time interval would subjectively feel longer than for a less impulsive person. For that reason alone, highly impulsive individuals may be expected to display greater temporal discounting, as a given delay interval would represent a higher cost against which the reward must be traded off than for those with less contracted time perception.

Evidence in support of this hypothesis comes from a number of sources. First, otherwise healthy individuals with higher levels of impulsivity have been observed to overestimate the duration of time intervals compared to less impulsive individuals.¹⁰⁰ Second, sleep deprivation, which has been shown to transiently induce impulsive behaviours and disinhibition,¹⁰¹ has been found to cause both increa-

⁹⁶ Treisman 1963.

⁹⁷ Droit-Volet & Meck 2007.

⁹⁸ Wittmann & Paulus 2008.

⁹⁹ Wittmann & Paulus 2008.

¹⁰⁰ Lawrence & Stanford 1998.

¹⁰¹ Anderson & Platten 2011, Kumari & Ettinger 2020 and Meldrum et al. 2015.

sed temporal discounting and alterations in timing task performance suggestive of faster subjective time experience.¹⁰² Finally, patients with borderline personality disorder, drug dependence or frontal lobe damage show both impulsive behaviours and faster subjective time processing, and often associations between the two.¹⁰³

In conclusion, there is evidence of interrelationships between heightened impulsivity, faster subjective experience of time and increased temporal discounting. Put simply, these associations may be interpreted to suggest that impulsive individuals show enhanced temporal discounting due to a faster subjective sense of time flow. However, it should be noted that the relationship between time processing and impulsivity is likely to be complex and has not yet been fully understood. Specifically, the reasons for faster subjective time experience in people with high levels of impulsivity remain unclear, with both a faster pacemaker and enhanced attention to time as well as interactions amongst them postulated as explanations, whereas impaired inner clock mechanisms per se are considered less likely.¹⁰⁴ It should also be noted that the overlap between psychometric, self-report impulsivity and temporal discounting is of small magnitude,¹⁰⁵ suggesting that substantial variance in either measure remains unexplained by a postulated, shared alteration in temporal processing sensitivity.

5. Conclusions and Outlook

In this chapter, we focussed on intertemporal choice as a manifestation of temporal cognition. In particular, we introduced key methodological approaches and concepts in the study of intertemporal choice and provided an overview of both experimental and individual differences-based findings that relate intertemporal choice to aspects of temporal information processing.

With regard to experimental effects, we discussed framing effects in which temporal information processing is experimentally manipulated, namely, the episodic future thinking (EFT) and date/delay

¹⁰² Reynolds & Schiffbauer 2004.

¹⁰³ Berlin et al. 2004, Berlin & Rolls 2004 and Wittmann et al. 2007.

¹⁰⁴ Wittmann & Paulus 2008.

¹⁰⁵ Keidel et al. 2021.

manipulations. The effects that result from those manipulations underscore the importance of temporal information processing in intertemporal choice. Specifically, both detailed thinking about specific, often positive future events and a change from the common expression of a reward delay as a time period to a specific date in the future cause reduced temporal discounting. These effects are well replicated in the experimental literature¹⁰⁶ and thus present an ideal starting point for further research into processes of temporal cognition that may underlie temporal discounting. Specifically, the specific psychological mechanisms underlying these effects remain unclear. Whilst a number of hypotheses have been put forward (see above), it remains unclear how exactly temporal discounting is affected by EFT and the date/delay manipulation—both at the psychological and the neural level.

Following our discussion of group-level effects of temporal manipulations on intertemporal choice, we summarised evidence that individual differences in temporal discounting, impulsivity and their association may to some extent be explained by inter-individual variation in sensitivity to time. Specifically, we followed the argument that both temporal discounting and impulsive behaviour may be viewed as choosing immediate, smaller gratification over delayed, larger gratification, and outlined the hypothesis that impulsive actions may result from an overestimation of durations. Within this framework, a highly impulsive person would experience a given time interval as longer than a less impulsive individual, which could underlie the observed, greater temporal discounting in highly impulsive individuals.

A number of caveats remain in this area of research. First and foremost, impulsivity is increasingly recognised as a poorly defined, heterogeneous construct,¹⁰⁷ necessitating careful consideration of which aspect of impulsivity is being investigated. Second, associations between impulsivity and intertemporal choice tend to be of only small magnitude, requiring samples with appropriate statistical power.¹⁰⁸ Third, the cognitive and neural mechanisms that underlie these associations remain to be characterised in more detail. More research is needed in this area.

¹⁰⁶ Rung & Madden 2018.

¹⁰⁷ Strickland & Johnson 2021.

¹⁰⁸ Keidel et al. 2021.

To conclude, intertemporal choice is a decision-making task that yields the widely reproducible, cross-species phenomenon of temporal discounting. The task can be placed firmly in the temporal cognition literature, and both group-level experimental effects and individual differences studies point to a substantial role of temporal perceptual and cognitive processes in intertemporal choice. More refined characterisation of the cognitive and neural mechanisms that underlie these findings is required in the future.

References

- Aichert, Désirée S., Wöstmann, Nicola M., Costa, Anna, Macare, Christine, Wenig, Johanna R., Möller, Hans-Jürgen, Rubia, Katya & Ettinger, Ulrich (2012): Associations between trait impulsivity and prepotent response inhibition, in *Journal of Clinical and Experimental Neuropsychology* 34(10), 1016–1032. <https://doi.org/10.1080/13803395.2012.706261>
- Ainslie, George (1975): Specious reward: A behavioral theory of impulsiveness and impulse control, in *Psychological Bulletin* 82(4), 463–496. <https://doi.org/10.1037/h0076860>
- Amlung, Michael, Marsden, Emma, Holshausen, Katherine, Morris, Vanessa, Patel, Herry, Vedelago, Lana, Naish, Katherine R., Reed, Derek D. & McCabe, Randi E. (2019): Delay discounting as a transdiagnostic process in psychiatric disorders: A meta-analysis, in *JAMA Psychiatry* 76(11): 1176–1186. <https://doi.org/10.1001/jamapsychiatry.2019.2102>
- Anderson, Clare & Platten, Charlotte R. (2011): Sleep deprivation lowers inhibition and enhances impulsivity to negative stimuli, in *Behavioural Brain Research* 217(2), 463–466. <https://doi.org/10.1016/j.bbr.2010.09.020>
- Atance, Cristina M. & O'Neill, Daniela K. (2001): Episodic future thinking, in *Trends in Cognitive Sciences* 5(12), 533–539. [https://doi.org/10.1016/S1364-6613\(00\)01804-0](https://doi.org/10.1016/S1364-6613(00)01804-0)
- Ballard, Kacey & Knutson, Brian (2009): Dissociable neural representations of future reward magnitude and delay during temporal discounting, in *NeuroImage* 45(1), 143–150. <https://doi.org/10.1016/j.neuroimage.2008.11.004>
- Bartra, Oscar, McGuire, Joseph T. & Kable, Joseph W. (2013): The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value, in *NeuroImage* 76, 412–427. <https://doi.org/10.1016/j.neuroimage.2013.02.063>
- Berlin, Heather A. & Rolls, Edmund T. (2004): Time perception, impulsivity, emotionality, and personality in self-harming borderline personality disorder patients, in *Journal of Personality Disorders* 18(4), 358–378. <https://doi.org/10.1521/pe.2004.18.4.358>

- Berlin, Heather A., Rolls, Edmund T. & Kischka, Udo (2004): Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbito-frontal cortex lesions, in *Brain* 127(5), 1108–1126. <https://doi.org/10.1093/brain/awh135>
- Cho, Sang Soo, Pellicchia, Giovanna, Aminian, Kelly, Ray, Nicola, Segura, Barbara, Obeso, Ignacio & Strafella, Antonio P. (2013): Morphometric correlation of impulsivity in medial prefrontal cortex, in *Brain Topography* 26(3), 479–487. <https://doi.org/10.1007/s10548-012-0270-x>
- Choi, James J., Laibson, David, Madrian, Brigitte C. & Metrick, Andrew (2006): Saving for retirement on the path of least resistance. In McCaffery, Edward & Slemrod, Joel (eds.), *Behavioral Public Finance*. Russell Sage Foundation, 304–351. <http://www.jstor.org/stable/10.7758/9781610443852.14>
- Cyders, Melissa A. & Coskunpinar, Ayca (2011): Measurement of constructs using self-report and behavioral lab tasks: Is there overlap in nomothetic span and construct representation for impulsivity?, in *Clinical Psychology Review* 31(6), 965–982. <https://doi.org/10.1016/j.cpr.2011.06.001>
- Dehaene, Stanislas (2003): The neural basis of the Weber-Fechner law: A logarithmic mental number line, in *Trends in Cognitive Sciences*, 7(4), 145–147. [https://doi.org/10.1016/S1364-6613\(03\)00055-X](https://doi.org/10.1016/S1364-6613(03)00055-X)
- DeHart, William Brady & Odum, Amy L. (2015): The effects of the framing of time on delay discounting, in *Journal of the Experimental Analysis of Behavior* 103(1), 10–21. <https://doi.org/10.1002/jeab.125>
- Dickman, Scott J. (1990): Functional and dysfunctional impulsivity: Personality and cognitive correlates, in *Journal of Personality and Social Psychology*, 58(1) 95–102. <https://doi.org/10.1037/0022-3514.58.1.95>
- Droit-Volet, Sylvie & Meck, Warren H. (2007): How emotions colour our perception of time, in *Trends in Cognitive Sciences* 11(12), 504–513. <https://doi.org/10.1016/j.tics.2007.09.008>
- Dshemuchadse, Maja, Scherbaum, Stefan & Goschke, Thomas (2013): How decisions emerge: Action dynamics in intertemporal decision making, in *Journal of Experimental Psychology: General* 142(1), 93–100. <https://doi.org/10.1037/a0028499>
- Ebert, Jane E. J. & Prelec, Drazen (2007): The fragility of time: Time-insensitivity and valuation of the near and far future, in *Management Science* 53(9), 1423–1438. <https://doi.org/10.1287/mnsc.1060.0671>
- Evenden, John L. (1999): Varieties of impulsivity, in *Psychopharmacology* 146(4), 348–361. <https://doi.org/10.1007/PL00005481>
- Eysenck, Sybil B. G., Pearson, Paul R., Easting, G. & Allsopp, J. F. (1985): Age norms for impulsiveness, venturesomeness and empathy in adults, in *Personality and Individual Differences* 6(5), 613–619. [https://doi.org/10.1016/0191-8869\(85\)90011-X](https://doi.org/10.1016/0191-8869(85)90011-X)
- Frederick, Shane, Loewenstein, George & O'Donoghue, Ted (2002): Time discounting and time preference: A critical review, in *Journal of Economic Literature* 40(2), 351–401. <https://doi.org/10.1257/002205102320161311>

- Frost, Ralph, & McNaughton, Neil (2017): The neural basis of delay discounting: A review and preliminary model, in *Neuroscience and Biobehavioral Reviews* 79, 48–65. <https://doi.org/10.1016/j.neubiorev.2017.04.022>
- Guo, Yiqun, Chen, Zhiyi, & Feng, Tingyong (2017): The effect of future time perspective on delay discounting is mediated by the gray matter volume of vmPFC, in *Neuropsychologia* 102(2), 229–236. <https://doi.org/10.1016/j.neuropsychologia.2017.06.021>
- Hall, Robert E. (1988): Intertemporal substitution in consumption, in *Journal of Political Economy* 96(2), 339–357. <https://doi.org/10.1086/261539>
- Hamilton, Kristen R., Mitchell, Marci R., Wing, Victoria C., Balodis, Iris M., Bickel, Warren K., Fillmore, Mark, Lane, Scott D., Lejuez, Carl W., Littlefield, Andrew K., Luijten, Maartje, Mathias, Charles W., Mitchell, Suzanne H., Napier, T. Celeste, Reynolds, Brady, Schütz, Christian G., Setlow, Barry, Sher, Kenneth J., Swann, Alan C., Tedford, Stephanie E., ... Moeller, F. Gerard (2015): Choice impulsivity: Definitions, measurement issues, and clinical implications, in *Personality Disorders: Theory, Research, and Treatment* 6(2), 182–198. <https://doi.org/10.1037/per0000099>
- Ikuta, Toshikazu, del Arco, Alberto & Karlsgodt, Katherine H. (2018): White matter integrity in the fronto-striatal accumbens/ventral tract predicts impulsivity, in *Brain Imaging and Behavior*, 12(5): 1524–1528. <https://doi.org/10.1007/s11682-017-9820-x>
- Jiang, Jingya & Dai, Junyi (2021): Time and risk perceptions mediate the causal impact of objective delay on delay discounting: An experimental examination of the implicit-risk hypothesis, in *Psychonomic Bulletin & Review* 28, 1399–1412. <https://doi.org/10.3758/s13423-021-01890-4>
- Kable, Joseph W. & Glimcher, Paul W. (2007): The neural correlates of subjective value during intertemporal choice, in *Nature Neuroscience* 10(12), 1625–1633. <https://doi.org/10.1038/nn2007>
- Kable, Joseph W. & Glimcher, Paul W. (2010): An «as soon as possible» effect in human intertemporal decision making: Behavioral evidence and neural mechanisms, in *Journal of Neurophysiology* 103(5), 2513–2531. <https://doi.org/10.1152/jn.00177.2009>
- Kalenscher, Tobias & Pennartz, Cyriel M. A. (2008): Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making, in *Progress in Neurobiology* 84(3), 284–315. <https://doi.org/10.1016/j.pneurobio.2007.11.004>
- Keidel, Kristof, Ettinger, Ulrich, Murawski, Carsten & Polner, Bertalan (2022): The network structure of impulsive personality and temporal discounting, in *Journal of Research in Personality* 96, Article 104166. <https://doi.org/10.1016/j.jrp.2021.104166>
- Keidel, Kristof, Rramani, Qëndresa, Weber, Bernd, Murawski, Carsten & Ettinger, Ulrich (2021): Individual differences in intertemporal choice, in *Frontiers in Psychology* 12, Article 643670. <https://doi.org/10.3389/fpsyg.2021.643670>

- Kim, B. Kyu & Zauberman, Gal (2009): Perception of anticipatory time in temporal discounting, in *Journal of Neuroscience, Psychology, and Economics* 2(2), 91–101. <https://doi.org/10.1037/a0017686>
- Kim, Hyunji, Schnall, Simone & White, Mathew P. (2013): Similar psychological distance reduces temporal discounting, in *Personality and Social Psychology Bulletin* 39(8), 1005–1016. <https://doi.org/10.1177/0146167213488214>
- Kumari, Veena & Ettinger, Ulrich (2020): Controlled sleep deprivation as an experimental medicine model of schizophrenia: An update, in *Schizophrenia Research* 221, 4–11. <https://doi.org/10.1016/j.schres.2020.03.064>
- Lawrence, Jennifer B. & Stanford, Matthew S. (1998): Impulsivity and time of day: Effects on performance and cognitive tempo, in *Personality and Individual Differences* 26(2), 199–207. [https://doi.org/10.1016/S0191-8869\(98\)00022-1](https://doi.org/10.1016/S0191-8869(98)00022-1)
- LeBoeuf, Robyn A. (2006): Discount rates for time versus dates: The sensitivity of discounting to time-interval description, in *Journal of Marketing Research* 43(1), 59–72. <https://doi.org/10.1509/jmkr.43.1.59>
- Lempert, Karolina M., Johnson, Eli & Phelps, Elizabeth A. (2016): Emotional arousal predicts intertemporal choice, in *Emotion* 16(5), 647–656. <https://doi.org/10.1037/emo0000168>
- Lempert, Karolina M. & Phelps, Elizabeth A. (2016): The malleability of intertemporal choice, in *Trends in Cognitive Sciences* 20(1), 64–74. <https://doi.org/10.1016/j.tics.2015.09.005>
- Lin, Henry & Epstein, Leonard H. (2014): Living in the moment: Effects of time perspective and emotional valence of episodic thinking on delay discounting, in *Behavioral Neuroscience* 128(1), 12–19. <https://doi.org/10.1037/a0035705>
- Liu, Lei, Feng, Tingyong, Chen, Jing & Li, Hong (2013): The value of emotion: How does episodic prospection modulate delay discounting?, in *PLoS ONE* 8(11), Article e81717. <https://doi.org/10.1371/journal.pone.0081717>
- Loewenstein, George (1996): Out of control: Visceral influences on behavior, in *Organizational Behavior and Human Decision Processes* 65(3), 272–292. <https://doi.org/10.1006/obhd.1996.0028>
- Lopez-Guzman, Silvia, Konova, Anna B., Louie, Kenway & Glimcher, Paul W. (2018): Risk preferences impose a hidden distortion on measures of choice impulsivity, in *PLoS ONE* 13(1), Article e0191357. <https://doi.org/10.1371/journal.pone.0191357>
- Lynam, Donald R., Smith, G. T., Whiteside, S. P. & Cyders, Melissa A. (2006): The UPPS-P: Assessing five personality pathways to impulsive behavior (Technical Report). Purdue University.
- MacKillop, James, Weafer, Jessica, Gray, Joshua C., Oshri, Assaf, Palmer, Abraham & de Wit, Harriet (2016): The latent structure of impulsivity: Impulsive choice, impulsive action, and impulsive personality traits, in *Psychopharmacology* 233(18), 3361–3370. <https://doi.org/10.1007/s00213-016-4372-0>

- Malkoc, Selin A. & Zauberan, Gal (2006): Deferring versus expediting consumption: The effect of outcome concreteness on sensitivity to time horizon, in *Journal of Marketing Research* 43(4), 618–627. <https://doi.org/10.1509/jmkr.43.4.618>
- Malkoc, Selin A. & Zauberan, Gal (2019): Psychological analysis of consumer intertemporal decisions, in *Consumer Psychology Review* 2(1), 97–113. <https://doi.org/10.1002/arcp.1048>
- Mas-Colell, Andreu, Whinston, Michael D. & Green, Jerry R. (1995): *Microeconomic Theory*, Oxford: Oxford University Press.
- Matthews, William J. & Meck, Warren H. (2016): Temporal cognition: Connecting subjective time to perception, attention, and memory, in *Psychological Bulletin* 142(8), 865–907. <https://doi.org/10.1037/bul0000045>
- Meldrum, Ryan C., Barnes, J. C. & Hay, Carter (2015): Sleep deprivation, low self-control, and delinquency: A test of the strength model of self-control, in *Journal of Youth and Adolescence* 44(2): 465–477. <https://doi.org/10.1007/s10964-013-0024-4>
- Mischel, Walter (2014): *The Marshmallow Test: Mastering Self-Control*, Boston : Little, Brown and Company.
- Mischel, Walter, Shoda, Yuichi & Rodriguez, Monica L. (1989): Delay of gratification in children, in: *Science* 244(4907), 933–938. <https://doi.org/10.1126/science.2658056>
- Modigliani, Franco & Brumberg, Richard (1954): Utility analysis and the consumption function: An interpretation of cross-section data, in Kurihara, Kenneth K. (ed.), *Post Keynesian Economics*, New Brunswick: Rutgers University Press, 388–436. <https://doi.org/10.4324/9781315016849>
- Moeller, F. Gerard, Barratt, Ernest S., Dougherty, Donald M., Schmitz, Joy M. & Swann, Alan C. (2001): Psychiatric aspects of impulsivity, in *The American Journal of Psychiatry* 158(11), 1783–1793. <https://doi.org/10.1176/appi.ajp.158.11.1783>
- Noël, Xavier, Jaafari, Nematollah & Bechara, Antoine (2017): Addictive behaviors: Why and how impaired mental time matters?, in *Progress in Brain Research* 235, 219–237. <https://doi.org/10.1016/bs.pbr.2017.07.011>
- Noorani, Imran & Carpenter, Roger H. S. (2016): The LATER model of reaction time and decision, in *Neuroscience and Biobehavioral Reviews* 64, 229–251. <https://doi.org/10.1016/j.neubiorev.2016.02.018>
- Paasche, Cecilia, Weibel, Sébastien, Wittmann, Marc & Lalanne, Laurence (2019): Time perception and impulsivity: A proposed relationship in addictive disorders, in *Neuroscience and Biobehavioral Reviews* 106: 182–201. <https://doi.org/10.1016/j.neubiorev.2018.12.006>
- Patton, Jim H., Stanford, Matthew S. & Barratt, Ernest S. (1995): Factor structure of the Barratt Impulsiveness Scale, in: *Journal of Clinical Psychology* 51(6): 768–774. [https://doi.org/10.1002/1097-4679\(199511\)51:6<768::AID-JCLP2270510607>3.0.CO;2-1](https://doi.org/10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1)

- Peters, Jan & Büchel, Christian (2010): Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions, in: *Neuron* 66(1), 138–148. <https://doi.org/10.1016/j.neuron.2010.03.026>
- Peters, Jan & Büchel, Christian (2011): The neural mechanisms of inter-temporal decision-making: Understanding variability, in: *Trends in Cognitive Sciences* 15(5), 227–239. <https://doi.org/10.1016/j.tics.2011.03.002>
- Pleskac, Timothy J., Diederich, Adele & Wallsten, Thomas S. (2015): Models of decision making under risk and uncertainty, in Busemeyer, J. R., Wang, Z., Townsend J. T. & Eidels, A. (eds.), *The Oxford Handbook of Computational and Mathematical Psychology*, Oxford : Oxford University Press, 209–231. <https://doi.org/10.1093/oxfordhb/9780199957996.013.10>
- Rachlin, Howard & Green, Leonard (1972): Commitment, choice and self-control, in: *Journal of the Experimental Analysis of Behavior* 17(1), 15–22. <https://doi.org/10.1901/jeab.1972.17-15>
- Radu, Peter T., Yi, Richard, Bickel, Warren K., Gross, James J. & McClure, Samuel M. (2011): A mechanism for reducing delay discounting by altering temporal attention, in *Journal of the Experimental Analysis of Behavior* 96(3), 363–385. <https://doi.org/10.1901/jeab.2011.96-363>
- Read, Daniel (2001): Is time-discounting hyperbolic or subadditive?, in *Journal of Risk and Uncertainty*, 23(1), 5–32. <https://doi.org/10.1023/A:1011198414683>
- Read, Daniel, Frederick, Shane, Orsel, Burcu & Rahman, Juwaria (2005): Four score and seven years from now: The date/delay effect in temporal discounting, in *Management Science*, 51(9): 1326–1335. <https://doi.org/10.1287/mnsc.1050.0412>
- Read, Daniel & Scholten, Marc (2017): Future-oriented decisions: Intertemporal choice, in Ranyard, R. (ed.), *Economic Psychology*, Oxford: Wiley-Blackwell, 35–50. <https://doi.org/10.1002/9781118926352.ch3>
- Reynolds, Brady & Schiffbauer, Ryan (2004): Measuring state changes in human delay discounting: An experiential discounting task, in *Behavioural Processes* 67(3): 343–356. <https://doi.org/10.1016/j.beproc.2004.06.003>
- Rung, Jillian M. & Madden, Gregory J. (2018): Experimental reductions of delay discounting and impulsive choice: A systematic review and meta-analysis, in *Journal of Experimental Psychology: General* 147(9), 1349–1381. <https://doi.org/10.1037/xge0000462>
- Rung, Jillian M. & Madden, Gregory J. (2019): Demand characteristics in episodic future thinking II: The role of cues and cue content in changing delay discounting, in *Experimental and Clinical Psychopharmacology* 27(5): 482–495. <https://doi.org/10.1037/pha0000260>
- Schacter, Daniel L., Benoit, Roland G. & Szpunar, Karl K. (2017): Episodic future thinking: Mechanisms and functions, in *Current Opinion in Behavioral Sciences* 17, 41–50. <https://doi.org/10.1016/j.cobeha.2017.06.002>

- Scholten, Hanneke, Scheres, Anouk, de Water, Erik, Graf, Uta, Granic, Isabela & Luitjen, Maartje (2019): Behavioral trainings and manipulations to reduce delay discounting: A systematic review, in *Psychonomic Bulletin and Review* 26(6), 1803–1849. <https://doi.org/10.3758/s13423-019-01629-2>
- Schüller, Canan Beate, Kuhn, Jens, Jessen, Frank & Hu, Xiaochen (2019): Neuronal correlates of delay discounting in healthy subjects and its implication for addiction: An ALE meta-analysis study, in *American Journal of Drug and Alcohol Abuse* 45(1), 51–66. <https://doi.org/10.1080/00952990.2018.1557675>
- Sharma, Leigh, Markon, Kristian E. & Clark, Lee Anna (2014): Toward a theory of distinct types of »impulsive« behaviors: A meta-analysis of self-report and behavioral measures, in *Psychological Bulletin* 140(2), 374–408. <https://doi.org/10.1037/a0034418>
- Shoda, Yuichi, Mischel, Walter & Peake, Philip K. (1990): Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions, in *Developmental Psychology* 26(6), 978–986. <https://doi.org/10.1037//0012-1649.26.6.978>
- Snider, Sarah E., LaConte, Stephen M. & Bickel, Warren K. (2016): Episodic future thinking: Expansion of the temporal window in individuals with alcohol dependence, in *Alcoholism: Clinical and Experimental Research* 40(7), 1558–1566. <https://doi.org/10.1111/acer.13112>
- Stein, Jeffrey S., Tegge, Allison N., Turner, Jamie K. & Bickel, Warren K. (2018): Episodic future thinking reduces delay discounting and cigarette demand: An investigation of the good-subject effect, in *Journal of Behavioral Medicine* 41(2), 269–276. <https://doi.org/10.1007/s10865-017-9908-1>
- Stein, Jeffrey S., Wilson, A. George, Koffarnus, Mikhail N., Daniel, Tinuke Oluyomi, Epstein, Leonard H. & Bickel, Warren K. (2016): Unstuck in time: Episodic future thinking reduces delay discounting and cigarette smoking, in *Psychopharmacology* 233(21–22), 3771–3778. <https://doi.org/10.1007/s00213-016-4410-y>
- Strickland, Justin C., & Johnson, Matthew W. (2021): Rejecting impulsivity as a psychological construct: A theoretical, empirical, and sociocultural argument, in *Psychological Review* 128(2), 336–361. <https://doi.org/10.1037/rev0000263>
- Sutter, Matthias (2014): *Die Entdeckung der Geduld. Ausdauer schlägt Talent [The discovery of patience. Perseverance beats talent]*, Salzburg: Ecowin.
- Sze, Yan Yan, Stein, Jeffrey S., Bickel, Warren K., Paluch, Rocco A. & Epstein, Leonard H. (2017): Bleak present, bright future: Online episodic future thinking, scarcity, delay discounting, and food demand, in *Clinical Psychological Science* 5(4), 683–697. <https://doi.org/10.1177/2167702617696511>
- Takahashi, Taiki (2005): Loss of self-control in intertemporal choice may be attributable to logarithmic time-perception, in *Medical Hypotheses* 65(4), 691–693. <https://doi.org/10.1016/j.mehy.2005.04.040>
- Treisman, Michel (1963): Temporal discrimination and the indifference interval: Implications for a model of the »internal clock«, in *Psychological Monographs: General and Applied* 77(13), 1–31. <https://doi.org/10.1037/h0093864>

- Trope, Yaacov & Liberman, Nira (2003): Temporal Construal, in *Psychological Review* 110(3), 403–421. <https://doi.org/10.1037/0033-295X.110.3.403>
- Watts, Tyler W., Duncan, Greg J. & Quan, Haonan (2018): Revisiting the Marshmallow Test: A conceptual replication investigating links between early delay of gratification and later outcomes, in *Psychological Science* 29(7), 1159–1177. <https://doi.org/10.1177/0956797618761661>
- Whiteside, Stephen P. & Lynam, Donald R. (2001): The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity, in *Personality and Individual Differences* 30(4), 669–689. [https://doi.org/10.1016/S0191-8869\(00\)00064-7](https://doi.org/10.1016/S0191-8869(00)00064-7)
- Wittmann, Marc (2013): The inner sense of time: How the brain creates a representation of duration, in *Nature Reviews Neuroscience* 14(3), 217–223. <https://doi.org/10.1038/nrn3452>
- Wittmann, Marc, Leland, David S., Churan, Jan & Paulus, Martin P. (2007): Impaired time perception and motor timing in stimulant-dependent subjects, in *Drug and Alcohol Dependence* 90(2–3), 183–192. <https://doi.org/10.1016/j.drugalcdep.2007.03.005>
- Wittmann, Marc & Paulus, Martin P. (2008): Decision making, impulsivity and time perception, in *Trends in Cognitive Sciences* 12(1), 7–12. <https://doi.org/10.1016/j.tics.2007.10.004>
- Yi, Richard, Stuppy-Sullivan, Allison, Pickover, Alison & Landes, Reid D. (2017): Impact of construal level manipulations on delay discounting, in *PLoS ONE* 12(5), Article e0177240. <https://doi.org/10.1371/journal.pone.0177240>
- Zauberman, Gal, Kim, B. Kyu, Malkoc, Selin A. & Bettman, James R. (2009): Discounting time and time discounting: Subjective time perception and intertemporal preferences, in *Journal of Marketing Research* 46(4), 543–556. <https://doi.org/10.1509/jmkr.46.4.543>
- Zauberman, Gal & Urminsky, Oleg (2016): Consumer intertemporal preferences, in *Current Opinion in Psychology* 10, 136–141. <https://doi.org/10.1016/j.copsyc.2016.01.005>

Contributors

Marius Bartmann is Research Associate at the German Reference Centre for Ethics in the Life Sciences (DRZE), University of Bonn.

André Beauducel is Professor of Psychology at the Department of Psychology, University of Bonn.

Ulrich Ettinger is Professor of Psychology at the Department of Psychology, University of Bonn.

Charlotte Gauvry is Research and Teaching Associate at the Department of Philosophy, University of Bonn.

Bert Heinrichs is Professor of Philosophy at the Institute of Science and Ethics (IWE), University of Bonn, and at the Institute for Neurosciences and Medicine, Research Centre Juelich.

Michael Herzog is Professor of Psychophysics at the Laboratory of Psychophysics, École Polytechnique Fédérale de Lausanne (EPFL).

Christoph Horn is Professor of Philosophy at the Department of Philosophy, University of Bonn.

Kristof Keidel is PhD Student at the Department of Psychology, University of Bonn, and at the Department of Finance, Centre for Brain, Mind & Markets, University of Melbourne.

Carsten Murawski is Professor of Finance at the Department of Finance, Centre for Brain, Mind & Markets Laboratory, University of Melbourne.

Dieter Sturma is Professor of Philosophy at the Department of Philosophy and at the Institute of Science and Ethics (IWE), University of Bonn.

