Theory and History in the Human and Social Sciences

Davood Gozli
Jaan Valsiner *Editors*

Experimental Psychology

Ambitions and Possibilities

Springer

Theory and History in the Human and Social Sciences

Series Editor
Jaan Valsiner
Communication and Psychology
Aalborg University
AALBORG, Denmark

Theory and History in the Human and Social Sciences will fill in the gap in the existing coverage of links between new theoretical advancements in the social and human sciences and their historical roots. Making that linkage is crucial for the interdisciplinary synthesis across the disciplines of psychology, anthropology, sociology, history, semiotics, and the political sciences. In contemporary human sciences of the 21st there exists increasing differentiation between neurosciences and all other sciences that are aimed at making sense of the complex social, psychological, and political processes. Thus new series has the purpose of (1) coordinating such efforts across the borders of existing human and social sciences, (2) providing an arena for possible inter-disciplinary theoretical syntheses, (3) bring into attention of our contemporary scientific community innovative ideas that have been lost in the dustbin of history for no good reasons, and (4) provide an arena for international communication between social and human scientists across the World.

Davood Gozli • Jaan Valsiner
Editors

Experimental Psychology

Ambitions and Possibilities

Springer

Editors
Davood Gozli
University of Macau
Taipa, Macau S.A.R., China

Jaan Valsiner
Communication and Psychology
Aalborg University
AALBORG, Denmark

ISSN 2523-8663 ISSN 2523-8671 (electronic)
Theory and History in the Human and Social Sciences
ISBN 978-3-031-17052-2 ISBN 978-3-031-17053-9 (eBook)
https://doi.org/10.1007/978-3-031-17053-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Chapter 1
Finding the Place of Experimental Psychology: Introduction

Davood Gozli and Jaan Valsiner

The present volume puts together a diverse set of viewpoints, all of which are addressing fundamental concerns in psychological science. Each chapter on its own provides a pathway into thinking about experimental psychology, its promises and strengths, and its limits and potential risks. We read about the historical roots of—and early debates regarding—experimental psychology (Chap. 2), the role of concepts and operational definitions (Chap. 3 and 4), the problems of external validity (Chap. 5), the organization of behavior in an experiment (Chap. 6), the interpretation of behavior in an experiment (Chap. 7), and broader philosophical frameworks that could warrant or undermine a particular line of research (Chap. 8). Together, the chapters will equip the reader to think about experimental research in a balanced, complex, and cautious manner.

From the perspective of someone strongly attached to a particular method of research, there might be no apparent limit to the application of the method. When confronting various objects of study, instead of becoming aware of the limits of the method, the researcher only considers how objects are given within the framework of the method. Insisting that universal applicability of the method, researchers unknowingly *distort* and *limit* their view of the phenomena. It would be fair to ask whether such an attitude, such rigid application of method, which never raises the question of limitation and suitability, should be called "research," given that it is more akin to the exercise and extension of power within a domain. Initial statements, like "our method works here," are soon replaced by "*only* our method *ought* to work here!"

Regardless of the particular positions taken by the authors of this volume, what is more important is the very engagement with fundamental issues. The type of writing represented in the following chapters goes beyond the strict forms in which the

D. Gozli (✉)
University of Macau, Taipa, Macau S.A.R., China

J. Valsiner
Communication and Psychology, Aalborg University, AALBORG, Denmark

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9_1

findings of experimental psychology are presented and discussed. As such, they can serve as an opening for dialogue both for people within the field who have philosophical interests and for people outside of the field who are interested in a philosophical and critical engagement with experimental psychology. Questions regarding fundamental issues are rarely raised by experimental researchers, but when they are raised, we can soon see how current research practices are connected, and arise from, certain tacit answers—given without much reflection—to those fundamental questions. How should we think about the relationship between research participant and researchers? How is that relationship grounded in our view of self/other distinction? Is there any connection and resemblance between the experimental situations, the communication between the two parties involved, and the broader context of our social-political lives outside of the laboratory? Is it possible for experimental researchers to fixate on topics that are "artificial" inventions of the field, and if so, what is the way out of such fixations?

The chapters that follow begin with a historical contextualization. Russo Krauss (this volume) gives an account of the development of scientific psychology during the late nineteenth century, with particular reference to a disagreement between Wilhelm Wundt and Richard Avenarius (see also Araujo, 2016; Russo Krauss, 2019). While that debate is relevant to questions about the place of experimentation in psychology, Russo Krauss shows why it arises from disagreements over fundamental epistemological issues, including how we think about the self/other distinction. Wundt regarded experimentation as a necessary supplement to self-observation, whereas Avenarius saw experimentation as the primary method of investigation. For Avenarius, psychology must begin with the experience of others and must maintain a third-person perspective toward the phenomena under investigation. This is, for the most part, the stance adopted by contemporary psychological science, which provides an easy bridge to neuroscientific discourse about experience and behavior. This approach encourages a third-person view of even oneself. When one thinks about one's mental life or feelings in terms of brain activity, one is adopting such a third-person scientific perspective, which might side-step the view of oneself as a person among, and in relationships with, other people.

Wundt's insistence on first-person experience, and the necessity to ground psychological science in intelligible first-person experience, is instructive for contemporary psychologists. With Avenarius, the experimental situation involves one person observing another. With Wundt, one person is setting up the conditions to help another make observation. Thus, for Wundt, the subject of the experiment remains at the center, while the two persons work together to understand one kind of experience. The distinction between first- and third-person perspectives loses its significance for Wundt, and this is reinforced by the close positions the two adopt in relation to each other, i.e., as fellow researchers. Russo Krauss concludes by reminding us that the history of psychology should teach us about the importance of maintaining a relationship between philosophy and psychology.

Heath Matheson (Chap. 3) centers his discussion on concepts. Concepts, he argues, serve as the foundation of any science, just as they serve as the foundation of our social-political lives. Matheson draws attention to the sensorimotor, embodied basis of concepts. Pointing out the sensorimotor aspect of conceptual understanding involves acknowledging a set of constraints, which leads to the recognition of both intra-individual and inter-individual differences in understanding. It would not be surprising, Matheson argues, that crises about truth arise, if we appreciate the underlying psychological capacities that enable our evaluation of truth. His discussion shows the relevance of our topic with broader issues, e.g., political debates. He advocates moving from a realist approach to a pragmatist approach to science and taking conceptual disagreements seriously at the outset of our investigations. Matheson's contribution emphasizes concepts, largely postponing questions of methodology. It is, nonetheless, a useful demonstration of the connection between philosophy of science and psychological science. Whether or not we agree with Matheson, his argument shows that the ambitions and optimism of a psychological scientist must be grounded in some view of how science operates, why scientists disagree, and how they ought to address their disagreements.

Complementing Matheson's chapter, Roland Pfister (Chap. 4) shifts the center of discussion to methods. He points out two problematic tendencies in experimental psychology: (1) downplaying the role of tasks (i.e., method) in research and conversely (2) investigating tasks for their own sake. Downplaying the role of tasks will create a naïve sense of external validity for the research. For example, in a line of research related to human memory, not paying attention to experimental tasks can strategically replace the statement, "we are studying how participants remember these items in task X, Y, Z" with "we are studying *memory*!" At the same time, Pfister rightly points out, tasks can be over-emphasized, playing too large a role in motivating research, to a degree that the goals for which they were originally designed (i.e., knowledge about psychological phenomena) are forgotten. It might be difficult for some readers to see how these two tendencies can co-exist in the same research community, given that they seem to contradict each other. The two tendencies, however, represent two styles of engagement adopted at different times and in different settings.

When discussing research with members of their own field (at conferences, in articles published in specialty journals), discussion is focused heavily on methods and tasks, such that the tasks themselves—their limitations and the uncertainties regarding their use—creates the motivation to continue research. By contrast, when researchers take a position to talk with general audiences (writing articles in more popular journals, writing grant proposals), they set aside concerns about the tasks and refer to their research as if it has perfect external validity. Consequently, we have researchers who switch from paying too little attention to their method (highlights the aims of their research with unrealistic optimism) to paying exclusive attention to their method (forgetting the aim of their research). Pfister proposes a solution. He argues that introducing variety in operational definitions can help overcome this problem. If one psychological concept, such as rule violation, is defined in three different ways, it becomes difficult to ignore the differences in those three definitions, insofar as we remain interested in the concept that unifies them, and it becomes easier to remember why those (different) methods were constructed in the first place. Pfister also argues in favor of paying close attention to the details of action, as well as arguing in favor of continued critical engagement with psychological research.

From a more critical standpoint, Carol Ting (Chap. 5) points out the inevitable tradeoff between, on the one hand, the ease of identification, and on the other hand, the artificiality of the behavior under investigation. In general, the easier it is for researchers to categorize behavior, the more artificial the situation has become. Ting chooses the case of dishonesty to clarify the tradeoff. Dishonesty has a social-relational dimension, which is essential to it; it is dynamic and continually responsive to what is going on. Some might continue a lie for a length of time, trying to adjust, extend, and elaborate the lie while remembering what the recipients of the lie already know. The meaning and consequences of dishonesty change with context. These consequences might include social harm, harm to one's self-image or social status, punishment, and harm to the existing trust in communities and institutions which operate on the basis of trust. Ting reviews several approaches to the experimental study of dishonesty and shows what they are missing and what they systematically exclude, by placing "dishonesty" in an experimental setting. For instance, experiments with "dishonesty" often take place in rather contrived situations, with researchers going out of their way to ensure anonymity of the participants. The experimenters isolate the "dishonest act" as much as possible from the social context, and they encourage its occurrence, in many cases, by tying it to monetary reward. Ting's argument is applicable to other areas of experimental research, whenever there is discrepancy between the operational definition of the behavior and its meaning and consequence outside of the laboratory (Gozli, 2019). It also has implication for the discussion of how concepts differ within and outside of experimental settings. Beginning with Ting's arguments, researchers could see how they could apply a similar analysis to other fields of research.

In Chap. 6, Hazeltine, Dykstra, and Schumacher (this volume) trace the development of the notion of task in recent decades in cognitive psychology. They observe that a reliance on stimulus-response (SR) associations, or "task set," is insufficient for understanding the existing evidence. A better understanding of task, they argue, is in terms of an organization of SR associations. Unfortunately, the word "response" is somewhat ambiguous, and it is not easy to see how much meaning can or should be attached to a response when we describe behavior in an experimental situation. Even though the whole (tasks) determines the meaning of the part (response), the part has a role in determining the character of the whole. What we usually mean by "response" is closer to the physical description of the movement, which risks neglecting the character of the whole (task) and the meaning assigned to individual responses. Hazeltine et al. propose that the organization of the behavior is maintained by internal representation, rather than the environment, which is why they propose the idea of "task files." The chapter includes an experiment that demonstrates how switch costs relate to task structures. The authors provide a visual representation of the association (grouping) between individual responses, which is an effective demonstration of the limits of the SR approach to understanding tasks. Moreover, their demonstration shows the limits of relying on response times as a one-dimensional variable.

Paying close attention to several lines of research, and especially attending to their methods, Davood Gozli (Chap. 7) addresses the problem of interpretation within the experimental situation. How research participants make sense of the experimental situation can change without being noticed by the researchers. Consequently, the categories researchers use in describing and explaining the findings of an experiment may not fit what happens in the experiment. For instance, while researchers assume they are comparing two conditions, which are equivalent except for the experimental manipulation, in fact there could be two qualitatively different conditions, each associated with a different set of task rules, strategies, and experiences. The reason why this major blind spot tends to go unnoticed by experimental researchers is because they impose a set of fixed criteria, as part of designing the experiment, for describing the experimental situation, while leaving out any explicitly discussion of those criteria (Gozli, 2017, 2019; see also Mammen, 2017).

In his remarkably thorough argument, Aaro Toomela (Chap. 8) turns to basic questions about science, epistemology, and methodology. Central to the chapter is Toomela's position that research methods cannot be applied without at the same time considering research questions. A method is associated with theoretical presuppositions about the object of study, whether or not researchers are aware of those presuppositions, and those could be incompatible with what is being studied. Toomela addresses along the way questions about what science is, the relationship between methods, methodology, and knowledge, drawing on Aristotle, Francis Bacon, Lev Vygotsky, and others. Turning to psychology, Toomela challenges what is often taken as the basis of many operational definitions, namely, the correspondence between behavior and psychic phenomenon. Believing we understand psychic phenomena based on isolated behaviors in the lab is not only neglecting the structural and systemic nature of psychic processes but also carrying the prejudices of researchers into the field of research. Thus, rather than adding something new to our psychological knowledge, researchers end up offering a particular demonstration of their own prejudices. Toomela reminds us that "certain methods are absolutely necessary to construct scientific knowledge but not *all* knowledge achieved by using such methods is necessarily scientific." That is analogous to pointing out that not all grammatically correct sentences are meaningful. Applying the rules of grammar on its own is not a guarantee that we are making sense. Toomela's valuable chapter shows how the questions about experimental psychology are related to more fundamental questions about knowledge. Interested readers could then explore the argument in connection to the author's earlier works on science and methodology (e.g., Toomela, 2007, 2019, 2020; Valsiner & Toomela, 2010).

Central themes and questions that emerge from the chapters include concepts and their treatment within the experimental situation, external validity, (mis)interpretation of behavior, the neglected organization of behavior—when we focus on isolated "responses," and unexamined theories of science that could keep researchers attached to a set of methods. A major strength in the chapters that follow comes from how they demonstrate the line of reasoning with a close connection to a particular research question. This is perhaps clearest in the case of Pfister (Chap. 4) and Ting (Chap. 5), who engage with experimental research on rule violation and dishonesty, respectively, and such close engagement is necessary for fleshing out an argument, demonstrating points of contact between the mainstream psychology and theoretical-critical psychology, namely, particular cases of research.

Alongside comparable recent contributions to the critique of psychological research (e.g., Lamiell & Slaney, 2020; Uher, 2021; Valsiner, 2017), we hope this volume stimulates further reflection and dialogue regarding experimental psychology, the place of methodology in psychological science, and a turn toward foundational questions, which would not only bring depth of understanding to our scientific thinking, but also relevance to the wider context of our social-political lives. Finally, we hope the present chapters remind students and researchers that psychological writing does not have to follow the strict conventions of empirical research (i.e., the well-known sequence of Introduction, Methods, Results, & Discussion). Thoughtful critique of empirical findings, with careful attention to their presuppositions and interpretations, presented in reflective forms of writing, could just as effectively open up new pathways of thought and enable genuine advancements in scholarship.

References

Araujo, S. F. (2016). *Wundt and the philosophical foundations of psychology: A reappraisal*. Springer.

Gozli, D. (2017). Behaviour versus performance: The veiled commitment of experimental psychology. *Theory & Psychology, 27*(6), 741–758.

Gozli, D. (2019). *Experimental psychology and human agency*. Springer.

Lamiell, J. T., & Slaney, K. L. (2020). *Problematic research practices and inertia in scientific psychology: History, sources, and recommended solutions*. Routledge.

Mammen, J. (2017). *A new logical foundation for psychology*. Springer.

Russo Krauss, C. (2019). *Wundt, Avenarius, and scientific psychology: A debate at the turn of the twentieth century*. Palgrave Macmillan.

Toomela, A. (2007). Culture of science: Strange history of the methodological thinking in psychology. *Integrative Psychological and Behavioral Science, 41*(1), 6–20.

Toomela, A. (2019). *The psychology of scientific inquiry*. Springer.

Toomela, A. (2020). Psychology today: Still in denial, still outdated. *Integrative Psychological and Behavioral Science, 54*(3), 563–571.

Uher, J. (2021). Problematic research practices in psychology: Misconceptions about data collection entail serious fallacies in data analysis. *Theory & Psychology, 31*(3), 411–416.

Valsiner, J. (2017). *From methodology to methods in human psychology*. Springer.

Valsiner, J., & Toomela, A. (Eds.). (2010). *Methodological thinking in psychology: 60 years gone astray?* Information Age Publishing.

Chapter 2
From Introspection to Experiment: Wundt and Avenarius' Debate on the Definition of Psychology

Chiara Russo Krauss

Aim of the Paper

This paper provides an account of the debate between Wilhelm Wundt and Richard Avenarius on the definition of psychology. It shows that – despite the fame of the former as the founder of experimental psychology – it was the latter who first defined this science on the basis of the experimental method. Moreover, the paper reconstructs how Avenarius elevated physiological experiment to the rank of a paradigm, using it to define not only psychology but also the relationships between this science and knowledge in general. In so doing, Avenarius elaborated a groundbreaking conception of psychology that anticipated several topics of later debate on this science. Finally, we will show how Avenarius' attention to the interactions between philosophy, psychology, and the concrete practice of science can still be instructive today.

Historical Background

Since the Scientific Revolution in the sixteenth to seventeenth centuries, the definition of science was based on two fundamental and interrelated requirements: the mathematization of knowledge and the adoption of the experimental method, which is one of the conditions for this mathematization. Accordingly, physics was regarded as the model for all other sciences. In the following centuries, as new branches of knowledge started to develop, the question of whether they could be considered

C. Russo Krauss (✉)
University of Naples "Federico II", Naples, Italy
e-mail: chiara.russokrauss@unina.it

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9_2

sciences in the proper sense of the word overlapped with the question of the applicability of mathematics and experimentation in their fields.

In the case of psychology, this approach is evident in the work that became a reference point for the entire debate on the establishment of this science in the nineteenth century: Immanuel Kant's Preface to the *Metaphysical Foundations of Natural Science*. In it, Kant stated "that in any special doctrine of nature there can be only as much proper science as there is mathematics therein" (Kant, 1786/2004, 6). Consequently, for Kant psychology, "the empirical doctrine of the soul" could not aspire to "the rank of a properly so-called natural science," for two reasons, because "mathematics is not applicable to the phenomena of inner sense and their laws" and because the method of "systematic art of analysis or experimental doctrine" does not fit with "inner observation," and "still less does another thinking subject suffer himself to be experimented upon to suit our purpose" (Kant, 1786/2004, 7).

In the nineteenth century, several thinkers tried to challenge Kant's negative answer to the question of the fate of scientific psychology.[1] Johann Friedrich Herbart believed that physics was possible thanks to its metaphysical foundation, i.e., the construction of the ideas of matter and force to make sense of the contradictions of experience and to provide the necessary basis for calculation. Consequently, he wanted to do the same for psychology, proposing a mathematized "psychical mechanics" based on the dynamics between "representations" that oppose each other like "forces" (Herbart, 1816/1901, 6).[2] However, despite Herbart's undeniable lasting influence, his attempt was soon abandoned as Gustav Theodor Fechner proposed a new psychophysical and experimental approach. Unlike Herbart, he believed that science should adhere to experience. For him, physics was based on outer observation, and psychology on inner observation. Therefore, Fechner believed that it was impossible to *directly* experience the connection between the physical and the psychical, since "no one can have an outer and inner perspective on the same thing at once" (Fechner, 1860, 5). Nonetheless, he admitted the possibility of experimentally investigating the *indirect* but lawful relationship between the two. Given the measurability of physiological data, the psychical phenomena could be expressed as a function of those data. On this basis, Fechner established "psychophysics as the exact science of the functional relationships between body and soul" (Fechner, 1860, 8).[3]

As we can see, Kant, Herbart, and Fechner, apart from their differences, were all convinced that physics was the model for all sciences and that psychology, in order to ascend to the rank of science, should imitate this model. However, during the

[1] For a brief account of the early debate on scientific, experimental psychology, see Kim, 2009. For a more comprehensive analysis of this debate in the nineteenth century, see Teo, 2005, 39–92, and Fahrenberg, 2015.

[2] For more details on Herbart, see Beiser, 2014, 89–141. About Herbart's psychology and his fortune, see Boudewijnse et al., 1999 and 2001.

[3] On Fechner see Heidelberger 2004. However, Fechner maintained some elements of Herbart approach (on this topic see Murray & Bandomir, 2001).

nineteenth century, a new attitude toward this issue began to develop. More and more thinkers started to distinguish the so-called *Geisteswissenschaften* (i.e., the "spiritual," "moral," or "human" sciences) from the natural sciences, stressing that the two had different objects, goals, and methods. The best example of this trend is Wilhelm Dilthey.[4] He criticized the "constructive psychology" that adopted the "explanatory" (*erklärende*) method of natural sciences, i.e., the reduction of phenomena to a set of simplest elements (such as atoms or sensations) and their laws of connection. For Dilthey, this method was suitable only for the knowledge of natural phenomena, since we experience them from the outside and separately, without access to their inner relationships. Conversely, in the case of the human phenomena that are the subject of spiritual sciences, we have a direct "understanding" (*Verstehen*) of the "living nexus" that intrinsically characterizes them. Thus, Dilthey supported a "descriptive psychology" based on the immediate living experience of the interconnectivity that is typical of spiritual phenomena. For him, such a psychology was not only a spiritual science but the very foundation of all spiritual sciences (see Dilthey, 1894/2010).

Wilhelm Wundt Between Introspection and Experiment

In textbooks, Wilhelm Wundt (1832–1920) is usually regarded as the father of experimental and physiological psychology.[5] While Fechner established psychophysics as the science that studies the relationship between mind and body, Wundt adopted Fechner's psychophysics to use as a method for the investigation of purely psychological phenomena. In the Preface to his seminal *Principles of Physiological Psychology*, Wundt explicitly stated his intention "to mark out a new domain of science," i.e., "the experimental treatment of *psychological* problems" (Wundt, 1874, III, emphasis mine). Moreover, in 1879 Wundt founded the Institute of Experimental Psychology in Leipzig (Germany) which – apart from being one of the first of its kind – became the Mecca of the new psychology, attracting students and researchers from all over the world. Finally, in his writings on the status of psychology, Wundt criticized the method of self-observation, comparing it to the attempt of Baron Munchausen to save himself from drowning in a swamp by pulling his own hair (Wundt, 1882/1906, 198).

---

[4] On Dilthey's critique of psychology, see Teo, 2005, 78–84, Hodges, 2000, 196–224.

[5] For a thorough analysis of Wundt's reception, especially in textbooks, see Fahrenberg, 2011, 125–130, and Fahrenberg, 2020, 218–264. Many accounts of Wundt's ideas in the histories of psychology were distorted by blatant errors and incomprehensions. However, in the last decades, some authors have produced works that correct these mistakes and provide a correct understanding of Wundt's conceptions, such as Rieber & Robinson, 2001, Araujo, 2016, and the already mentioned Fahrenberg, 2020. These works have been preceded by the researches of Kurt Danziger, who paved the way for all the recent investigations on Wundt (see Danziger, 1979, 1980a, b, 1987, 2001).

In view of this, the situation seems clear: Wundt is and should be considered a representative of the tendency that aimed at making psychology scientific (in the sense of physics), by rejecting introspection in favor of the new physiological-experimental method. However, on closer examination, this account of Wundt's position turns out to be false.

First of all, Wundt believed that psychology occupied a "mediating position between the natural and the spiritual sciences." Or, more precisely, he sided with Dilthey, who regarded it as merely "related" to the natural sciences but as the very "fundamental discipline of the spiritual sciences." The reason was that "every expression of the human spirit has its last cause in the elementary phenomena of the inner experience," which is the subject matter of psychology (Wundt, 1874, 4).[6] Indeed, his criticism of self-observation did not challenge the assumption that introspection is the fundamental method of psychology. Wundt rejected self-observation only insofar as it implies intention and effort, since attention dissolves the mental phenomenon we are trying to analyze. For him, "the more we strive to observe ourselves, the more certain we can be that we are observing nothing at all." So, instead of self-observation, he suggested that psychology should rely on "fortuitous inner perception" (Wundt, 1882/1906, 197–198).[7] Needless to say, he was aware that this "uncertain ground" was not sufficient to establish a science. And this is where the experiment came into play, allowing for a "deliberate renewal of inner processes" "under the same or voluntarily modified conditions." Hence, Wundt recognized the need for psychological experiments as "auxiliary means" to support the method of self-observation (Wundt, 1888a, 301–303).

Accordingly, Wundt distinguished three phases in the history of psychology. The first is the "physiological phase." In it, as Kant said, "inner experience is regarded as a field inaccessible to experimental method, and therefore to all exact investigation." Consequently, "the only task of experimental method is considered the investigtion of the physiological basis of the psychical." The second is the "psychophysical phase," represented by Fechner. In this phase, it is still held that "no experiment can be applied to purely psychical interactions." Still, assuming the "functional relationships" between physiological and psychical phenomena, the experimental method is extended to the investigation of the "psychophysical interactions" "between body and soul." Finally, in the third and purely "psychological phase," the "physical causes *no longer count as members of a functional relationship*, since, strictly speaking, such a relationship is possible only between members *of the same kind*, i.e. between physical and physical, or between psychical and psychical elements."

[6] The definition of psychology as the science that deals with the "immediate experience," which is often associated with Wundt, was introduced in a second phase of his career, as a direct reply to the new definition of psychology developed by Avenarius and Mach, and adopted by several of Wundt's pupils (see Russo Krauss, 2019, 113–117).

[7] The ambiguity of Wundt's position – who apparently rejects self-observation, only to admit it immediately afterward in the form of inner perception – created misunderstandings even then. Johannes Volkelt criticized Wundt's rebuttal of self-observation (Volkelt, 1887), thus pushing Wundt to reply, to explain that he was actually favorable to introspection (Wundt, 1888a).

Accordingly, the physical causes employed in physiological experiments "are regarded as auxiliary means to produce *psychical* processes at will" (Wundt, 1895, 172–174, emphasis mine).

It is clear from the above that Wundt's conception of psychology differed considerably from Fechner's. Wundt did not base experimental psychology on the functional relationship between physiological processes and mental activity, but even denied the existence of such a relationship. He agreed with Dilthey that "psychical connections and physical connections are different and incomparable" (Wundt, 1895, 155). In fact, Wundt reversed the meaning of the very concept of psychophysical parallelism. For him, it did not mean perfect coordination, even in the absence of causal connection, between the physical and the psychical events. Rather, parallelism meant to him that "there are phenomena on the physical side which have no counterpart in psychical elements, and vice versa, that there are features on the psychical side for which we cannot discover or presuppose any physical correlate phenomena" (Wundt, 1895, 253).

For Wundt, the object of psychology are inner experiences. Between these experiences there are purely psychological connections, which form a "psychical causality" that is independent of the parallel physiological activity. Therefore, the goal of psychology is the investigation of these purely psychological connections. Since the object and aim of psychology are purely psychological, the physiological-experimental method has only a subsidiary function. As Wundt writes, "True as it is that a deeper knowledge of nervous and cerebral functions can be beneficial to the understanding of psychical processes, this benefit is always possible *only insofar as it encourages a deeper psychological analysis*, otherwise, the physiology of the brain can cause *nothing but confusion*" (Wundt, 1880–1883, 483, emphasis mine).

The relative independence between physical and psychical processes and the consequent auxiliary role of experimental method in psychological investigation become clearer when one looks more closely at the theoretical framework that formed the basis for the research carried out in Wundt's laboratory. Wundt identified five moments in the process of human reaction to stimuli: "(1) the transmission from the sensory organ to the brain, (2) the entry into the field of consciousness [*Blicksfeld*], or perception, (3) the entry into the focus of attention [*Blickspunkt*] of attention, or apperception, (4) the time of the will, which is required to initiate the reaction movement in the central organ, (5) the transmission of the motor excitation thus produced to the muscles and the increase of energy in the latter" (Wundt, 1874, 727). For Wundt, the experimental-physiological method consisted of the deliberate production of these fivefold processes through the application of physiological stimuli to the subject of the experiment. Since the first and the last moments were physiological events, it was possible to calculate the time interval between them. By doing so, the experimenter obtained the measure of the duration of the three central *psychological* moments. However, the real object of the investigation was not the connection between the physiological stimulation and the psychological response. The object of the investigation were the central and purely psychological moments; moreover, despite being produced and calculated with the aid of the experimental-physiological method, their account was provided by an introspective description. The subject of the experiment was supposed to give a recollection of his mental contents during the process, distinguishing the three phases of perception, apperception, and voluntary act.[8]

So, even though psychology required the physiological experiment, for Wundt it was essentially independent of physiology, both in terms of its object of study and its methods. If this independence was already evident in the field of simple reactions to stimuli, it was even more so in the case of the more complex mental functions. For Wundt, the experimental method was useless for the investigation of these higher psychical phenomena, even as an auxiliary means. As he wrote, "psychology has two exact methods: the first is the experimental method, which serves for the analysis of simple psychical processes; and the second, the observation of the universally valid spiritual product, which serves for the investigation of higher psychical processes and developments" (Wundt, 1896, 28). For him, the latter method was typical of the other branch of psychology: the so-called *Völkerpsychologie* (literally, the "psychology of people," but more accurately would be "cultural psychology").[9] This psychology was supposed to deal with "all the spiritual phenomena connected with the life of people in community," such as language, myth, and culture in general (Wundt, 1888b, 2).

In view of the above, Wundt's statement about the position of psychology in relation to the natural and the spiritual sciences becomes clearer. Psychology is a spiritual science; indeed, it is the very basis of all spiritual sciences, for its field of investigation embraces all mental phenomena, the entire inner life, from the simplest sensations to the greatest cultural creations. But the simpler the phenomena, the more important is the physiological-experimental method, and the closer psychology gets to the natural sciences. This makes physiological psychology the "intermediary" between the natural and the spiritual sciences (Wundt, 1874, 4). However, Wundt stated that "such intermediate disciplines, for their very nature, have only a transitory value." He believed that once all the physiological investigations will be assigned to physiology, it will become apparent that the physiological-experimental method is just an aid for the investigation of purely psychological phenomena and "physiological psychology in the present sense *will no longer exist*" (Wundt, 1895, 230).[10]

[8] More information on Wundt's experimental method in Robinson, 2001.

[9] For a more detailed account of Wundt's *Völkerpsychologie*, see Araujo (2018) and Jüttemann (2006).

[10] Given the purpose of this paper, we are not considering the inner evolution of Wundt's thought. However, it can be shown that his opinion on the limits of the physiological approach changed over time, as he increasingly emphasized the need for a purely psychological investigation (see Russo Krauss, 2019; van Hoorn & Verhave, 1980).

Richard Avenarius and the Physiological Experiment as a Paradigm

Despite Wundt's enormous influence on the development of psychology in the nineteenth century, the rapid changes in the field soon relegated him to the role of the respected but outmoded father of the discipline. Over the years, even Wundt's own students began to distance themselves from him.[11] Richard Avenarius (1843–1896) was one of the first Wundtians to propose a conception of psychology radically different from Wundt's. In fact, Avenarius was not a psychology student from Wundt's laboratory, but an enthusiast of the new experimental psychology. He had studied philosophy at various German universities before settling in Leipzig, where in 1875 he defended his habilitation thesis before Moritz Drobisch, Max Heinze, and Wundt, who had just been appointed professor of philosophy after a career as a psychophysiologist.[12]

Wundt's appointment was the first, but not the last, time an experimental psychologist was nominated for a chair of philosophy. Traditional philosophers began to perceive such appointments as a threat. Thus, at the turn of the twentieth century there was an escalating debate about the place of psychology within philosophy departments. Given the opposition to experimental psychologists like Wundt, support from a philosopher like Avenarius was valuable.[13] Although Avenarius had a traditional philosophical training, he shared Wundt's idea that the new experimental-physiological psychology could help philosophy, and the theory of knowledge in particular. For this reason, the two began to cooperate. When Avenarius founded the journal *Vierteljahrsschrift für wissenschaftlichen Philosophie* (Quarterly for Scientific Philosophy) in 1877, Wundt agreed to serve on the editorial board. More importantly, Wundt helped Avenarius obtain the chair for philosophy at the Zurich University.[14]

However, after a series of articles published in the first issues of his journal, Avenarius took a 10-year hiatus from publishing. When he finally published the *Kritik der reinen Erfahrung* (Critique of pure experience) in 1888–1890, *Der menschliche Weltbegriff* (The human concept of the world) in 1891, and the series of articles *Bemerkungen über den Begriff des Gegenstandes der Psychologie* (Remarks on the concept of object of psychology) in 1894–1895, these works presented a conception of psychology that differed considerably from that of Wundt.

First of all, Avenarius rejected the traditional distinction between inner and outer experience that underlay the traditional concept of psychology. Kant, Fechner, Dilthey, and Wundt (to name only those we cited in the first paragraph) all assumed that there were two fundamentally distinct experiences, whatever name they used

[11] On the controversy between Wundt and his pupils, see Danziger (1979) and Russo Krauss (2019).

[12] For an account of Wundt's carreer before his appointment in Leipzig, see Diamond (2001).

[13] On the dispute between psychologists and traditional philosophers about academic appointments, see Ash, 1980; Kusch, 1991.

[14] For more details on the relationship between Avenarius and Wundt, see Russo Krauss, 2019.

for them: inner and outer sense, internal and external observation, immediate and mediated experience, and so on. Accordingly, psychology could be defined as the science that studies the phenomena of inner experience.

Avenarius, on the other hand, believed that experience is essentially homogeneous. We experience all sorts of different contents, but these are not experienced *in different ways* (see Avenarius, 1891/1905, 82). According to him, the concept of "inner" (inner experience, inner world, inner observation…) is the result of a fundamental mistake, which he called "introjection." Introjection arises from an incorrect interpretation of the experience of the fellow-man. The fellow-man is like me; this means that he too has an experience. The problem is that we do not experience his experience, so the question arises: *Where* is the fellow-man's experience? Introjection answers this question by saying that the fellow-man's experience is hidden *inside within* him. Thus, thanks to a "creation out of nothing," introjection gives rise to the idea of a specific dimension of interiority, that is opposed to the external material world (Avenarius, 1894–1895, 18, 159). Worse still, as soon as we begin to ascribe an inner life to our fellow-men, we begin to view our own experience in the same way. Consequently, the original unity of our own experience is split into two parts: inner and outer world (Avenarius, 1891/1905, 25–62).

Avenarius believed that, in order to avoid introjection and its consequences, we must stick to what we actually experience. We do experience the fellow-man's movements (words, gestures, facial expressions), and – because of the similarity between us and the fellow-man – we assume that these movements have not only a "mechanical meaning" but also a "more-than-mechanical meaning," i.e., a linguistic meaning. We ascribe to the movements of the fellow-man the same meaning that we do experience in relation to our own movements. Consequently, the experience of the fellow-man is not a hidden, unexperienceable inner world, but rather his "contents of assertions," which Avenarius also called "*E*-values" (Avenarius, 1891/1905, 6–10).

Once we reject introjection and return to the original unity of experience, there is no longer room for the dualism between the psychical and the physical, the inner and the outer experience. Therefore, we can no longer define psychology by specifying a particular field of investigation. For this reason, Avenarius proposed a new way of defining psychology, which distinguishes it from other sciences by the specific point of view from which it regards the experience. According to Avenarius, "the object of empirical psychology is every experience, *insofar as it* […] *is regarded as dependent on the individual in relation to whom* […] *it is an experience* (Avenarius, 1894–1895, 18, 417). This means that even the 'tree in front of us,' the 'movement of the leaves,' or the whole 'moving material word' can become the object of psychology, insofar as we can somehow think of them in connection with the speaking individual, and – in this connection – as somehow (logically) dependent on the features of this individual" (Avenarius, 1894–1895, 18, 414). In particular, the "features of the individual" on which experience depends are the processes in the central nervous system, which Avenarius called "system C." Consequently, it can be briefly said that the "object of psychology is experience in general, regarded as dependent on the system C" (Avenarius, 1894–1895, 18, 418).

As we have seen, Wundt regarded the physiological processes related to mental functions as mere means to produce the psychical contents which are the true and only object of psychology. Conversely, Avenarius' definition of psychology states that the object of this science is the investigation of mental activity in its dependence on the brain. Thus, in contrast to Wundt, Avenarius interpreted psychophysical parallelism as the assumption that there is no single event in experience that is not entirely dependent on the nervous system (Avenarius, 1891/1905, 18–19). In so doing, Avenarius left no room for anything like Wundt's "psychical causality," and stated that the sequence of experiences can only be explained by parallel physiological brain activity.

Needless to say, at first glance Avenarius' view may seem similar, if not identical, to that famously proposed by Ernst Mach in the same period. Indeed, both shared the conviction that the traditional distinction between the physical and the psychical is invalid because experience is essentially unitary. And they both proposed to replace it with a distinction between two different perspectives. As Mach said, "The traditional gulf between physical and psychological researches exists only for the habitual stereotyped method of observation. A color is a physical object so long as we consider its dependence upon its luminous source, upon other colors, upon heat, upon space, and so forth. Regarding, however, its dependence upon the retina it becomes a psychological object, a sensation. Not the subject, but *the direction of our investigation*, is different in the two domains" (Mach, [1886] 1897, 14–15). Indeed, Mach devoted an entire chapter of his *The Analysis of Sensations* to "My Relation to Richard Avenarius and Other Thinkers," in which he stated that he "attaches the greatest importance to our agreement [with Avenarius] in the conception of the relation between the physical and the psychical," and that, for him, "this is the main point at issue" (Mach, [1900] 1914, 50). Because of these analogies, at the time the two thinkers were grouped together under the common label of "empiriocriticists."

However, despite the undeniable similarities, Avenarius' position remains unique. First of all, we must note that Mach opposed the psychological perspective with the perspective of physics. Indeed, Mach's goal throughout his life was to refute the materialist metaphysical interpretation of physics, according to which this science deals with the true reality beyond the sheer appearances of our experience. To this end, Mach stated that *all* sciences have the same object and the same goal: to organize empirical data in such a way that suits the economic drive that comes from our biological needs. Therefore, according to Mach, there were no essential differences between the various disciplines.

Although Avenarius too rejected the opposition between the psychical and the physical, the inner and the outer world, he did not define the psychological perspective by setting it against the perspective of physics. Being a philosopher by training, he pursued different goals than Mach. Avenarius' main goal was to reconcile the apparent contradiction between philosophy and the new experimental and physiological psychology.

For Avenarius, philosophy must acknowledge the "immediate givenness of consciousness" as its only legitimate starting point (Avenarius, 1891/1905, X). We cannot go beyond our own experiences; whatever we know, we know it from what is empirically given to us, in our consciousness. Therefore, philosophy is ultimately based on an "idealistic" point of view (Avenarius, 1891/1905, IX). On the other hand, psychology is based upon a "realist" framework, according to which consciousness and experience are not the *original* and unsurpassable ground of all knowledge but the *result* of certain physiological processes, in which the brain reacts to the environment. This means that, for psychology, the brain and the environment must somehow lie *before* and *beyond* consciousness and experience.[15]

Avenarius believes that philosophy, being based on an idealistic point of view which contrasts with the realism of psychology, is incapable of providing a foundation for the latter. On the other hand, he concedes that we cannot simply do away with philosophical idealism, because psychology itself seems to lead to idealistic conclusions. In fact, by the early nineteenth century, the first physiological investigations of the sensory system by Johannes Peter Müller and his students had led to a reaffirmation of Kant's idealism. Müller's law of specific nerve energy – according to which the specific quality of any sensation reflects not the specific quality of the stimulus, but the specific character of the stimulated nerve – was interpreted as a confirmation of Kant's belief that we do not perceive the objective features of the world, but only our own subjective intuitions of it.[16] Against this background, Avenarius wrote: "I believe that there are quite a number of representatives of philosophical idealism, *who have been trained in the natural sciences*, and who would consider the restoration of their former 'realism' as a relief. They would gladly allow this, if they only knew how to break away from 'idealism' with a clear conscience, from a logical point of view. But for them, it is an undeniable fact that – as soon as we reflect upon things – we come across the scheme of cause and effect, in which the things are the causes and the 'sensations' = 'perceptions'= 'conscious phenomena' are the effects, and these effects are 'idealistic' contents, and these 'idealistic' contents are what is 'immediately given' and, therefore, 'the only given,' from which we might 'infer' 'what lies beyond consciousness,' even though 'everything that is inferred' should in turn be only 'in our consciousness'" (Avenarius, 1891/1905, 108–109, emphasis mine).

So, what was Avenarius' solution to this fundamental conflict between experience as the ultimate *starting point* of all knowledge and experience as the *result* of the interaction between brain and environment? Who is right? Philosophy and its idealism, even if it does not seem to be able to lay the foundation for the physiological investigations of the conditions of experience? Or psychology and its realism, even if it seems to lead back to idealism?

Avenarius answered that both are right, insofar as they do not refer to the same experience. According to him, philosophy refers to *my* experience. Conversely,

[15] In this context Avenarius uses "experience" and "consciousness" as synonyms. When he speaks of "psychology," he means the physiological-experimental psychology of the time.

[16] On the Kantian readings of the physiological investigations in the early XIX century see Edgar (2015), and Beiser (2014), especially the chapters "The Interim Years, 1840–1860" and "The coming of age."

psychology deals with the experience *of the fellow-man*. Of course, the latter is also a part of my experience. As we saw when we talked about introjection, the experience of the fellow-man is given in my experience through the linguistic assertions of the fellow-man (*E*-values), i.e., in the meaning of his sounds, movements, gestures, expressions, etc. For Avenarius, it is *my* experience - the experience *in the first person*, as we may say - that is the inescapable horizon of all knowledge. Whatever is given must be given to my consciousness; otherwise, it is virtually non-existent. On the other hand, psychology does not deal with *this* experience, but with the experience *of the fellow-man*, the experience *in the third person*. If we keep this distinction in mind, many of the classical problems of philosophy and psychology simply disappear as false problems.

Let us consider the problems that arise when I claim that *my* experience of a tree depends on my brain perceiving the tree. First, in my experience, the tree and my perception of the tree are not two separate contents: there is only *one* tree, so I cannot say that one depends on the other. Moreover, I cannot claim that an experience depends on my brain, because either the brain is regarded as something that is *beyond* my experience or as an empirical content that is *part of* my experience. In the first case, the brain is not an empirical content, but becomes a sort of metaphysical entity. In the second case, we fall into two logical absurdities: because the whole of my experience depends on one of its parts and because the experience of the brain also depends on the brain, so that the brain is the cause of itself. Conversely, I *can* say that my fellow-man's experience of the tree depends on his brain perceiving the tree. First, in my experience, his experience of the tree (i.e., his assertion about it) and the tree are actually two different things, so there is no problem in stating the dependence of one on the other. And the same is true for the dependence of his experience on his brain and even for the dependence of his experience "brain" on his brain. In each case, we simply establish that his assertions depend on his brain. So, we are simply stating the dependence between two empirical contents.

However, we should note that Avenarius did not really claim that we cannot claim the dependence of our own experiences upon our brains. The point is rather that, when we claim such dependence, we are not talking about our experiences in the first person. Even without knowing it, we view ourselves from a third-person perspective, we look at ourselves as if we were another person, a fellow-man. As Avenarius writes, "The moment a person wants to conceive of 'something found by himself' as dependent on his own 'brain,' he must look at himself from the relative perspective. However, since he cannot put himself in a point of view outside his own, *in this self-observation he is only able to imitate the observation of other people*" (Avenarius, 1891/1905, 89–90, emphasis mine).

Having presented the core of Avenarius' conception, it is apparent that the similarity with Mach is only superficial. Undoubtedly, they shared the definition of psychology as the science that regards experience as dependent on the individual. Therefore, they both believed that the goal of psychology is the investigation of mental life in its connection with the brain. In other words, unlike Wundt, they both maintained that physiological psychology is not just a branch of psychology, but psychology itself, the only psychology that exists. On the other hand, Avenarius not only claimed that psychology regards experience in its dependence on the brain. He also claimed that we can state this dependence only in reference to the fellow-man if we do not want to fall back into the old antinomies between idealism and realism.

We must stress that for Avenarius the distinction between first-person and third-person perspective was not just a theoretical trick to solve some abstract philosophical problem. With his conception of psychology, Avenarius wanted to stay true to experience. In fact, we do not observe the dependence of our own experiences on ourselves, but we know of this dependence only by observing of other people. As Avenarius wrote, "Referring to the fellowman has the advantage that we *actually* obtain the analytical aspects of the human person from other individuals. What I know about my body's internal constitution – about the blood, the nerves, and ultimately about the brain – I know, to a great extent, *only* through the analysis of foreign bodies, which is *then* transferred to me" (Avenarius, 1891/1905, 22, emphasis mine).

We may say that Avenarius' conception of psychology is modeled on the actual practice of the experimental-physiological psychology of his time. During the experiment, the psychologist observes the motor responses of the subject of the experiment. The psychologist does not consider such motor responses as merely physiological-mechanical movements, but he assumes that they have a linguistic meaning too, insofar as they express the experience of the subject of the experiment. So, during the experiment, the psychologist empirically observes the connections between the contents of assertions of the subject of the experiment (*E*-values) and the physiological processes in his nervous system (system C). This means that the experimental investigation consists solely of the observation of a third person by the experimenter. Even when the experiment involves not only the observation of a motor response but also an "introspective" account by the subject of the experiment (as in the case of the experiments conducted in Wundt's laboratory), it is still the observation, by the experimenter, of certain connections between the physiological activity of a third person and the content of assertions made by this third person.

In view of the above, we may say that Avenarius completely rejected the traditional conception of psychology as the science of inner observation. For him, the paradigm of this science was no longer the *single* individual looking within himself. In, Avenarius elevated to new paradigm the experimental setting, which always involves *two* persons: the experimenter, who observes, and the subject of the experiment, who is observed. In so doing, Avenarius eliminated all the problems related to the old notion of "inner" experience. There is no longer the problem of the subjectivity and unreliability of inner self-observation, since introspection is replaced by objective observation of the connections between physiological processes and the experiences asserted by a third person. The alleged unobservability of other people's "inner" experience is not a problem either, since the experience in the third person is nothing but the observable assertions of the fellow-man. In so doing, psychology is indeed established as an empirical science, since it relies entirely on observable facts.

As Wladyslaw Heinrich, one of Avenarius' students, pointed out: the "rejection of self-analysis" was a consequence both of "the deficiencies of this method," which is not "controllable," and of the need for an "exact method." Unfortunately, "the implementation of the experiment instead of the defective self-investigation" at first did not lead to "ask anew the question of *what* was being investigated." Avenarius and his school tried to fill this gap. Their answer was that "the objective investigation of others has placed the psychology among the other natural sciences." For them, "the experimental research cannot be the investigation of the foreign consciousness," because "what is accessible to our immediate observation are just the assertions of men in the form of communications or actions" (Heinrich, 1896, 348–349, emphasis mine).[17]

Wundt's Reply to Avenarius

In the late nineteenth century, the definition of psychology as the science that regards experience in its dependency on the brain began to spread. In fact, even some of Wundt's students adopted this definition in their works: Hugo Münsterberg and Edward B. Titchener, who were among the founding fathers of American experimental psychology, as well as Wundt's own laboratory assistant Oswald Külpe, later the founder of the so-called Würzburg school (Russo Krauss, 2019, 59–111).

Wundt did not stand idly by. First, he interrupted his collaboration with Avenarius and withdrew his name from the editorial board of his colleague's journal. Then, despite Avenarius' death in 1896, he wrote a series of papers against his views, culminating in the long and harsh essay *Über naiven und kritischen Realismus. II. Der Empiriokritizismus* (On naïve and critical realism. The empiriocriticism). In this work, Wundt attacked the two pillars on which Avenarius had established psychology as physiological-experimental psychology: the investigation of the dependence of the experience on the brain activity and the elevation of the experimental observation of the fellow-man to the rank of new paradigm.

As for the first issue, Wundt dismisse the overvaluation of the role of the physiological investigations as a form of metaphysical materialism. According to him, if Avenarius really wanted to stay true to the empirical facts, he should have acknowledged "the relationships between the central functions and the cognitive and emotional values" only "where they are empirically demonstrated." And he should have also acknowledged "the mutual relationships of dependency *between the psychological values*" (Wundt, 1898, 47, emphasis mine). Instead, Avenarius denied the existence of purely psychological connections and postulated that *every* psychical occurrence depends upon the brain activity. This proves that he was not following the experience, but rather the "metaphysical urge to link the everchanging flow of events to an immutable being, a substance" (Wundt, 1898, 47). Thus, Avenarius' arguments were "essentially none other than the long-familiar arguments that

[17] Following Avenarius' footsteps, Heinrich further developed his project of a purely experimental psychology in his *Die moderne physiologische Psychologie in Deutschland* (1899). Among the works of Avenarius' pupil that deal with the problem of psychology, we may cite Carstanjen (1894), Kodis (1895), Cornelius (1897), and Willy (1899).

constantly reappear in the materialistic literature of the 18th century," so much so that "one cannot even say that they are really presented in a new light" (Wundt, 1898, 46). For Wundt, the only merit of Avenarius was having "attempted to express the assumption of the exclusive dependency of spiritual life upon the central nervous system in a more exact way in comparison to the older forms of materialism" (Wundt, 1898, 353). But precisely for this reason, Avenarius "provided an eloquent proof of the unsustainability of this general perspective" (Wundt, 1898, 365).

As regards the second issue, Wundt attacked the "misconception, according to which the 'assertions' about something experienced in first-person [*Selbsterlebtes*] can count as an equivalent of the latter." Wundt conceded that "even in the description of subjective observations we cannot do without language." Yet, for him, "the auxiliary means of linguistic communication" aimed at "awaking in the other person the adequate concepts and at making possible the *comparison with own observations*" (Wundt, 1898, 13, 55n, emphasis mine). According to Wundt, every psychological discourse ultimately must lead to the inner, personal experience of the subject to be intelligible at all. Therefore, the inner, personal experience remains the fundamental ground of psychology.

However, we must notice that Avenarius did not dispute the epistemological priority of the first-person experience, which he held as the basis of *all* sciences, psychology included. Moreover, he acknowledged that the linguistic meaning of the fellow-man movements arises from an analogy with the meaning of our own movements. The real conflict between the two was that Wundt believed that psychology deals with the first-person experience, whereas Avenarius believed that it deals with the third-person experience.

This is all more evident when one looks at Wundt's rebuttal of the idea that the experiment necessarily involves two persons. For him, it is wrong to assume "that the peculiarity of the 'experimental psychology' is that the experimenter put *other people* in certain conditions, under which they must make certain assertions about what they have observed," because "*the observer is not the so-called experimenter, but the 'subject of the experiment,'* and the setting of the experiment serves just put him in the favorable conditions for the *subjective observation*" (Wundt, 1898, 13, 55n, emphasis mine). In other words, for Wundt, the experiment is not characterized by the distinction between the observer and the one who is observed (the experimenter and the subject of the experiment). Since psychology is based on self-observation, *even in the case of the experiment*, the observer and the one who is observed are the same person. Vice versa, Avenarius stated that psychology is based on the observation of the dependency between the physiological activity of a third person and his assertions. Thus, he believed that *even when we are considering ourselves* from a psychological perspective, there must be such a distinction between the observer and the one who is observed, insofar as we regard ourselves as if we were another person.

Wundt's rejection of the idea that the psychological experiment implies a distinction between the observer and the one who is observed is confirmed by his criticism of the experiments with hypnosis. According to him, hypnosis is inherently flawed because "the deep hypnotic sleep makes the *self-observation* impossible." Moreover, since the observer is the subject of the experiment, rather than the experimenter, the hypnosis would work only if it was possible *for the psychologist* to be hypnotized. But the psychologist's awareness of the scope and circumstances of the experiment hinders his transition into hypnosis (Wundt, 1892, 40).

Groundbreaking Aspects of Avenarius' Conception of Psychology

To sum up, despite Wundt's fame as the founding father of experimental psychology, he still maintained the conception of psychology as the science of inner observation. This is reflected in his interpretation of the experimental method. For him, the experiment was just an auxiliary means to support self-observation, which remained the fundamental basis of all psychology. Accordingly, he held that, even in the experimental setting, the observer and the one who is observed are the same person, i.e., the subject of the experiment, who must provide his introspective account of his mental states.

On the other hand, during those same years, Avenarius proposed a new foundation for psychology. Instead of trying to fit the old introspective conception into the new experimental practice, Avenarius embraced the latter, redesigning on its basis the very notion of psychology. Like Mach, Avenarius defined it as the science that deals with the experience in its dependency upon the individual. In so doing, they both identified psychology with the physiological psychology first established by Fechner. However, unlike Mach, Avenarius also stressed that the investigation of this dependency could only regard the fellow-man. Psychology does not deal with the experience in the first person, but with the experience in the third person or, more precisely, with the empirical observation of the connections between the physiological processes of the fellow-man and his assertions. Thus, the experimental setting, with the separation between the observer and the one who is observed, the experimenter, and the subject of the experiment, is elevated to the rank of new paradigm of psychology.

In light of the above, we may say that Avenarius' conception of psychology presents several groundbreaking aspects. First of all, he clearly anticipates Watson's behaviorism, insofar as they both deny that the investigation of inner mental states is the fundamental object and method of psychology.[18]

Secondly, by separating the first-person and the third-person perspectives, Avenarius defends the physiological approach without falling into a reductionist position. Indeed, psychology can and must aspire to discover the physiological conditions of every mental occurrence. However, even if psychology ever fulfills this task, this will not mean that it has made any step further in explaining the

[18] Titchener pointed out Avenarius' priority in his reply to Watson's behaviorist manifesto. According to Titchener, the "unhistorical character" of Watson's manifesto hid the fact that "behaviorism is neither so revolutionary nor so modern as a reader unversed in history might be led to imagine" (Titchener, 1914, 4–5). Hence, Titchener remarked that definitions of psychology analogous to that of Watson had already been proposed by himself, Ward, Avenarius, Külpe, and Ebbinghaus (Titchener, 1914, 1–2).

first-person experience. The first-person experience remains beyond reach, and it cannot be the object of science, being rather the starting point of all sciences. In this respect, Avenarius' arguments resonate with those famously presented by Thomas Nagel in his *What Is It Like to Be a Bat?* (1974). They both stress the fundamental difference between the subjective point of view from which we regard our own experience and the objective point of view from which we regard the experiences of other people. And they both agree that, even though language can help us switch perspective, putting ourselves in other people's shoes, so to speak, it does not eliminate this fundamental difference between the first-person point of view and the point of view toward other people. Finally, they both acknowledge that a complete scientific account of the neurological conditions of mental life does not and cannot regard our first-person experience.

Needless to say, Avenarius' philosophy is not without defects. Considering the paramount role played by the "contents of assertions" in his system, the most evident flaw is definitely the lack of a proper theory of meaning. His treatment of the topic appears very naïve. The only works about logic he cited were the ones by Drobisch (1836), Sigwart (1873–1878), and Wundt (1880–1883) that all shared a psychological approach, insofar as they linked logical contents with the psychological acts of thought. This indicates that Avenarius was more interested in the psychological aspects of the act of meaning, rather than in a purely logical theory of meaning, like the one Frege had developed at that time. However, this deficiency should not induce us to underestimate the innovative and influential aspects of his conception of language. The fact that he regarded language as a potential source of philosophical problems (introjection) and, at the same time, as a way out from those problems (thanks to the distinction between the first-person experience and the experience as contents of assertions of the fellow-man) gives us a glimpse of issues that were bound to become popular after the coming linguistic turn.

What Can We Learn from the Debate Between Avenarius and Wundt About Experimental Psychology?

The debate between Wundt and Avenarius demonstrates that scientific methods may act at very different levels in the conception of science. On a first level, they can be regarded as mere "'toolbox' of ready-made (and often 'standardized') concrete methods that can be borrowed at a researcher's will without much consideration to the phenomena to which they are applied, for the purposes of producing 'data'" (Valsiner, 2017, 5). For example, the physiological-experimental method first developed by Fechner was adopted by Wundt as a mere practice employed to obtain data, without influencing his overall conception of this science.

However, further considerations soon reveal that every method is always part of a broader conception of a methodology, where "abstract and concrete features of the research act are intricately intertwined" (Valsiner, 2017, 5). In our case, the fact that Wundt regarded the physiological-experimental method as a mere auxiliary means does not prove that methods are theoretically neutral. On the contrary, it proves that they are always applied and interpreted within specific frameworks that shape their meaning and the meaning of the data obtained through them. Wundt regarded the experiment as a mere support to self-observation because he maintained an introspective conception of psychology. As stressed by Danziger, "Wundt's laboratory produced a large number of studies whose data base was entirely 'behavioral,' mostly in the form of various kinds of reaction time measures. What was 'mentalistic' about these studies was the theoretical interpretation of the results, not the data base itself" (Danziger, 1980b, 248).

Then, in Avenarius we observe a further level of interaction between a method and the more general conception of science. The method does not just receive its meaning within a certain conception of a given science, but it directly contributes to shaping this conception. For Avenarius, the experimental setting, where the experimenter observes the assertions of the subject of the experiment in reaction to his stimulation, becomes the basis upon which psychology itself is defined as the science that investigates the dependency of the contents of assertions (*E*-values) upon the brain (system C).

Moreover, we may say that Avenarius shows us that a method can interact not only with the general conception of a given science but also, on an even higher level, with the relationship between a certain science and knowledge in general. Indeed, Avenarius not only defined psychology on the basis of the experimental method, but he used the experimental method as a starting point to separate two fundamental attitudes toward experience: the first-person perspective and the third-person perspective. His reflection on the experimental method employed by the psychologists of that time led him to what he held as a possible solution for the philosophical antinomies between idealism and realism.

However, we must notice that Avenarius' attention toward the interactions between methods, methodology, the definitions of particular sciences, and knowledge in general was rather typical for the German culture of the nineteenth century. At the time, sciences like economics, sociology, history, and psychology were all involved in the so-called *Methodenstreit* (literally, the dispute over methods).[19] Indeed, after Kant, the thinkers of that era intended philosophy as *Wissenschaftslehre* or *Methodenlehre* (doctrine of sciences or doctrine of methods), i.e., as the discipline the purpose of which was to "critically examine the principles and *methods* of knowledge, of science in general and of the singular sciences in particular," as well as "to bring knowledge and science to full consciousness about their *practice* (*Tun*), their essence, and their limits" (Eisler, 1904, 804–805, emphasis mine).

In conclusion, what we can learn from that era and its representatives, like Wundt and Avenarius, is to keep alive the interaction between philosophy and science, not just in the sense of the theoretical reflection on the premises and the results of science but also on its concrete practices and methods. Because the latter are not just neutral instruments to collect data, but define each science, in itself and in its dynamic relation to the whole system of knowledge.

[19] On the *Methodenstreit*, with particular focus on psychology, see Nerlich, 2000.

References

Araujo, S. D. F. (2016). *Wundt and the philosophical foundations of psychology: A reappraisal*. Springer.

Araujo, S. D. F. (2018). Völkerpsychologie as cultural psychology: The place of culture in Wundt's psychological project. In G. Jovanović, L. Allolio-Näcke, & C. Ratner (Eds.), *The challenges of cultural psychology. Historical legacies and future responsibilities* (pp. 75–84). Routledge.

Ash, M. G. (1980). Academic politics in the history of science: Experimental psychology in Germany, 1879–1941. *Central European History, 13*(3), 255–286.

Avenarius, R. (1888–1890). *Kritik der reinen Erfahrung* (Vol. 2 vols). Fues.

Avenarius, R. (1891/1905). *Der menschliche Weltbegriff*. Reisland.

Avenarius, R. (1894–1895). Bemerkungen zum Begriff des Gegenstandes der Psychologie. *Vierteljahrsschrift für wissenschaftliche Philosophie, 18*, 137–161, 400–420, 19: 1–18, 129–145.

Beiser, F. (2014). *The genesis of neo-Kantianism, 1796–1880*. Oxford University Press.

Boudewijnse, G.-J. A., Murray, D. J., & Bandomir, C. A. (1999). Herbart's mathematical psychology. *History of Psychology, 2*(3), 163–193.

Boudewijnse, G.-J. A., Murray, D. J., & Bandomir, C. A. (2001). The fate of Herbart's mathematical psychology. *History of Psychology, 4*(2), 107–132.

Carstanjen, F. (1894). Richard Avenarius's Biomechanische Grundlegung der neuen allgemeinen Erkenntnistheorie. *Eine Einführung in die "Kritik der reinen Erfahrung"*. Th. Ackermann.

Cornelius, H. (1897). *Psychologie als Erfahrungswissenschaft*. B.G. Teubner.

Danziger, K. (1979). The positivist repudiation of Wundt. *Journal of the History of the Behavioral Sciences, 15*(3), 205–230.

Danziger, K. (1980a). On the thresold of the new psychology: Situating Wundt and James. In W. G. Bringmann & R. D. Tweney (Eds.), *Wundt studies. A centennial collection* (pp. 363–379). Hogrefe.

Danziger, K. (1980b). The history of introspection reconsidered. *Journal of the History of Behavioral Sciences, 16*, 241–262.

Danziger, K. (1987). Wilhelm Wundt and the emergence of experimental psychology. In G. N. Cantor, J. R. R. Christie, M. J. S. Hodge, & R. C. Olby (Eds.), *Companion to the history of modern sciences* (pp. 396–408). Routledge.

Danziger, K. (2001). The unknown Wundt: Drive, apperception and volition. In R. W. Rieber & D. K. Robinson (Eds.), *Wilhelm Wundt in history. The making of a scientific psychology* (pp. 95–120). Springer.

Diamond, S. (2001). Wundt before Leipzig. In R. W. Rieber & D. K. Robinson (Eds.), *Wilhelm Wundt in history. The making of a scientific psychology* (pp. 1–68). Springer.

Dilthey, W. (1894/2010). The ideas for a descriptive and analytic psychology. In R. A. Makkreel & F. Rodi (Eds.), *Wilhelm Dilthey: Selected works, Volume II: Understanding the human world* (pp. 115–210). Princeton University Press.

Drobisch, M. (1836). *Neue Darstellung der Logik nach ihren einfachsten Verhältnissen*. Voss.

Edgar, S. (2015). The physiology of the sense organs and early neo-Kantian conceptions of objectivity: Helmholtz, Lange, Liebmann. In F. Padovani, A. Richardson, & J. Y. Tsou (Eds.), *Objectivity in science, new perspectives from science and technology studies* (pp. 101–122). Springer.

Eisler, R. (1904). *Wörterbuch der philosophischen Begriffe*. Mittler und Sohn.

Fahrenberg, J. (2011). *Wilhelm Wundt – Pionier der Psychologie und Außenseiter? Leitgedanken der Wissenschaftskonzeption und deren Rezeptionsgeschichte*. Online: http://www.jochen-fahrenberg.de

Fahrenberg, J. (2015). *Theoretische Psychologie. Eine Systematik der Kontroversen*. Pabst. Online: http://www.jochen-fahrenberg.de

Fahrenberg, J. (2020). *Wilhelm Wundt (1832–1920). Intoduction, quotations, reception, commentaries, attempts at reconstruction*. Pabst.

Fechner, G. T. (1860). *Elemente der Psychophysik* (Vol. I). Breitkopf und Härtel.
Heinrich, W. (1896). Die Aufmerksamkeit und die Funktion der Sinnesorgane. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane, 9*, 343–388.
Heinrich, W. (1899). *Die moderne physiologische Psychologie in Deutschland. Eine historisch-kritische Untersuchung mit besonderer Berücksichtigung des Problems der Aufmerksamkeit*. E. Speidel.
Herbart, J. F. (1816/1901). *A text-book in psychology. An attempt to found the science of psychology on experience, metaphysics, and mathematics* (M.K. Smith, Trans.). Appleton.
Hodges, H. A. (2000). *The philosophy of Wilhelm Dilthey*. Routledge.
Jüttemann, G. (2006). *Wilhelm Wundts anderes Erbe. Ein Missverständnis löst sich auf.* Vandenhoeck und Ruprecht.
Kant, I. (1786/2004). *Metaphysical foundations of natural science* (M. Friedman, Trans. & Ed.). Cambridge University Press.
Kim, A. (2009). Early experimental psychology. In J. Symons & P. Calvo (Eds.), *The Routledge companion to philosophy of psychology* (pp. 41–58). Routledge.
Kodis, J. (1895). Die Anwendung des Functionsbegriffes auf die Beschreibung der Erfahrung. *Vierteljahrsschrift für wissenschaftliche Philosophie, 19*, 359–367.
Kusch, M. (1991). The sociological deconstruction of philosophical facts. The case of psychologism. *Science & Technology Studies, 4*, 45–59.
Mach, E. (1886/1897). *Contributions to the analysis of the sensations*, trans. by C. M. Williams. Open Court.
Mach, E. (1900/1914). *The analysis of sensations, and the relation of the physical to the psychical*, trans. by C. M. Williams. Open Court.
Murray, D. J., & Bandomir, C. A. (2001). Fechner's inner psychophysics viewed from both a Herbartian and Fechnerian perspective. In E. Sommerfeld et al. (Eds.), *Fechner day 2001: Proceedings of the 17th annual meeting of the international society for psychophysics* (pp. 49–54). Pabst.
Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review, 83*(4), 435–450.
Nerlich, B. (2000). Coming full (hermeneutic) circle: The debate about methods in psychology. In Z. Todd, B. Nerlich, S. McKoweon, & D. Clarke (Eds.), *Mixing methods in psychology. The integration of qualitative and quantitative methods in theory and practice* (pp. 16–36). Routledge.
Rieber, R. W., & Robinson, D. K. (Eds.). (2001). *Wilhelm Wundt in history. The making of a scientific psychology*. Springer.
Robinson, D. K. (2001). Reaction-time experiments in Wundt's institute and beyond. In R. W. Rieber & D. K. Robinson (Eds.), *Wilhelm Wundt in history. The making of a scientific psychology* (pp. 161–204). Springer.
Russo Krauss, C. (2019). Wundt, Avenarius, and scientific psychology. A Debate at the Turn of the Twentieth Century.
Sigwart, C. (1873–1878). *Logik*. Tübingen.
Teo, T. (2005). *The critique of psychology from Kant to postcolonial theory*. Springer.
Titchener, E. B. (1914). On 'psychology as the behaviorist views it'. *Proceedings of the American Philosophical Society, 53*(213), 1–17.
Valsiner, J. (2017). *From methodology to methods in human psychology*. Springer.
van Hoorn, W., & Verhave, T. (1980). Wundt's changing conception of a general and theoretical psychology. In W. G. Bringmann & R. D. Tweney (Eds.), *Wundt studies. A centennial collection* (pp. 71–113). Hogrefe.
Volkelt, J. (1887). Psychologische streitfragen. I. Selbstbeobachtung und psychologische analyse. *Zeitschrift für Philosophie und philosophische Kritik, 90*, 1–49.
Willy, R. (1899). *Die Krisis in der Psychologie*. O.R. Reisland.
Wundt, W. (1874). *Grundzüge der physiologischen Psychologie*. Engelmann.
Wundt, W. (1880–1883). *Logik: eine Untersuchung der Principien der Erkenntniss und der Methoden wissenschaftlicher Forschung*. Enke.

Wundt, W. (1882/1906). Die Aufgaben der experimentellen Psychologie. In W. Wundt (Ed.), *Essays* (pp. 187–209). Engelmann.
Wundt, W. (1888a). Selbstbeobachtung und innere Wahrnehmung. *Philosophische Studien, 4*, 292–309.
Wundt, W. (1888b). Ueber Ziele und Wege der Völkerpsychologie. *Philosophische Studien, 4*, 1–27.
Wundt, W. (1892). *Hypnotismus und suggestion*. W. Engelmann.
Wundt, W. (1895). *Logik. Eine Untersuchung der Principien der Erkenntniss und der Methoden wissenschaftlicher Forschung* (Vol. II/2, 2nd ed.). F. Enke.
Wundt, W. (1896). *Grundriss der Psychologie*. W. Engelmann.
Wundt, W. (1898). Über naiven und kritischen Realismus. Der Empiriokritizismus. *Philosophische Studien, 13*, 1–105. 323–433.

Chapter 3
Truth and Mind: How Embodied Concepts Constrain How We Define Truth in Psychological Science

Heath Matheson

'Is truth dead?' reads the cover of *Time* magazine in April 2017 (Pine, 2017). Culture wars, fake news, alternative facts, post-truth politics, echo chambers, anti-intellectualism and the increasing primacy of lived experience over empirical data define the context of this alarming question (see McIntyre, 2018). But what motivates the question 'Is truth dead?', and what does it mean for experimental psychology? Are we really living in a 'relativist' world in which claims about mind, brain and behaviour cannot be arbitrated? If someone suggests that depression leads to increased creativity or that the anterior temporal lobe of the cortex is the basis of semantic memory, can we not evaluate the 'truth' of those statements?

It is deceptively easy for psychologists to dismiss these questions as rhetoric—of course we can arbitrate truth! We use the scientific method, inferential statistical tools and our own ingenuity to understand truth. Anyone challenging our claims regarding the objective truths of the human mind can be put in their place with the use of logic and data. Or so we would think. Although we like to suppose that experimental psychology is destined to arbitrate reality and uncover truths, psychology suffers from its own set of truth crises: first, the replication crisis has shown that many major findings are unreliable (Pashler & Wagenmakers, 2012); second, the theory crisis has demonstrated that the discipline has failed to develop any major or overreaching laws or theories (Oberauer & Lewandowsky, 2019); and finally, the practicality crisis highlights how our findings are rarely applied into domains of life (Berkman & Wilson, 2021). Overall, we have to wonder if psychology lacks three keystones of scientific progress: (a) accumulating knowledge that (b) helps us generate predictive theories and that (c) allow us to intervene on what happens in the world (see Newell, 1973; van Rooij & Baggio, 2021). Indeed, the number of 'truths' from experimental psychology that are being upturned by failed replications

H. Matheson (✉)
The University of Northern British Columbia, Prince George, BC, Canada
e-mail: heath.matheson@unbc.ca

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9_3

continues to increase (see Shrout & Rodgers, 2018, for a thorough review), and such failures increase mistrust both within and outside of psychology (e.g. Anvari & Lakens, 2018; see also Fawcet & Matheson, 2019). Despite some debate about the existence and extent of the replication crisis and other crises (see Pashler & Harris, 2012, for a discussion), it is clear that confidence in psychological science has been shaken to its core. Given psychology's truth crises, I would not be surprised to see the *Time* magazine cover co-opted by any of the field's most impactful journals. For many, psychology is in need of revolution (Spellman, 2015).

How can psychology address its truth crises? In this chapter, I hope to show that empirical considerations borne from experimental psychology and neuroscience regarding the learning and use of *concepts* (a.k.a. categories) are the key to mitigating psychology's turmoil and understanding how we define truth.[1] My argument has a number of steps. First, I examine two definitions of truth. Next, I give a brief introduction to the theoretical framework of embodied cognition and discuss how neural and behavioural research within this framework shows that concepts are constrained by embodied experience. Finally, I will discuss the implications of these constraints on what it means to say that something is 'true'—for how we conduct scientific inquiry and, more broadly, for our sociopolitical lives.[2] Ultimately, I hope to show that truth is not dead and that psychological science, by embracing embodiment, can escape its crises.

What Is Truth?

We are all engaged in truth-seeking activities. When a politician suggests that we simply look at the facts—and that someone else's facts are fake—they are making a strong commitment to an objective reality. Similarly, when a psychologist reports that depressed people are more creative or that they have discovered the neural basis of semantic memory, they are making a commitment to a reality. An objective, reality is one that exists independently of human experience and can be known both formally through scientific method and informally by 'seeing it with your own eyes'. This stance towards reality is philosophical realism—the idea that things

[1] Here I tackle the issue not as an analytic philosopher concerned with the specific concepts we use but rather how our understanding of how concepts are structured in the mind/brain shape what we can call truth.

[2] Before beginning, it is essential to point out that psychology and related neuroscientific fields are the only scientific disciplines in which the phenomenon of interest studies itself: that is, minds/brains studying minds/brains. We are people with poor eyesight attempting to remove our eyeglasses to study the said eyeglasses. Using psychological science to critique psychological science necessarily entails that the critique is applicable to itself. Despite this, I am feeding the observations from our scientific practice back into themselves, and in doing so, I hope to provide a method for understanding truth that will help us deal with the truth crises we are currently facing in science and in our sociopolitical lives without succumbing to various radical conclusions about the nature of reality.

exist and have properties regardless of human minds (i.e. reality is mind-independent). Accordingly, if one person is correct about reality, then another person with a conflicting viewpoint is necessarily incorrect. In science, this idea is famously captured by the quip 'Gravity is not a version of truth. Gravity is truth. Anyone who doubts it is invited to jump out a tenth-story window' (attributed to Richard Dawkins).

In the philosophical literature on truth (see Dowden & Swartz, 2020, for an introduction), realism is consistent with the correspondence theory of truth, where truth is defined by how accurately it reflects reality (see Davidson, 1984, for review and critique). You can perform an experiment where you test the major predictions of gravity by jumping out of a window; if gravity is true, then your death verifies it, and we can conclude that the concept of gravity in our minds/brains corresponds to the mind-independent reality.

However, there is at least one major problem with applying this perspective to psychological science: Most of our psychological scientific questions are not readily arbitrated by death as they are in the jump-out-the-window experiment (a $p > 0.05$ in an experiment on creativity or memory won't lead to death of any psychologists, regardless of what they think it will do to their careers). Instead, psychologists rely on *concepts* and attempt to operationally define 'depression', 'creativity' and 'semantic memory', and they use statistical inferences to draw conclusions related to these concepts. Further, they do so in particular contexts and with particular goals in mind, usually involving the capacity to predict how people will behave on average in particular contexts that matter to the researcher.

If death doesn't serve to arbitrate truth claims in psychological science, what does this mean for the kind of truths we can discover? William James, a progenitor of western psychology, helped establish the pragmatic theory of truth (James, 1907). In short, the pragmatic theory suggests that truths are evaluated by whether they lead to successful actions by groups of people over time. I argue that neuroscientific and behavioural research on concepts—that is, the categories we think about and use to make decisions and try to measure in experiments—strongly supports a pragmatic theory of truth.[3]

A very general way to define concepts is to say that they are groupings of things that serve some purpose (see Margolis & Laurence, 1999). In this way, science is about concepts. Variables are concepts (i.e. depression, creativity). When we measure behaviour, we measure it with concepts (e.g. accuracy). When we hypothesize and reason about causes and relationships (e.g. depression leads to enhanced creativity), we do so with concepts. And when we make statements to other scientists about experimental findings, we are attempting to convey concepts. Similarly, our sociopolitical life is defined by the use of concepts. When we make decisions on

[3] William James introduced his lecture on pragmatism with the subheading 'a new name for some old ways of thinking', and this chapter does something similar, updating and (I hope) enriching an old idea with current research results and current theorizing. Like James, it will seek an ethically/metaphysically satisfying way to approach truth and science while still retaining 'the richest intimacy with facts' (James, 1907).

policy, the policies are articulated as collections of concepts (e.g. an increase in taxes; ‘increase’ and ‘taxes’ are concepts). When we vote for platforms, they are articulated as a set of concepts (e.g. the concept of ‘freedom’). The news media report events using sets of concepts. Science and democracy themselves are both concepts.

This observation might seem mundane, but concepts are often described as *the* fundamental building block of psychological functioning, and for some the learning and use of concepts *is* psychological functioning ne plus ultra (e.g. Millikan, 2017). Thus, to understand human behaviour, including scientific and sociopolitical behaviour, is to, in part, understand how we conceptualize our world. Our concepts carve out the ontology of our science—the entities that we think are real or valuable, like depression and semantic memory—and these ontologies guide our interests, research questions and interventions. To have a psychological theory is to have a set of concepts that, we hope, illuminates truth.[4]

Embodiment and Grounding: A Brief Introduction

Because concepts are so central to our scientific and sociopolitical lives, it is imperative that we have a way of characterizing what they are and how they work. Importantly, psychological science has been engaged in this endeavour since its beginnings. In the ‘classic cognitivist’ approach to concepts, ‘thinking’ is best understood as computational operations on concepts that are represented as abstract symbols in the mind/brain (see Newell, 1980, for discussion). For instance, the concept BLACK BEAR is understood by combining the symbol for the concept BLACK and the symbol for the concept BEAR. These ‘mental symbols’ are thought to be stored and manipulated independently of sensorimotor experience; that is, they are ‘amodal’ and do not depend on the biological systems of vision, audition, somatosensation, etc. that are identified by perceptual psychologists and neuroscientists. Further, it is thought that they are only arbitrarily related to their referents in the ‘real world’, are primarily defined with respect to each other and can be manipulated in thought with a type of neural syntax. Note that this perspective on concepts aligns well with the correspondence theory of truth: simply put, to have mental/neural symbols for BLACK and BEAR means that mental/neural symbols correspond to blackness and the bears out there in the mind-independent world. According to this view, doing psychological science is about aligning the symbols we have in our minds/brains (and in our journal articles, textbooks, etc.) with the entities in the real world.

However, critics have shown that there is at least one major issue with this view of concepts (e.g. Barsalou, 1999): If the storage and manipulation of concepts are

[4] The history of thinking about concepts is vast, defining some of humanity’s earliest writings on the workings of mind and resulting in multiple life time’s worth of discussion (see Murphy, 2004; Prinz, 2004; Margolis & Laurence, 1999, for book-length treatments).

divorced from experience (and from the perceptual and motor systems of humans that give rise to experience), and if these symbols are only defined with respect to each other and are arbitrarily related to the world, it is not clear how these purely abstract symbols come to have *meaning* (see Harnad, 1990); that is, why do concepts mean what they do? Indeed, in a computer or a dictionary, a symbol's meaning is provided by the programmers and users. In the mind/brain, if concepts are represented by abstract symbols in this way, who (or what) programs the meaning for *us*[5]?

To deal with this issue, recent frameworks have turned to the body as a way of grounding meaning (Shapiro, 2019).[6] The embodied cognition framework rejects the computer metaphor of the mind and, instead, identifies the body as the starting point for grounding concepts (i.e. creating meaning; see Matheson & Barsalou, 2018, for a review). While currently there is no single, agreed-upon model of embodiment or of how meaning is grounded mechanistically (e.g. Foglia & Wilson, 2013; Hommel, 2015; Wilson, 2002; see Chemero, 2011; Anderson, 2014; Engel et al., 2015; Feldman, 2008; Gibbs Jr, 2005; Lakoff & Johnson, 1999; Prinz, 2004, for book-length treatments), there are some general features of this view shared by some embodied theorists. Here, I briefly review some of these general features, specifically the nature of *how* the mind/brain represents, and *what* is represented in the mind/brain. We will have to take a short detour and delve into details of neural organization[7] to make sense of the implications embodiment has for our understanding of concepts.

How the Brain Represents[8]: Maps and Cognitive Controllers

The cortex of the brain is organized structurally and functionally. Neuronal cells in the eyes and ears and skin transform energy in the world (i.e. light, soundwaves, pressure) into signals that are passed on to other cells in the cortex for further

[5] Assuming that, as many psychologists readily do, that it is not a god or that we are not a computer simulation of some super intelligent alien species.

[6] Much of the short account in this and the previous paragraph is drawn from discussions in Shapiro (2019), especially Chaps. 4 and 5, which provide a thorough discussion of the historical and contemporary issues in embodied cognitive science and the nature of concepts in particular. Please see that text for elaboration on the summary I provide here.

[7] Note I am not looking to *reduce* concepts, truth or scientific and sociopolitical discourse to neural activity, but to show how a consideration of the brain, in the larger brain-body-environmental context, allows us to make sense of meaning. Indeed, embodied approaches are almost all aligned in their proposal that complex psychological constructs like concepts/memory/planning, etc. are *not* reducible to neural events. In that way the present argument is neurophilic, but not neurocentric. Importantly, the argument here does not rest on the fidelity of the specific models I discuss here but rather on the general insights regarding the relationships between brain and behaviour.

[8] Readers familiar with the broader literature on 4e cognition (embodied, extended, embedded, enactive) will understand that many research programs under this framework are anti-representationalist. Here, I use the word simply as a convenience to suggest that we can reliably correlate patterns of brain activity with aspects of experience, not to suggest that the brain has localized, high-fidelity, picture-like representations of things in the world.

processing. These cells and their connections make up a ‘sensorimotor systems’, organized into different ‘modalities’ (i.e. vision, audition, touch). The main function of these systems is to extract information from the environment (e.g. shapes, pitches, patterns of vibration). Embodied neurocognitive models of brain organization propose that these sensorimotor systems of the brain (e.g. those involved in vision, audition, touch, etc.) are relatively specialized to ‘map’ information in the world.

One such model is presented by Meyer and Damasio (2009), though other models share similar features (e.g. Barsalou et al., 2003; Fernandino et al., 2016; Versace et al., 2014). According to this model, the major sensorimotor systems in the cortex are hierarchically organized, with cells lower in the hierarchies processing simple information and cells and networks higher in the hierarchies processing increasingly complex types of information. For instance, cells in the primary visual cortex (i.e. the first cortical region where information from the eyes is processed) process edges in the environment (i.e. lines), while cells in the inferotemporal cortex process complex collections of lines that make up faces and hands (see Orban et al., 2004, and Grill-Spector & Malach, 2004, for reviews of functional organization of occipitotemporal cortex; see Fuster, 2004 for frontal cortex organization; see Mesulam, 2012 for general review of cortical connectivity). Eventually, in the processing pathways of the cortex, sensorimotor information is combined in complex ways in ‘multimodal’ areas—areas that process information across multiple modalities. Importantly, a ‘sensorimotor-to-multimodal gradient’ is a primary organizational principle of the cortex (Huntenburg et al., 2018). Further, it is known that, anatomically and functionally, these hierarchies are reciprocally connected; that is, cells in lower areas communicate with cells in higher areas, sending information ‘upstream’, and cells in higher areas also communicate backwards or ‘downstream’. Thus, the cortex is made up of hierarchically organized, reciprocally connected networks that map sensorimotor information gleaned from the world in a simple-to-complex gradient.

Models like that of Meyer and Damasio (2009) hypothesize that the mind/brain’s representation of concepts is *distributed* in the information that is activated across *all* levels of these hierarchies. In general, because of known physiological processes (i.e. Hebbian process of ‘cells that fire together, wire together’), experiences that we have (e.g. of a bear) lead to coactivation of cells across the modalities (e.g. within somatosensory, motor, visual and auditory maps), increasing the likelihood that these maps are activated together. Because of reciprocal connectivity, the activation of one map (e.g. the visual information regarding the shape of a bear) can result in the activation of others (e.g. the somatosensory information regarding the feel of bear fur). This type of probabilistic ‘retroactivation’ is thought to be under the coordination of higher areas in the hierarchy that help keep track of the coactivation patterns across different sensorimotor maps. We can think of these regions as ‘cognitive controllers’. These multimodal controller regions can retroactively activate distributed components of sensorimotor experience depending on our goals. In these embodied models, then, there is no single locus of the representation of a concept, and the representation of the world is activated in a task-dependent manner. Overall, the learning and use of concepts is supported by this 'maps and controllers' structure of the mind/brain (see Fig. 3.1 for further details).

These types of models are consistent with the broader research literature in cognitive psychology and neuroscience showing that concepts are represented dynamically (see Yee & Thompson-Schill, 2016 for a review). Indeed, we know that concepts are flexibly applied: We can create ad hoc categories like 'the category of things you bring to the beach', and these category structures can shift when taking the perspectives of different people (see Barsalou, 1987, for a review). Importantly, recent neuroscientific research is showing how such flexibility and contextual sensitivity are instantiated in the cortex; specifically, the neural representation of different categories of objects is shaped by task goals (Çukur et al., 2013). Thus, current goals determine what information is activated to constitute a representation of any given concept in context. Overall, then, the picture of *how* concepts are represented that emerges from this framework is one in which concepts are represented by the activity of distributed, hierarchical neural maps that are non-consciously activated in probabilistic ways in response to the things in the environment.

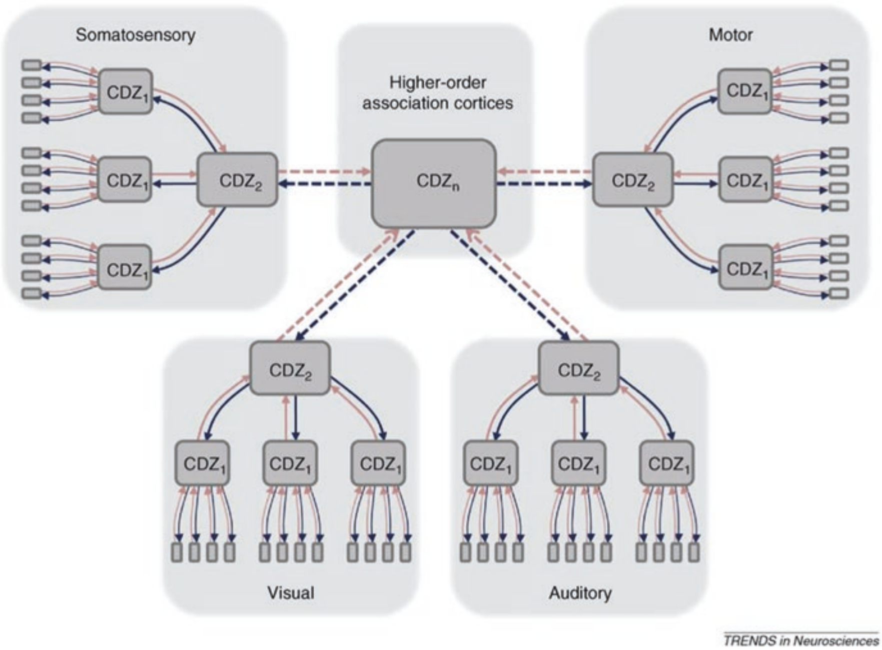

Fig. 3.1 A schematic representation of the neural organization that supports our conceptual thinking. (Taken from Meyer & Damasio, 2009). Sensorimotor systems are organized hierarchically, with cells in lower regions processing simple information (e.g. lines in vision) and cells in higher regions processing more complex information (i.e. shapes). These sensorimotor maps converge in higher regions (here, labelled CDZs) that control the reactivation of information throughout this system. For these types of models, concepts are understood as distributed representations of information across these sensorimotor systems that are activated in a context dependent manner

What the Brain Represents: Affordances

I have shown that the neural organization of sensorimotor systems constrain *how* concepts are represented in the mind/brain. Importantly, this organization is not just a structural curiosity because it ultimately constrains *what*, exactly, is represented. Embodied views emphasize that sensorimotor systems are embedded in a body with a particular morphology (in humans this means bipedal, dexterous appendages, forward facing, of particular mass, etc.; Varela et al., 2016) and constitute a system geared towards homeostatic action; that is, we act to bring ourselves into favourable relationship with the world (Friston, 2010). Importantly, the main consequence of this embedding is that our sensorimotor systems seek to represent the affordances within the world. Here, affordances are defined as information in the world that allows for particular actions (e.g. the shape of a hammer allows for particular grasping and manipulation, the sound of a bear allows for certain approach/avoid activities, etc.; see Tucker & Ellis, 1998).[9] One way of thinking about this is to suggest that we live in a world of affordances, and what we 'see' or 'hear' is not bears per se, but merely the consequences of extracting the affordances for action (i.e. runawayability, manipulability, etc.). The radical consequence of this perspective is that the contents of our minds/brains (i.e. of our distributed sensorimotor representations) are not actually about *things* but are about *opportunities*.

Neurophysiological research in non-human primates supports this idea. For instance, Cisek and Kalaska (2010) have shown that there are cells within higher regions of the motor system that represent different action affordances. For instance, some cells help control a right-ward grasp, while other cells control a left-ward grasp. Importantly, these cells become active automatically in response to stimuli that afford such actions; for instance, when the animal is shown an object it has learnt to use, these cells are active before a behavioural response is made. Thus, different action affordances are activated automatically in response to the environment even before an action takes place. This suggests that the brain is constantly assessing the action affordances of the environment; for the brain, to interpret the world is to, in part, interpret it for action opportunities, preparing the body for effective action in the world.

The idea that concepts are constituted by affordance information is consistent with the general research literature on concepts. Specifically, we know objects can

[9] Note that the term affordance is used to describe a wide number of relationships/structures relevant to physics and neuroscience (see Osiurak et al., 2017, for a taxonomy). Gibson (1979) discussed the idea in an effort to avoid representationalist models of the mind and suggested that affordances are a property of the possibilities of an organism/environmental interaction; here, I adopt a definition that is more general and used more widely—perhaps in violation of Gibson's proposal (Goldman, 2012)—in cognitive neuroscience. In this case, affordances are descriptions of organism/environmental possibilities for action and are predictively correlated with neural activity. Importantly, the notion of representing affordances used here—a correlative relationship between distributed brain activity and information in the environment—is not necessarily incompatible with other ecological conceptions of mind (see Golonka & Wilson, 2019).

be categorized at three levels: superordinate (e.g. tool), basic (e.g. hammer) and subordinate (e.g. claw hammer) (see Rosch, 1978). These categorizations serve different purposes. It is known that people categorize objects most commonly at the *basic* level (i.e. people will more readily identify a picture of a hammer as a hammer than as a tool or as a sledge hammer; children learn to successfully identify hammers as hammers before they identify them as claw hammers, sledge hammers, etc. or even as tools per se). Why does the basic level have such primacy? Rosch et al. (1976) have demonstrated that basic categorization tends to maximize the affordance similarities of objects. Claw hammers and sledge hammers are interacted within more or less the same way, and therefore their basic-level categorization is most relevant for organizing behaviour. Thus, our concepts are organized around the basic categories which are defined by actions and our bodily interactions.

Overall, then, concepts are not only constituted by distributed, hierarchically organized, sensorimotor information in the mind/brain, but, because we have a particular body and occupy particular environments, *what* is represented is information relevant to the organization of successful action—the affordances of the environment and our interactions in it. This is what it means for concepts to be 'embodied'.

Consequences of Embodiment for Our Understanding of Concepts

I have shown that the framework of embodied cognition leads to a particular view of how the mind/brain represents concepts and what, specifically, is represented. Concepts are constituted by distributed, hierarchically organized, sensorimotor information in the mind/brain. Further, because we have a particular body and act in particular environments, what is represented in the mind/brain is information relevant to the organization of successful action—the affordances of the environment and our interactions in it. Thus, the meaning of our concepts is grounded in embodied action.

There are two important consequences of this perspective on concepts (see also Casasanto & Lupyan, 2015). First, if distributed, hierarchical, sensorimotor representations are activated probabilistically and in context, then there is *within-person* variation in how conceptual thinking is supported. For instance, the concept BLACK and BEAR cannot be understood as simply adding the abstract neural symbols for BLACK and for BEAR but requires activating, under the control of multimodal neural regions (i.e. the controllers), distributed information pertaining to colour, shape, sound and perhaps aspects of somatosensation (the feel of fur) and motor activity (the steps used to keep distance or run away, i.e. the maps); that is, all the information pertaining to the affordances of the thing are relevant for our interactions with it. Because these activations are context-specific, it suggests that in some contexts thinking about BEARs might primarily involve activation of the visual shape information with little contribution from motor information; conversely, another context might support the reverse (more motor activation, little shape activation). Thus, the flexible activation of all of this information is constitutive of the concept BLACK BEAR in context—not an amodal representation that is activated every time we think about bears.

A second major consequence of this view is that people who have different experiences will have slightly different representations of the same concept; therefore, there is *between-person* variation in how conceptual thinking is supported. For instance, if you have never experienced the somatosensations of bear fur (because you have never petted one) but I have (because I grew up near a museum that had samples of bear fur), all other 'bear' experiences being equal, our distributed representations of BLACK BEAR will be differentiable, ever so slightly, by the unique information that constitutes them. Indeed, hammers afford building for some people with requisite sensorimotor experience (e.g. carpenters), but not for others; bears afford exciting wilderness experiences for people with requisite sensorimotor experience (seasoned hikers), but not for others. In this way, we can come to understand that experience, shaped by our bodies and interactions with the world, determines what affordances there are for us with respect to any given concept and what information the cortex represents to support our behaviours.

Overall, then, though it might *feel* like we are thinking the same thing every time we think of BLACK BEARS, and it *feels* like we are having the same thoughts when we share a discussion together about BLACK BEARS, embodied models suggest that each instance is constituted by slightly different types of sensorimotor information reflecting different affordances in context. Armed with this counterintuitive conclusion, we can now assess the role embodied concepts play in our understanding of truth.

Consequences of Embodied Concepts for Science and Truth

The embodied framework defines what it means to learn and use concepts. It also suggests limits to how we define truth. In some cases, two people will have similar experiences, and their concepts will be constituted by similar types of information and reflect similar affordances. However, the great variety of human experience will guarantee that this is not always the case. More specifically, it is only when there is *sufficient overlap* in our experiences from one moment to the next, and only when there is *sufficient overlap* in our experiences and another person's will we have sufficient (but not complete) overlap in what BLACK BEAR *means* to us (see Lakoff & Johnson, 1999). However, if we don't have sufficient overlap, two people's concepts won't *mean* quite the same thing. I am not simply suggesting that having similar experiences makes it easier for us to relate to one another, nor am I suggesting the mundane fact that different people might feel differently about bears; rather, I mean quite literally—in the sense of the representation of the concept in the mind/brain—that the *sharing and understanding of a concept requires sharing overlap in our sensorimotor experiences*.

The embodied view of concepts described here shows that understanding a concept is a type of agreement—not conscious and open for discussion, but activated non-consciously in our sensorimotor systems—between people. Thus, the truth statements we make are shaped by the completeness of this agreement. Imagine a bear encounter, and you say, 'There is a bear over there'. Is this true? What we can verify is that my distributed representation leads me to behave in similar enough ways to you in this context. That is, as long as my distributed representation allows me to say the word 'bear' when you and I are walking along the shore of the lake, and those words probabilistically activate a sufficiently shared distributed representation for you and me to coordinate a safe response, we have a shared understanding. This is what it means for us both to understand the 'truth' that 'There is a bear over there'. But note that this shared understanding arises not because we share the same abstract, mental symbol that consistently (though arbitrarily) corresponds to the bear, and therefore we cannot say we have homed in on a mind-independent reality. Rather, our shared understanding arises because the overlap in our distributed representations of the affordances of the environment, gleaned from our sensorimotor experience, is sufficient to support our coordinated behaviour. Indeed, someone with drastically different sensorimotor experiences, and extracting different affordances from the presence of a bear, is living in a different conceptual world.

Overall, this idea allows us to update James's (1907) message regarding pragmatism: 'Ideas…become true just in so far as they help us to get into satisfactory relations with other parts of our experience' and 'The truth of an idea is not a stagnant property inherent in it' but rather 'Truth *happens* to an idea. It *becomes* true, is *made* true by events' (emphasis in original). Rather than reflecting an immutable aspect of a mind-independent reality, using concepts entails assessing the overlap of our distributed representations through actions; that is, truth is determined only pragmatically. We share concepts to a greater or lesser degree depending on whether there is sufficient overlap in our representations to coordinate our behaviour. What is true cannot be simply a matter of aligning our symbolic concepts in the mind/brain to things out there in a mind-independent world. Instead, two (or more) people's concepts support their creation of a shared, pragmatically defined truth.

This presents a clear challenge to experimental psychology. We want to be able to say things like 'depressed people are more creative' or 'the anterior temporal cortex is the neural basis of semantic memory', but I have just argued that there is no way we can understand those statements and their concepts by matching symbols in our minds/brains with a mind-independent reality. So what are we actually doing? Pragmatism suggests that when I say that depressed people are more creative, I am stating that I expect that your distributed representation of the relevant concepts will be sufficiently overlapping with mine and that this overlap can support our coordinated behaviour. This is often the case, and, indeed, the purpose of operationally defining constructs in empirical study is to try to ensure this as much as possible. When this is occurring, we have the sense that we are identifying truths using our science. However, according to the view I lay out here, we can only ever hope for *sufficient overlap* of our distributed representations that leads *to pragmatic* understanding of the concepts, where pragmatic means taking empirical measurements and intervening on the world in similar ways with the same expectations of outcome for our behaviour. This is the only way for us to share concepts of creativity and depression and to identify truths from them.

However, as scientists, like people walking along a lake, we won't always have sufficient overlap in our experiences, and therefore we will often use different types of information to represent a concept. For example, say for you the sensorimotor information that constitutes your representation of 'depression' is primarily built from your sensorimotor experiences with physiological measurements in the lab; conversely, say for me it is primarily visual and built from my sensorimotor experiences watching the behaviour of people with a depression diagnosis. Further, because representations are about sensorimotor affordances, my representations are really about different opportunities for action than yours (e.g. how I would intervene in an attempt to increase creativity). Thus, despite using the same word (i.e. 'depression'), we are quite literally talking about different things in terms of how the concept is represented in our minds/brains. My ability to understand your writing up of an experiment, its conclusion and implications and to conceptually replicate the finding that depressed people are more creative in my lab is constrained by these differences in representation; despite using operationalized definitions to make measurements in the lab, our use of concepts in interpreting and communicating those definitions (i.e. the ways in which *we understand them*) will influence how we do science.

This constraint on understanding is especially relevant as concepts become increasingly abstract and extended beyond their operational definitions in the practices of scientists (e.g. giving popular science presentations). Most concepts psychologists use are operationally defined in multiple ways (e.g. 'creativity' is operationalized through numerous tasks, each with their own set of embodied concepts), and there are abundant concepts that are rarely (or never) subjected to attempts at operationalization at all in discussion sections of papers (or grant applications, etc.). For instance, how many psychologists (and their students) just 'know' what the concepts memory and mental health all *mean*? The embodied models reviewed here show us that, despite using same words and striving for operationalism, and despite there being sufficient overlap of conceptual representation in many cases, we do not share an understanding of the immutable, mind-independent reality of something called 'memory'. It's pragmatically defined embodied concepts all the way down.

These consequences extend to all areas of sociopolitical life and are exacerbated when effort is not made to understand the representations—the content and structure—of the concepts we use. Importantly, the role of conceptual understanding in moral and political life, and the requirement of sufficient overlap in representation in particular, has been pointed out by embodied researchers studying language for over a decade (though without concern for the neural mechanisms involved; see Lakoff, 2008, 2010). When a politician suggests that they stand for freedom, they hope that the word freedom activates sufficiently overlapping distributed representations in their constituents and that they can coordinate actions that are pragmatically effective. In politics, as in science, this cannot be guaranteed. For instance, what we call taxes provide certain opportunities for action (i.e. increase movements through the world) to some people and limit opportunities for action (i.e. decrease movements through the world) to others. Because opportunities for action are partially constitutive of an individual's understanding of concepts, it is really no surprise that you have people on all sides of the political spectrum criticizing their opponents for failing to see 'reality' about taxes.[10] The empirical research reviewed here helps us understand why this is so.

Overall, then, truth crises in science and society do not arise because some people have concepts that correctly correspond to reality and others don't. Instead, truth crises arise, partially, because there is some degree of multiplicity in every concept, in different people and over time, and there can thus be *insufficient* overlap in the representation of a concept to pragmatically support coordinated behaviour. When this happens, psychological science fails to accumulate reliable conclusions, and politicians have license to spread ideas about 'alternative facts'.

The Way Out of Psychology's Truth Crises

The embodied account of concepts discussed here puts a constraint on what truth is and accounts in a particular way for the multiplicity of meanings in conceptual thinking. Given this, how can we arbitrate statements like 'the anterior temporal lobe of the cortex is the basis of semantic memory?' Does the embodied approach suggest truth is 'relative'? But if truth is relative, why will I not try to prove it by jumping out of a ten-storey window?

Let's assume[11] the universe is full of some stable or semi-stable patterns, and the goal of science, and indeed the goal of any truth-seeking activity, is to detect, interpret and make use of these patterns. Let's also assume that death is a great arbiter of these patterns, as it can be seen as the ultimate pragmatic feedback from the world about the success of our actions. Given my body's sensorimotor constraints, and given that you and I share similar enough bodies that determine affordances in similar enough ways, and given that we find ourselves in similar enough contexts, we can agree that I will die if I jump out the window, and this agreement supports the pragmatically defined concept of gravity (and all the calculations of estimated velocity when I impact the ground, etc.) without the need to reify it as mind-independent reality. That is, the affordances of the situation, defined by our embodiment, are going to be similar enough for *most* people to agree that the concept of gravity reflects a pragmatically defined truth.

[10] Note again that I am not talking about debates about our personal opinions on taxes. I am talking about differences in the sensorimotor information and affordances that underlies the concept of taxes in different people that render discussions difficult because of a lack of shared representations in the mind/brain and therefore a lack of shared, pragmatic, understanding.

[11] Wishing to avoid the metaphysical discussion (and indeed I am not a metaphysician), this appears to me to be a basic, safe assumption, for most experimental psychologists.

Experimental psychology is also about detecting stable patterns in the universe, patterns pertaining to mind, brain and behaviour. Like with the truth of gravity, the truths of psychological science will be pragmatically determined by our shared experiences. However, most of our questions are not readily arbitrated by any jump-out-the-window test (or any of the practical tests seen in the physical sciences, biology, engineering, etc. that can result in catastrophic loss of life). Rather, the truths in psychology will be determined by how likely it is that our embodiment will ensure that we will converge on the same concepts. The more stable the pattern and the more we share embodied experiences, the more likely we are to pragmatically uncover truths. For instance, perhaps aspects of light spectra are relatively stable in the universe, and our shared embodiment increases the likelihood that we converge on the same truths regarding the perception of colour and the concepts we use to describe it. We can anticipate that uncovering truths about creativity and the prefrontal cortex will be more challenging, but they should still be possible, provided what we are homing in on is supported by consensus. However, the less overlap we share in our sensorimotor experiences, the less likely we can converge on a pragmatically shared truth. In such cases we will have non-overlap in our concepts, more difficulty communicating truths using these concepts and less consistent replication within our science. Thus, for some concepts in psychological science, we might not have any hope of pragmatically agreeing on some truths in this sense because they can be constituted by different types of sensorimotor information in too many ways. For instance, I can suggest that, given the multiple ways in which embodiment and sensorimotor experience might support various understandings of concepts such as memory and mental health, the likelihood of converging on single universal truths regarding these is quite low.[12]

Psychology, as a science, can address its truth crisis with an open acknowledgement of the multiplicity of conceptual understanding, and the ‘embodied pragmatism’ discussed here prevents us from succumbing to radical, relativist, conclusions that a shared understanding is arbitrary and/or hopeless. It allows us to reject correspondence theory and its realist connotations in favour of an approach to truths that remain arbitrated by our embodiment. To address the truth crisis, we need to seek to uncover the ways in which our conceptual thinking supports our design of experiments, what motivates them and how we interpret results. Abandoning a (implicit or explicit) commitment to correspondence theory of truth and its associated realism, but committing to pragmatism bolstered by the insights from embodied cognition, ensures that conceptual thinking leads to uncovering more truths in psychological science and better brings us into a satisfactory relationship with our world. This allows us to reduce our aggressive adherence to the truth of our findings and, more broadly, our sociopolitical discourse. Embodied concepts help explain our collective susceptibility to truth crises and bring us towards a more consciously

[12] However, it is worth keeping in mind that light—a presumably stable, non-random pattern in the universe—can be understood both as particles and as waves and that many valuable discoveries in our shared reality have ensued from both modes of conceptual understanding, e.g. the trichromatic and opponent process theories of colour vision.

shared—though not mind-independent—understanding of truth. From this perspective truth is far from dead.

Acknowledgements I would like to thank Davood Gozli, Jaan Valsiner and Nicole White for their detailed feedback on earlier drafts of this chapter, as well as Davy Mougenot for his constructive discussions of its central thesis. In addition, I would like to thank the 2020 class of PSYC 432 at UNBC for a term-long discussion of the main thesis and for their feedback on the earliest version of this chapter.

References

Anderson, M. L. (2014). *After phrenology* (Vol. 547). MIT Press.

Anvari, F., & Lakens, D. (2018). The replicability crisis and public trust in psychological science. *Comprehensive Results in Social Psychology, 3*(3), 266–286.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101–140). Cambridge University Press.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577–660.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7*(2), 84–91.

Berkman, E. T., & Wilson, S. M. (2021). So useful as a good theory? The practicality crisis in (social) psychological theory. *Perspectives on Psychological Science, 16*(4), 864–874. 1745691620969650.

Casasanto, D., & Lupyan, G. (2015). All concepts are ad hoc concepts. In E. Margolis & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 543–566). MIT Press.

Chemero, A. (2011). *Radical embodied cognitive science*. MIT Press.

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience, 33*, 269–298.

Çukur, T., Nishimoto, S., Huth, A. G., & Gallant, J. L. (2013). Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience, 16*, 763–770.

Davidson, D. (1984). *Truth and meaning. Engagements across philosophical traditions,* 69.

Dowden, B., & Swartz, N. (2020). Truth. The Internet Encyclopedia of Philosophy (IEP). ISSN 2161–0002, https://www.iep.utm.edu/, 2020, May 21.

Engel, A. K., Friston, K. J., & Kragic, D. (Eds.). (2015). *The pragmatic turn: Toward action-oriented views in cognitive science* (Vol. 18). MIT Press.

Fawcet, J., & Matheson, H. (2019). *Conducting research fast and slow: The importance of valuing scientific process over scientific products*. Retrieved from https://socialsciences.nature.com/posts/54628-conducting-research-fast-and-slow-the-importance-of-valuing-scientific-process-over-scientific-products

Feldman, J. (2008). *From molecule to metaphor: A neural theory of language*. MIT Press.

Fernandino, L., Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W. L., et al. (2016). Concept representation reflects multimodal abstraction: A framework for embodied semantics. *Cerebral Cortex, 26*(5), 2018–2034.

Foglia, L., & Wilson, R. A. (2013). Embodied cognition. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(3), 319–325.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127–138.

Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in Cognitive Sciences, 8*(4), 143–145.

Gibbs, R. W., Jr. (2005). *Embodiment and cognitive science*. Cambridge University Press.
Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton, Mifflin and Company.
Goldman, A. I. (2012). A moderate approach to embodied cognitive science. *Review of Philosophy and Psychology, 3*(1), 71–88.
Golonka, S., & Wilson, A. D. (2019). Ecological representations. *Ecological Psychology, 31*(3), 235–253.
Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience, 27*, 649–677.
Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena, 42*(1–3), 335–346.
Hommel, B. (2015). The theory of event coding (TEC) as embodied-cognition framework. *Frontiers in Psychology, 6*, 1318.
Huntenburg, J. M., Bazin, P. L., & Margulies, D. S. (2018). Large-scale gradients in human cortical organization. *Trends in Cognitive Sciences, 22*(1), 21–31.
James, W. (1907). *Pragmatism: A new name for some old ways of thinking*. Retrieved on https://www.gutenberg.org/files/5116/5116-h/5116-h.htm
Lakoff, G. (2008). *The political mind: A cognitive scientist's guide to your brain and its politics*. Penguin.
Lakoff, G. (2010). *Moral politics: How liberals and conservatives think*. University of Chicago Press.
Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought* (Vol. 640). Basic Books.
Margolis, E., & Laurence, S. (Eds.). (1999). *Concepts: Core readings*. MIT Press.
Matheson, H. E., & Barsalou, L. W. (2018). Embodiment and grounding in cognitive neuroscience. *Stevens' handbook of experimental psychology and cognitive neuroscience* (Vol. 3, pp. 1–27).
McIntyre, L. (2018). *Post-truth*. MIT Press.
Mesulam, M. (2012). The evolving landscape of human cortical connectivity: Facts and inferences. *NeuroImage, 62*(4), 2182–2189.
Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neurosciences, 32*(7), 376–382.
Millikan, R. G. (2017). *Beyond concepts: Unicepts, language, and natural information*. Oxford University Press.
Murphy, G. (2004). *The big book of concepts*. MIT Press.
Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W.G. Chase (Ed.), *Visual information processing: Proceedings of the eighth annual Carnegie symposium on cognition, held at the Carnegie-Mellon University, Pittsburgh, Pennsylvania*, 1972, May 19. Academic Press.
Newell, A. (1980). Physical symbol systems. *Cognitive Science, 4*(2), 135–183.
Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review, 26*(5), 1596–1618.
Orban, G. A., Van Essen, D., & Vanduffel, W. (2004). Comparative mapping of higher visual areas in monkeys and humans. *Trends in Cognitive Sciences, 8*(7), 315–324.
Osiurak, F., Rossetti, Y., & Badets, A. (2017). What is an affordance? 40 years later. *Neuroscience & Biobehavioral Reviews, 77*, 403–417.
Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science, 7*(6), 531–536.
Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science, 7*(6), 528–530.
Pine, D. W. (2017). *Is truth dead? Behind the time cover*. Retrieved from https://time.com/4709920/donald-trump-truth-time-cover/
Prinz, J. J. (2004). *Furnishing the mind: Concepts and their perceptual basis*. MIT Press.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Erlbaum.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382–439.

Shapiro, L. (2019). *Embodied cognition*. Routledge.

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology, 69*, 487–510.

Spellman, B. A. (2015). A short (personal) future history of revolution 2.0. *Perspectives in Psychological Science, 10*(6), 886–899.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 830.

van Rooij, I., & Baggio, G. (2021). Theory before the test: How to build high-verisimilitude explanatory theories in psychological science. *Perspectives on Psychological Science, 16*(4), 682–697.

Varela, F. J., Thompson, E., & Rosch, E. (2016). *The embodied mind: Cognitive science and human experience*. MIT Press.

Versace, R., Vallet, G. T., Riou, B., Lesourd, M., Labeye, E., & Brunel, L. (2014). Act-in: An integrated view of memory mechanisms. *Journal of Cognitive Psychology, 26*(3), 280–306.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–636.

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review, 23*(4), 1015–1027.

Chapter 4
Operationalization and Generalization in Experimental Psychology: A Plea for Bold Claims

Roland Pfister

Introduction

I read my first book on experimental psychology a little more than 15 years ago while eagerly waiting for my civil service to end. It probably was about time for me to dig into this subject as I read the book hoping to get a better grasp on what to expect from the degree program in psychology that I had just enrolled in quite naively. Having expected psychology to be a rather wordy and potentially woolly subject, the concept of using controlled experiments to unravel the inner workings of the minds was enticing. The science of experimental psychology promised astonishing methodological rigor. It promised a clearly defined and tractable world of controlled experimental conditions to base its conclusions on objective, transparent, and verifiable procedures. And it promised an exciting journey toward answering the big questions of human conduct.

This spirit does not quite align with how Davood Gozli's (2019) *Experimental Psychology and Human Agency* depicts the current state of the discipline. Unfortunately, his assessment appears to be careful, correct, and comprehensive. It

Author Note
The outline of this chapter was drafted during a research stay at the Center for Interdisciplinary Research, Bielefeld, and this chapter would not have come to life without their generous support. I thank Leonhard Höhnel for contributing his beautiful artwork, Felicitas Muth for her critical comments on an earlier version of this argument, Robert "How2Rulebreaker" Wirth for his continued obsession with research on rule-violation behavior, and Davood Gozli for taking this work at least as serious as we do.

R. Pfister (✉)
Department of Psychology III, University of Wuerzburg & Center for Interdisciplinary Research, Bielefeld University, Bielefeld, Germany
e-mail: roland.pfister@psychologie.uni-wuerzburg.de

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9_4

highlights some of the major issues surrounding many contemporary investigations in experimental psychology. That is, while mainstream discourse has been preoccupied mainly with empirical and statistical discussions over the last years, Gozli claims that experimental investigations of the human mind have fallen prey to a much more profound issue in that they neglect a proper theoretical assessment, discussion, and integration of research findings. I believe that this far-reaching claim warrants a closer look.

Tasks as Means, Tasks as Ends

The business of experimental psychology is to uncover how the mind does what it does, and its major tool to achieve this aim is to expose participants to tasks that they are instructed to perform. Following this view, the task becomes a critical means to assess theoretical models of how the mind works (Gozli, 2017, 2019; Hackman, 1969). This role of tasks as scientific tools to construct, test, and specify theoretical ideas presupposes that experimental psychologists should not be interested in a given experimental task per se, for the task is only a means to an end. They should instead be interested in formulating a theoretical idea that eventually applies to situations that transcend the specific task setting in which it happens to be investigated in the first place. This interest in removing task characteristics from the picture takes a prominent spot in Gozli's critique (2019). In fact, he offers two supposedly widespread motives for keeping the task in the background (pp. 12–13):

> It is important that we distinguish between the removal of task characteristics from psychological theories, as a scientific ambition, and the removal of task characteristics from description, as a rhetorical strategy. De-emphasizing the role of tasks in the production of research findings has rhetorical advantages. It allows us to make overly-general claims about human capacities. It leaves the social and normative dimension of the experiment out of consideration, treating participants, or models of the average participant, as isolated entities (Billig, 2013). By de-emphasizing the tasks, experimenters also de-emphasize the mutual understanding of the tasks, achieved through language, which then allows experimenters to maintain their focus to attributes of performance. By neglecting the social-normative dimension of the experiment, we ignore the fact that scientific activity is conducted within a social and cultural context.
>
> Another rhetorical advantage of keeping the task in the background is that it facilitates a type of bait-and-switch trick performed on the audience of the research, including funding agencies and incoming members of the discipline. This trick involves borrowing a concept from its everyday domain to justify the research project. The concept may have rich and varied meanings in its original contexts of use, but after it is operationalized as an attribute of an experimental task, its meaning changes (Smedslund, 1997; Teo, 2018). This creates a gap between the everyday meaning of the concept and the meaning within the experiment. If we de-emphasize the task, we can de-emphasize this gap, talking about the findings as if they apply equally to the experiment and to everyday contexts.

The remainder of this chapter is dedicated to reflecting on this excerpt. By doing so I do not intend to reduce Gozli's (2019) critique to this particular view, especially because many points raised in his book offer striking insights into the current state of the discipline. In fact, I found myself agreeing with many themes of this critique so enthusiastically that the only room for discussion arises from how he assesses the role of task characteristics in the psychological literature. Because tasks take such a prominent spot in experimental psychology, however, I believe that a thorough reflection on this particular quote – or rather, an exegesis of it – may highlight potential avenues to improve theorizing in the field. By extension, these avenues may also be apt to improve *empiricizing* at the same time. With the term empiricizing, I refer to the bread and butter of the empirical scientist: the business of conducting empirical work for any reason, be it to test a theoretical idea, to teach students how to conduct and interpret experiments, or any other motivation one can think of.

The following sections will thus reflect on the validity of Gozli's (2019) conclusions on how task characteristics are downplayed in psychological research. Based on a somewhat divergent perception of the field, I will argue that task characteristics are in fact over- rather than under-represented in some areas of experimental psychology. A way to resolve the limitations that come with neglecting and overly attending to task characteristics alike would be to re-embrace the power of explicit operationalization. I will apply this reasoning to the specific case of research on rule-violation behavior (due to my own preoccupation with the subject matter) before outlining a more general rationale of how theory and empirical work should interact in experimental psychology and related fields.

While doing so, I agree that both rhetorical strategies of the above quote – omitting task characteristics in order to derive overly general claims and omitting task characteristics to downplay the gap between experimental setups and everyday contexts – are certainly conceivable, and, anecdotally, I feel that I have made use of both of them myself (with varying levels of guilt for doing so). But how common are these strategies actually in the broader field of experimental psychology? First and foremost, they do not appear to be equally common across different researchers and different sub-disciplines. When browsing through recently published issues of relevant journals, it seems that some areas – e.g., experimental approaches in the social and developmental literature – do indeed tend to play down task characteristics quite routinely. The state of these fields thus seems to conform to the above quote.[1] I believe that matters appear quite different when turning to the field of cognitive psychology, however. Even though many of the studies discussed by Gozli (2019) fall into this domain, I feel that contemporary cognitive psychology is plagued by the very opposite limitation. Rather than playing down task characteristics, certain tasks have actually become the target of most scientific efforts in this field. That is, rather than seeing tasks as means to study some theoretically interesting idea, certain tasks have become an end of their own for many experimental psychologists sailing under the cognitive flag (see also Meiser, 2011).

[1] This claim might warrant additional evidence in the form of exemplary citations (quite a few of which readily come to mind) or a quantitative corpus analysis of a larger body of the published literature. I opted not to include any specific references because this chapter does not intend to point fingers at particular articles or journals. Instead, the assessment is meant to describe my subjective perception of the field.

The continued preoccupation with tasks is apparent in many colloquial presentations of research findings on conferences and workshops around the globe. These presentations would often start with icebreakers such as "Are you familiar with the negative priming paradigm?" or "I guess you have heard of task switching before!" Each of these paradigms might be used to study different theoretical twists, however (Frings et al., 2015; Kiesel et al., 2010), but these eventual theoretical implications tend to find themselves overshadowed by discussions over empirical and methodological fine print. More often than not, these fine-grained issues hide behind three-letter acronyms such as SOA (short for stimulus-onset asynchrony), ITI (inter-trial interval), or RSI (response-stimulus interval), and they often revolve around timing specifics and other variables that lend themselves to parametric manipulations (eccentricities, relative frequencies, sequential positions, and the like). Scrutinizing parametric manipulations of timing-related variables, even those with three-letter acronyms, may of course be viable and highly informative at times. This approach requires a thorough theory of mental processing to live up to expectations, though (for a positive example, see, e.g., Pashler & Johnston, 1989). One could of course argue that colloquial presentations omit such theoretical explications for rhetorical reasons and that these discussions help distill theoretically meaningful ideas in the end, but the published literature is similarly plagued by reports that seem to study certain experimental paradigms for their own sake without making meaningful connections to the theoretical landscape. One way to resolve this limitation of current practices is embracing the use of explicit operationalization much more strongly than currently done in the field.

Operationalization

In my judgment, it seems that the argument of playing down task characteristics does apply to certain sub-disciplines in experimental psychology, but it does not seem to map too closely onto the current state of cognitive psychology. Here researchers often seem to be preoccupied with studying task characteristics in the first place rather than systematically playing down such aspects of their work. Fortunately, the two strategies are not mutually exclusive when taking different reports of research findings into perspective, even reports published by the same researcher or group of researchers on different occasions. It seems perfectly possible to spend a major share of one's career empiricizing about highly specific and paradigm-centric methodological concerns but to jump to conclusions by playing down task characteristics on other occasions. I would thus argue that many experimental psychologists are indeed mindful of task characteristics and their role in producing empirical observations and that these factors are discussed at length in the field. This discussion is largely disconnected from occasions on which authors try to highlight theoretical ideas, however.

Perhaps surprisingly, the two different views of how task characteristics are discussed in the experimental literature converge on the same outcome when assessed in more general terms. Irrespective of whether task details become marginalized when reporting research findings or whether they become the main target of experimental endeavors, researchers adopting either one of these strategies will fail to build informed theories of human conduct due to a missing connection between theorizing and empirical work. In a similar vein, *Experimental Psychology and Human Agency* further highlights that findings from experimental psychology are not sufficiently scrutinized by philosophical and sociological analyses (Gozli, 2019). Ideally, such critique would go beyond addressing specifics of a certain task at hand, by scrutinizing how the field is shaped by the very fact of involving tasks in the first place (see also Ting, this volume). I fully agree that this dimension is hardly ever discussed in the field at present and would deserve a more prominent spot, even though I will focus on task-driven research in the following.

One particular technique to connect theory and empirical work is to derive an experimental setup – or "paradigm" as commonly dubbed in experimental psychology – from a theoretical idea. This process of explicitly operationalizing a theoretical idea seems to have fallen out of favor at least in the cognitive literature (here matters seem to be less critical in fields such as social psychology, again based on my subjective perception of the literature). Mainstream research in cognitive psychology seems to have reached a tacit agreement that certain paradigms can be employed without much justification. This is particularly apparent in research on cognitive control where specific paradigms such as the flanker task (Eriksen & Eriksen, 1974), the Simon task (Simon, 1990), and the Stroop task (1935) are routinely employed in a wide range of studies. Only rarely do researchers discuss other potential uses of these tasks – say, to study selective spatial attention rather than cognitive control in case of the flanker task – and other potential facets of cognitive control that might be captured by distinct experimental setups (e.g., Dignath et al., 2014; Miyake et al., 2000). This development has further resulted in collections of basic cognitive tasks that are commonly regarded as useful tools to carry out psychological research (e.g., Bermeitinger, 2012). The danger associated with such a development is that the tasks seem to become meaningful entities of their own. Researchers might thus be tempted to target the specifics of particular tasks while failing to address the broader picture.

A potential defense against this criticism might be the argument that such considerations had been voiced when the field of cognitive control was still in its infancy (Botvinick et al., 2001; Hommel et al., 2004). Maybe such theoretical groundwork may indeed help set a specific experimental setup into context, but I believe that the somewhat antique exercise of operationalization will always help provide the audience of a research finding with a good sense of its potential relevance. Going one step further, I would argue that explicitly communicating the logic behind the researcher's operationalization is an ideal way to justify general conclusions that aim to transcend the task at hand. I will try to give an example for this claim by turning to work on rule-violation behavior as discussed by Gozli (2017, 2019).

Bold Claims: The Case of Rule-Violation Behavior

Tasks are composed of a number of rules that define which type of situation requires which type of action from the participants, and such rules tend to take the form of stimulus-response mapping rules in the context of behavioral research. Not surprisingly, such task rules have been studied in considerable detail in the cognitive literature, e.g., by assessing how effectively a new, instructed task rule is established and by measuring interference from opposing rules (Kunde et al., 2003; Meiran et al., 2014; Waszak et al., 2013; Wenke et al., 2009). These findings suggest that rules become ingrained quite deeply into the human cognitive system and, once established, take considerable mental effort to be overcome (Dreisbach, 2012). Such observations resonate with classic findings from social psychology, which showed a general preparedness of human participants to follow rules and norms (Asch, 1951; Cialdini & Goldstein, 2004). This latter strand of research is commonly subsumed under the labels of "conformity" and "obedience," and corresponding studies have indicated that participants would even take extreme actions if repeatedly commanded to do so by an authority (Blass, 1999; Milgram, 1963, 1974).

Theorizing in cognitive and social psychology stands in stark contrast to theoretical ideas that have been developed in the economic literature on cheating and dishonesty (Fischbacher & Föllmi-Heusi, 2013; Gneezy, 2005; Hilbig & Thielmann, 2017). Here, participants were reported to show a strong tendency to break the rule of an experimental task whenever rule-violation behavior would promise to maximize their payoff (for a critique of these studies, see Ting, this volume). When asked to report the outcome of a hidden die roll, for instance, the mean reported outcome routinely exceeds chance level when participants can secure monetary rewards based on their reports (Fischbacher & Föllmi-Heusi, 2013). Economists therefore suggested that rule adherence requires time and mental effort to overwrite temptations from potential, motivationally relevant outcomes (Bereby-Meyer & Shalvi, 2015; Shalvi et al., 2012; but see Foerster et al., 2013).

These and other findings from behavioral economics may be taken to suggest that rules are merely relevant for informing about potential payoffs and punishment for certain behavioral options (Becker, 1968). A closer look at the data pattern emerging from recent studies on economic games suggests that participants do not blindly maximize their rewards, however, but that they rather tend to misreport their outcomes only slightly and not to the maximum possible extent (Hilbig & Hessler, 2013; Ting, 2020). Classic economic theories on cheating and dishonesty do not predict such a pattern of results, nor can they accommodate such qualifying observations ex post facto, because they regard rules as relevant only for defining likelihood and severity of punishment for rule-breaking. A similar idea lies at the heart of prominent sociological theories such as the "general theory of crime" (Gottfredson & Hirschi, 1990; Pratt & Cullen, 2000). This theory suggests that human agents have a strong tendency to maximize their own rewards and will do whatever it takes to secure such "gratification" (to use the term employed by the theory). This framework thus regards rules as a mere tool to define punishments in order to deter agents from engaging in criminal actions if such actions were to promise interesting rewards.

Which of the two assertions is true? Do rules mainly feed into rational decision processes that balance expectancies and values (rewards) as suggested by economic theorizing? Or is the human cognitive system geared toward absorbing and internalizing rules and norms as suggested in the psychological literature?

This question can be answered by the dedicated study of actions that aim at breaking a given rule or norm as commonly done in the economic literature. Answering this question further requires an approach that is able to assess an empirical proxy of how cognitive processing unfolds in the course of a rule violation, however.[2] The common focus on decision outcomes in studies on cheating and dishonesty appears to be too coarse-grained to capture such processing, and we thus proposed an alternative setup in a series of studies (Pfister et al., 2016; Wirth et al., 2016). Our reasoning was as follows: Rule-breaking at the very least requires a situation in which an agent is aware of the rule and the behavior it prescribes, and he or she deliberately performs a different course of action. Such a minimal definition allows for distilling an experimental paradigm that operationalizes precisely these two components, which appear both necessary and jointly sufficient to study rule-violation behavior.

In our paradigm, therefore, we presented participants with a simple stimulus-response classification task and asked them to perform a mouse movement from the bottom center of the screen to either the top-left or the top-right depending on an imperative stimulus that appeared on screen. We used a small set of only two stimuli, and the mapping rule prescribing the correct response to each of these stimuli was instructed explicitly to ensure that participants would be aware of the rule and the behavior it implied for each situation. Critically, participants either followed the rule or acted against the rule on different trials. This was achieved either by asking participants before each trial whether they wanted to abide by the rules or break the rules (Pfister et al., 2016, Exp. 1) or by instructing one type of behavior (Pfister et al., 2016, Exp. 2; Wirth et al., 2016, Exp. 1).[3] We then sampled the trajectory of

[2] The vocabulary employed in this sentence as well as the methods described in the context of the following studies may suggest a theoretical relation to "action dynamics" accounts (McKinstry et al., 2008). This resemblance is coincidental, however, and this work was not performed with such a theoretical perspective in mind.

[3] Instructing participants to break a similarly instructed rule may seem somewhat unorthodox, because rule-breaking then becomes nested in a meta-rule of either following or violating an instructed stimulus-response mapping (Gozli, 2017). This is especially the case if both instructions emerge from the same source, e.g., from the same experimenter as in the case of our experiments. We still opted to do so because relying on free choices between rule-following and rule-breaking is plagued by a general reluctance to opt for rule violations so that it is difficult to find a control condition which comes with a similar experience for one or the other response. Crucially, even this artificial situation conforms to the minimal definition of rule-breaking as behavior that does not align with a rule. Whether participants do construe it the same way is a different question, of course (Gozli, 2019). This concern would be especially relevant if the results of the free choice condition had not replicated for instructed violations. Observing a similar pattern of results for instructed violations, however, seems to validate the experimental design.

the mouse cursor while participants performed their action and assessed whether the resulting trajectories would follow a straight path to their eventual target location or whether they would be attracted toward the target location on the opposite side of the screen. While rule-abiding responses followed a relatively straight path, rule-violation responses were deflected toward the opposite target location, i.e., the target location that would have represented the rule-abiding option. This observation suggests that rule-based responses were indeed retrieved even against the agent's intention of enacting a different behavioral option. The experiments further included a control group in which the same procedure was introduced not as a choice between rule-following and rule-breaking but rather as a choice between a standard task ("Task 1") and an alternate task ("Task 2"). This procedure ensured that, from the outside, the participants performed the very same actions in response to the very same stimuli as for the rule-violation instructions, but their options were now labeled as equally rule-abiding. The trajectories of this control group showed only a small deflection when responding according to the alternate mapping as compared to the standard mapping, and this difference was substantially smaller than the difference observed under rule-violation instructions. This even held true when the alternate task was introduced not as a separate task ("Task 2") with its own mapping rule but rather as having the opposite mapping of the standard task so that participants would have to negate the task rule themselves (Wirth et al., 2016, Exp. 3).

We took these findings to suggest that "Merely defining a rule, however arbitrary and irrelevant it may be, thus seems to prompt a tendency toward following it" (Pfister et al., 2016, p. 97). This is a strong statement. In fact, it seems to be a perfect example for an attempt to make overly general statements by jumping from a rather abstract, experimental task to broad, theoretical claims about human conduct. This view is certainly correct in that our interpretation is likely too broad and too general to represent a full-fledged theoretical take on rule-violation behavior. I would still defend this claim for at least two reasons.

The first reason for defending our interpretation is that our conclusions are safeguarded by an explicit operationalization. That is, we took care to build an experimental design specifically to meet criteria that had emerged from a thorough analysis of the theoretical construct at hand (rule breaking = "an agent is aware of the rule and the behavior it prescribes, and deliberately performs a different course of action"), and we justified the reasons for doing so explicitly. This axiomatic procedure safeguards against criticisms in terms of the generality of the conclusions, because the task was built in a theoretically motivated attempt to capture the very essence of rule-violation behavior across a wide range of situations. This procedure does not safeguard against criticisms pertaining to how successful our operationalization was, of course, including potential reservations regarding an impact of the experimental situation itself that may reinforce tendencies to follow rather than break rules (Gozli, 2017). It also does not address criticisms of the general methodology of isolating specific cognitive processes (or aspects thereof) in controlled experimental designs (Gozli & Deng, 2018).

The second reason for defending our interpretation is that interpretations of scientific findings are necessarily provisional in nature. They should always be read as a statement of the authors' current beliefs based on the available evidence. Here, the results of our experiments do not suggest any indication for exceptions to the hypothesized retrieval of rule-based actions, and this also holds true when carefully assessing what is available in the published literature. The only hints for an exception in this direction seem to be repeated violations of the same rule in close temporal succession (Wirth et al., 2018) and situations that suggest the rule to be negated at times (Imhof & Rüsseler, 2019). The latter finding emerged from a setting in which participants were asked to break the rules on some trials, whereas they were asked to respond according to a negated rule on other occasions. One way to view these findings is that the simultaneous operation of both instructions ("break the rule" vs. "negate the rule") caused the participants to fuse both meanings either by re-framing rule-breaking into following the negated rule or vice versa. Such a re-framing appears likely given that breaking rules and following a negated rule implied the same behavioral response. This view is further in line with ideas that task instructions will affect the salience of certain aspects of the experimental task, the employed stimuli, or the potential response options (Gozli, 2019). The state of the evidence therefore does not allow for a strong case based on either of the two findings at present, neither is there a compelling theoretical alternative to explain the present database (note that accounts in terms of instruction-induced salience cannot easily explain the observed difference between "violation" instructions and instructions to select a separate "Task 2" without assuming a special role of the concept of rules). Additional evidence and observations will likely topple our interpretation sooner or later, but until such evidence is available, I believe it is useful – actually: desirable! – to aim at maximally general claims. After all, experimental work can only be conducted, interpreted, and discussed when assuming generalization of research findings. If one portrayed experimental observations to be specific to the precise circumstances from which they emerged – to name a few that apply to the studies discussed above: experiments conducted in the Röntgenring 11 buildings of Würzburg University in the time from November 2012 to January 2013, testing a WEIRD sample (Henrich et al., 2010) on an experiment that was displayed on an old-fashioned cathode-ray monitor and operated by a non-ergonomic Logitech™ optical-corded USB mouse, then there is no point conducting experimental work in the first place. Any empirical efforts will always come with an infinite number of specificities, and it will thus not be possible to assemble an exhaustive list of potential confounding variables. If one wanted to see meaning in experimental work – and this is to be expected from an experimental psychologist, then the possibility of generalization has to be the default mindset when interpreting research findings. Generalization, at least across time and space, but ideally also regarding any other variable on which the sample at hand comes with a specific value.

Now, if we were to accept the (preliminary) conclusion that rules become ingrained into the human cognitive system, does this conclusion speak against a direct and tempting influence of motivationally salient rewards as it is highlighted in economic theorizing? No, it does not. It seems eminently plausible that temptations may at times compete with tendencies to follow a rule or norm, and such temptations may readily bias participants to opt for violating a rule as has been shown in several experiments (e.g., Hilbig & Thielmann, 2017). Whether or not the presence of a strong motivational temptation will ever be sufficient to remove rather than counteract the hypothesized automatic retrieval of rule-based tendencies in their entirety is a different and currently open question, though. Tackling this question with suitable experimental designs will certainly inform our understanding of how rules are represented and how directly and immediately a rule is retrieved when encountering rule-related stimuli in the environment (see Pfister et al., 2019, for a first attempt in this direction).

Other potential moderators of retrieving rule-based response tendencies are also likely to exist and have been highlighted by research on conformity and by research on human factors alike. They include situational variables such as social support, e.g., by observing disobedient acts of a third party; interpersonal variables such as the individual propensity for risk-taking; as well as cultural variables, e.g., those that vary along the individualistic-collectivistic continuum (Chen et al., 2006; Elms & Milgram, 1966; Reason, 1990). Rule-violation behavior will further be shaped by prior acts of rule violation of one and the same agent (Jusyte et al., 2017; Ting, 2020; Wirth et al., 2018), and the cognitive burdens of enacting a rule violation are also likely to differ between different types of rule violations. Such types comprise hidden rule-breaking as in the case of cheating, open acts of rule-breaking as in the case of rebellious behavior, as well as norm-breaking acts of creativity and innovation (Gozli, 2019). As for the case of temptations, these potential moderators might either affect the strength with which a rule is retrieved upon encountering rule-related stimuli, or they might even prevent the rule of being retrieved entirely. Determining which of these two possibilities is the case for different potential moderators again requires empirical efforts. These empirical efforts of course call for more refined, more complex, and potentially also more externally valid experimental designs. They also call for different variants of operationalizing rule-violation behavior as only convergent operationalization allows drawing meaningful conclusions (Garner et al., 1956; Grace, 2001). One potential avenue is to compare different kinds of rules such as specific stimulus-response pairings of the type "Upon encountering stimulus X, perform action Y" for a fixed set of potential stimuli versus classification rules of the type "Upon encountering a stimulus with the feature X, perform action Y" that may be applied to an indefinite number of different stimuli or situations. Another potential avenue for arriving at convergent operationalization is to broaden the applied measures, with promising candidates being gaze behavior to study attentional allocation during rule violation as well as physiological responses such as skin conductance responses to evaluate potential affective implications. Undertaking such efforts will allow to judge the simple model of automatic rule retrieval, and corresponding results will likely call for modifications and theoretical refinements. Eventually, they will also allow to assess whether the topic at hand did indeed allow for an approach that tries to isolate "building blocks" and whether our initial operationalization did indeed capture what we had intended it to do (Gozli & Deng, 2018).

The use of different ways to operationalize the construct under investigation is especially relevant in the search of exceptions to our claim of immediate and necessary rule-retrieval during rule-violation behavior, because such work ultimately aims at providing evidence for the absence of an effect rather than documenting the absence of evidence. As mentioned above, a consistent bias toward choosing rule violation for certain individuals in certain situations cannot be seen as sufficient evidence as it omits the process of how an eventual decision came about. Suppose a child in the swimming pool, boasting in front of its peers that it has no issue whatsoever jumping from the highest diving platform. Observing this child to consistently muster the courage to jump from the platform on several occasions will only tell part of the story. If its peers wanted to assess whether the child's boasting was fully justified, they might be inclined to take a closer look at signs of hesitation when standing on top of the diving platform. Observing the child to step back from the platform several times before anxiously leaping over the edge will paint a different picture than observing the child standing on the platform relaxed and calm. Similarly, research on rule-violation behavior will at the very least have to ascertain that measures of action planning (e.g., response times), measures of action execution (e.g., movement trajectories), and additional proxies of cognitive and affective processing (eye movements, physiological arousal) jointly suggest an absence of rule retrieval upon encountering a rule-related stimulus to substantiate a case against the simple model advocated above.

While engaging in such research will offer many new insights in the representation of rules and its impact on human behavior, I believe that such efforts will have to be carried out in the spirit of narrowing down an ambitious and overly general theoretical idea, either by arriving at a more elaborate re-formulation of the theory or by pinpointing the situations to which the general theory applies sufficiently closely.

Generalization

> Science may be described as the art of systematic over-simplification. (Popper, 1988, p. 44)

Philosophy of science has brought a variety of accounts that aim at characterizing, justifying, and norming the dialogue between theoretical ideas and empirical findings. The view that I promoted with reference to research on rule representations might be seen as a clumsy import of structuralist ideas into Popper's logic of exposing a theory to attempts at falsifying its predictions (Balzer et al., 1987; Popper, 1935/2005, 1988; Westermann, 1987, 2017). Even though other modes of theoretical development are equally possible from a philosophical point of view, I believe that the agenda of embarking from bold theoretical claims in order to narrow down a theory or model is not only desirable in experimental psychology, but I believe that it is also without practical alternatives in this field as I will outline below (readily dismissing the Popperian insight that there is always a practical alternative when actually understanding a theory or problem; Popper, 1972). My commitment to such an approach in the context of experimental psychology does not intend to deny that more synthetic approaches for constructing theories will likely be viable on other occasions. Such synthetic approaches will further be able to inform experimental approaches (Valsiner, 2017). Still, I believe that experimentation will always require a sufficiently general theoretical idea as its bedrock. This even applies to experimental work that may be carried out with a number of other proximal goals in mind. A researcher might be inclined to address methodological shortcomings of previous work without necessarily subscribing to its theoretical tenets, or a lecturer might replicate an existing experiment with their students in order to demonstrate the intricacies of experimentation. However, such work will arise only if theory-driven work had been conducted in the first place.

It is a particular feature of experimental psychology – in fact, of any field that employs experimental methodology – that researchers carefully create the conditions they intend to investigate in order to elicit precisely the type of behavior they intend to study. This type of work can only be stimulated by a sufficiently general theory on the subject under investigation. Constructing a corresponding theory, however, requires a vision. It requires a vision of what to explain and how to explain it. It requires a vision of how to transcend empirical facts that have already been established. And it requires a vision to explain cognition and behavior along a sufficiently large range of situations.

A vision itself will only provide a broad theoretical outline, of course, ideally supplemented by an informed intuition on how to implement a first experimental take on the subject matter. Once this outline is in place, however, it allows for critical modifications based on new evidence. This process of modifying a general theory based on accumulating evidence can be compared to the process of sculpting as portrayed in Fig. 4.1 (Schenk, personal communication): The raw workpiece will typically define the scope of the final product regarding its outline (say, the shape of the rock that the sculptor starts to work on) and its makeup (say, the type of rock that is used).[4] Only continued labor will carve out a sufficiently pleasing result eventually. In the context of experimental psychology, the workpiece is a relatively general theory on any aspect of human conduct, while carving out takes place via continued empirical efforts that aim at critically testing the original theoretical idea. As in sculpting, different states of a theoretical idea call for different empirical tools. While coarse theoretical ideas will yield simple experiments, increasing theoretical refinements call for increasingly sophisticated and complex experimentation that further adapts to challenges along the way (Ting & Fitzgerald, 2020). The final product, be it a sculpture or a well-developed theoretical account, requires both types of work: Work that defines the broad outline with sufficiently coarse tools and

[4] I borrow the sculpting metaphor from an inspiring presentation by Thomas Schenk (personal communication) in the summer term 2014 on the two visual stream models. Like many other influential theories and models in psychology, this field of research started with an overly general claim on the neurophysiological and functional separation of two pathways for processing visual information (Milner & Goodale, 2006). Follow-up work then continued to show exceptions from this simple model, thus carving an elaborate understanding of visual processing in the context of conscious perception and action control alike (Schenk et al., 2011).

Fig. 4.1 The sculpting metaphor of theorizing in experimental psychology. Initial theories that are formulated when approaching a new subject may be overly general and miss many of the nuances that would be required for a satisfying description. Still, the original workpiece sets the scope of what the final product will eventually be like (left panel). Follow-up work that aims at replicating original findings with convergent operationalizations, work that aims at identifying limits and moderators, and work that aims at challenging the basic tenets of a theory will carve out a more precise and well-rounded theoretical approach (center). Focusing only on isolated components or specific details, however, will torpedo the whole enterprise (right)

work that carves out details to arrive at a desirable result. With regard to theorizing, this latter work takes place in the form of arguments that challenge, evaluate, and refine certain ideas.

The iterative process of narrowing down a general theoretical idea (or workpiece) also seems to come with several direct connections to scientific critique as advocated by Gozli (2019). Critique can curate theoretical approaches and corresponding findings to evaluate whether empirical (especially, experimental) researchers managed to find the right scope both with regard to theoretical progression and with regard to experimental manipulations. Critique can uncover connections to fields that might share goals or methodology, and it can also uncover methodological pitfalls such as experimental specificities that might affect the generalizability of research findings. Finally, critique can pressure experimental researcher not to lose touch with the real world (or at least aim to regain touch if running danger to miss the bigger picture). Crucially, however, these valuable contributions of scientific critique will shine most when researchers aim to formulate maximally general theories on a cognitive process of interest and clearly communicate how they derive an operationalization from their theoretical beliefs. The exercise of conducting empirical research is conditional on the belief that findings might generalize at least to some degree, and only with this mindset can we build an informative and useful model of the inner workings of the human mind.

References

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men* (pp. 177–190). Carnegie Press.

Balzer, W., Moulines, C. U., & Sneed, J. D. (1987). *An architectonic for science: The structuralist program*. Reidel.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy, 76*(2), 169–217. https://doi.org/10.1007/978-1-349-62853-7_2
Bereby-Meyer, Y., & Shalvi, S. (2015). Deliberate honesty. *Current Opinion in Psychology, 6*, 195–198. https://doi.org/10.1016/j.copsyc.2015.09.004
Bermeitinger, C. (2012). *Paradigmen der Kognitiven Psychologie: Affektive Reize I [The paradigms of cognitive psychology: Affective stimuli I]*. Uni-Edition.
Billig, M. (2013). *Learn to write badly: How to succeed in the social sciences.* Cambridge, UK: Cambridge University Press.
Blass, T. (1999). The Milgram paradigm after 35 years: Some things we now know about obedience to authority. *Journal of Applied Social Psychology, 29*(5), 955–978. https://doi.org/10.1111/j.1559-1816.1999.tb00134.x
Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652. https://doi.org/10.1037/0033-295X.108.3.624
Chen, S., Hui, N., Bond, M., Sit, A., Wong, S., Chow, V., et al. (2006). Reexamining personal, social, and cultural influences on compliance behavior in the United States, Poland, and Hong Kong. *The Journal of Social Psychology, 146*, 223–244. https://doi.org/10.3200/SOCP.146.2.223-244
Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015
Dignath, D., Kiesel, A., & Eder, A. B. (2014). Flexible conflict management: Conflict avoidance and conflict adjustment in reactive cognitive control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(4), 975. https://doi.org/10.1037/xlm0000089
Dreisbach, G. (2012). Mechanisms of cognitive control: The functional role of task rules. *Current Directions in Psychological Science, 21*(4), 227–231.
Elms, A. C., & Milgram, S. (1966). Personality characteristics associated with obedience and defiance toward authoritative command. *Journal of Experimental Research in Personality, 1*, 282–289.
Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143–149. https://doi.org/10.3758/BF03203267
Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise—An experimental study on cheating. *Journal of the European Economic Association, 11*, 525–547. https://doi.org/10.1111/jeea.12014
Foerster, A., Pfister, R., Schmidts, C., Dignath, D., & Kunde, W. (2013). Honesty saves time (and justifications). *Frontiers in Psychology, 4*, 473. https://doi.org/10.3389/fpsyg.2013.00473
Frings, C., Schneider, K. K., & Fox, E. (2015). The negative priming paradigm: An update and implications for selective attention. *Psychonomic Bulletin & Review, 22*(6), 1577–1597. https://doi.org/10.3758/s13423-015-0841-4
Garner, W. R., Hake, H. W., & Eriksen, C. W. (1956). Operationism and the concept of perception. *Psychological Review, 63*(3), 149–159. https://doi.org/10.1037/h0042992
Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review, 95*(1), 384–394. https://doi.org/10.1257/0002828053828662
Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford University Press.
Gozli, D. G. (2017). Behaviour versus performance: The veiled commitment of experimental psychology. *Theory & Psychology, 27*, 741–758. https://doi.org/10.1177/0959354317728130
Gozli, D. G. (2019). *Experimental psychology and human agency*. Springer. https://doi.org/10.1007/978-3-030-20422-8
Gozli, D. G., & Deng, W. (2018). Building blocks of psychology: On remaking the unkept promises of early schools. *Integrative Psychological and Behavioral Science, 52*, 1–24. https://doi.org/10.1007/s12124-017-9405-7
Grace, R. C. (2001). On the failure of operationism. *Theory & Psychology, 11*(1), 5–33. https://doi.org/10.1177/0959354301111001
Hackman, J. R. (1969). Toward understanding the role of tasks in behavioral research. *Acta Psychologica, 31*, 97–128. https://doi.org/10.1016/0001-6918(69)90073-0

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hilbig, B. E., & Hessler, C. M. (2013). What lies beneath: How the distance between truth and lie drives dishonesty. *Journal of Experimental Social Psychology, 49*(2), 263–266. https://doi.org/10.1016/j.jesp.2012.11.010

Hilbig, B. E., & Thielmann, I. (2017). Does everyone have a price? On the role of payoff magnitude for ethical decision making. *Cognition, 163*, 15–25. https://doi.org/10.1016/j.cognition.2017.02.011

Hommel, B., Proctor, R. W., & Vu, K. P. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological Research, 68*(1), 1–17. https://doi.org/10.1007/s00426-003-0132-y

Imhof, M. F., & Rüsseler, J. (2019). Performance monitoring and correct response significance in conscientious individuals. *Frontiers in Human Neuroscience, 13*, 239. https://doi.org/10.3389/fnhum.2019.00239

Jusyte, A., Pfister, R., Mayer, S. V., Schwarz, K. A., Wirth, R., Kunde, W., & Schönenberg, M. (2017). Smooth criminal: Convicted rule-breakers show reduced cognitive conflict during deliberate rule violations. *Psychological Research, 81*(5), 939–946. https://doi.org/10.1007/s00426-016-0798-6

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching–a review. *Psychological Bulletin, 136*(5), 849–874. https://doi.org/10.1037/a0019842

Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition, 88*(2), 223–242. https://doi.org/10.1016/S0010-0277(03)00023-4

McKinstry, C., Dale, R., & Spivey, M. J. (2008). Action dynamics reveal parallel competition in decision making. *Psychological Science, 19*(1), 22–24. https://doi.org/10.1111/j.1467-9280.2008.02041.x

Meiran, N., Pereg, M., Kessler, Y., Cole, M. W., & Braver, T. S. (2014). The power of instructions: Proactive configuration of stimulus–response translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(3), 768. https://doi.org/10.1037/xlm0000063

Meiser, T. (2011). Much pain, little gain? Paradigm-specific models and methods in experimental psychology. *Perspectives on Psychological Science, 6*(2), 183–191. https://doi.org/10.1177/1745691611400241

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67*(4), 371–378. https://doi.org/10.1037/h0040525

Milgram, S. (1974). *Obedience to authority*. Harper & Row.

Milner, D., & Goodale, M. (2006). *The visual brain in action* (2nd Ed.). OUP.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *The Quarterly Journal of Experimental Psychology, 41*(1), 19–45. https://doi.org/10.1080/14640748908402351

Pfister, R., Wirth, R., Schwarz, K., Steinhauser, M., & Kunde, W. (2016). Burdens of non-conformity: Motor execution reveals cognitive conflict during deliberate rule violations. *Cognition, 147*, 93–99. https://doi.org/10.1016/j.cognition.2015.11.009

Pfister, R., Wirth, R., Weller, L., Foerster, A., & Schwarz, K. A. (2019). Taking shortcuts: Cognitive conflict during motivated rule-breaking. *Journal of Economic Psychology, 71*, 138–147. https://doi.org/10.1016/j.joep.2018.06.005

Popper, K. R. (1935/2005). *Logik der Forschung [The logic of scientific discovery]*. Mohr Siebeck.

Popper, K. R. (1972). *Objective knowledge: An evolutionary approach*. Clarendon.

Popper, K. R. (1988). *The open universe: An argument for indeterminism* (Vol. II). Routledge.

Pratt, T. C., & Cullen, F. T. (2000). The empirical status of Gottfredson and Hirschi's general theory of crime: A meta-analysis. *Criminology, 38*(3), 931–964. https://doi.org/10.1111/j.1745-9125.2000.tb00911.x

Reason, J. (1990). *Human error*. Cambridge University Press.

Schenk, T., Franz, V., & Bruno, N. (2011). Vision-for-perception and vision-for-action: Which model is compatible with the available psychophysical and neuropsychological data? *Vision Research, 51*(8), 812–818. https://doi.org/10.1016/j.visres.2011.02.003

Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science, 23*, 1264–1270. https://doi.org/10.1177/0956797612443835

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus–response compatibility. An integrated perspective* (pp. 31–86). Elsevier. https://doi.org/10.1016/S0166-4115(08)61218-2

Smedslund, J. (1997). *The structure of psychological common sense.* Mahwah, NJ: Lawrence Erlbaum.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662. https://doi.org/10.1037/h0054651

Teo, T. (2018). *Outline of theoretical psychology: Critical investigations.* London, UK: Palgrave Macmillan.

Ting, C. (2020). The feedback loop of rule-breaking: Experimental evidence. *The Social Science Journal*, 1–14. https://doi.org/10.1016/j.soscij.2018.11.004

Ting, C. (this volume). *The role of social context in experimental studies on dishonesty*.

Ting, C., & Fitzgerald, R. (2020). The work to make an experiment work. *International Journal of Social Research Methodology, 23*(3), 329–345. https://doi.org/10.1080/13645579.2019.1694621

Valsiner, J. (2017). *From methodology to methods in human psychology*. Springer.

Waszak, F., Pfister, R., & Kiesel, A. (2013). Top-down vs. bottom-up: When instructions overcome automatic retrieval. *Psychological Research, 77*(5), 611–617. https://doi.org/10.1007/s00426-012-0459-3

Wenke, D., Gaschler, R., Nattkemper, D., & Frensch, P. A. (2009). Strategic influences on implementing instructions for future actions. *Psychological Research, 73*(4), 587–601. https://doi.org/10.1007/s00426-009-0239-x

Westermann, R. (1987). Wissenschaftstheoretische Grundlagen der experimentellen Psychologie [Epistemological foundations of experimental psychology]. In G. Lüer (Ed.), *Allgemeine Experimentelle Psychologie* (pp. 4–42). Fischer.

Westermann, R. (2017). *Methoden psychologischer Forschung und evaluation [Methods for psychological research and evaluation] (Chapter 7)*. Kohlhammer.

Wirth, R., Pfister, R., Foerster, A., Huestegge, L., & Kunde, W. (2016). Pushing the rules: Effects and aftereffects of deliberate rule violations. *Psychological Research, 80*(5), 838–852. https://doi.org/10.1007/s00426-015-0690-9

Wirth, R., Foerster, A., Herbort, O., Kunde, W., & Pfister, R. (2018). This is how to be a rule breaker. *Advances in Cognitive Psychology, 14*(1), 21–37. https://doi.org/10.5709/acp-0235-2

Chapter 5
The Role of Social Context in Experimental Studies on Dishonesty

Carol Ting

Introduction

Dishonesty is an emotionally charged word. Accusations of dishonesty are often considered as powerful weapons by the accusing and serious insults by the accused. Across cultures, people are imbued with values against dishonesty since childhood, so much so that most of us do not need people to call out our dishonest behavior—just being aware of our own dishonesty often brings significant discomfort: fear of getting caught, guilt, and shame. Despite emotional costs, people still engage in all kinds of dishonest behavior big and small because cheating or breaking rules can bring significant material and/or psychological payoffs. Many of the high-profile cases of dishonesty in organizations can be told as stories of individuals or groups of people succumbing to the temptation of money or power. Unfortunately, misconducts at this level not only often result in substantial financial harm to others (Dyck et al., 2013), but they also undermine rules and norms, thereby corrupting institutions and compromising their long-term effectiveness (Gächter & Schulz, 2016; Shalvi, 2016; Stapel, 2016).

The emotional charge and harm to others underpin the social nature of dishonesty: it makes no sense to talk about dishonesty in a society made of only one person—she could make all kinds of false statements, and no one is there to call out her false statements; no one will be hurt so there is no responsibility, guilt, or shame. Put differently, dishonesty matters when it can potentially cause social harm, and that is why people feel uneasy and want to conceal dishonest behaviors. Preserving the social nature of dishonesty in the lab, however, poses a significant challenge to experimentalists because examining dishonesty in the lab requires inducing this covert behavior while making it transparent—a very fine line to walk on.

C. Ting (✉)
Department of Communication, University of Macau, Taipa, Macau S.A.R., China
e-mail: tingyf@um.edu.mo

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9_5

Although researchers have come up with various experiment designs to tackle this challenge, the extent to which these designs preserve the social nature of dishonesty is an issue that has not been explored. This article argues that dominant experimental paradigms of dishonesty research are susceptible to experimental artifacts that distort the social nature of dishonesty. The next section samples recent literature on dishonesty and identifies major experimental paradigms of dishonesty research. Taking a closer look, the three sections that follow analyze the social contexts of these paradigms and discuss their methodological implications.

Major Experimental Paradigms of Dishonesty Research

This section covers studies on dishonesty, lying, and cheating since the experimental literature tends to treat them as synonyms (Gozli, 2019; Jacobsen & Fosgaard, 2017; Köbis et al., 2019). Based on the nature and procedures of the experimental tasks, we can divide methods for inducing and measuring dishonesty in the lab into four categories: performance-misreporting tasks, stochastic tasks, social tasks, and instructed intention tasks.

Performance Misreporting Tasks

Performance misreporting experiments involve effortful tasks with built-in incentives (usually performance-based monetary incentives) for participants to over-report their performance. A prime example of performance misreporting tasks is Mazar et al. (2008), where performance refers to the number of matrix problems participants solve in a given amount of time. Each of those 3×4 matrices contains 12 three-digit numbers (e.g., 1.23), two of which add up to exactly 10, and the participants' task is to find those two numbers for each matrix. At timeout, participants report their performances and are paid accordingly. There are a few variants to this method (Gino et al., 2009; Gino & Ariely, 2012). General knowledge questions are sometimes used in place of matrix questions, and the procedures for submitting answers and receiving payment vary slightly. The most important commonality is that, with performance misreporting tasks, the induced level of dishonesty depends on the verifiability of participants' reported performances, which varies with experimental conditions. For example, in Mazar et al. (2008), participants' answers are verified in the control condition so there is no chance for cheating. In contrast, in the treatment condition, participants indicate their performance on a separate answer sheet and are instructed to put away their original test sheet for recycling, creating an opportunity to cheat.

In experiments based on performance misreporting tasks, participants report their performances anonymously, and researchers infer group-level dishonesty by comparing the performances between treatment and control groups. On the other hand, researchers also have the option of tracking individual-level dishonesty by linking individual worksheets to reported performances. This can be done in various ways: adding small print ID numbers to worksheets (Gino & Ariely, 2012), marking them with invisible ink (Vincent et al., 2013), and using a hidden camera to record and match worksheets to participants (Yaniv et al., 2019).

Stochastic Tasks

In contrast to performance misreporting tasks, stochastic tasks do not involve effort and effort-based achievement. Instead, participants only need to perform simple actions that generate random outcomes such as rolling a die (Fischbacher & Föllmi-Heusi, 2013) or flipping a coin (Bucciol & Piovesan, 2011). While these random variables have predictable distributions, dishonesty is incentivized by paying participants more for certain outcomes (e.g., \$1 if the die-roll result is 1, \$2 for 2, and so on), thereby inducing deviation in reported outcomes from the natural distributions of the random variables. Of course, people will be concerned about whether their reported outcomes signal dishonesty, so anonymity is an important element in such tasks. To assure participants of their anonymity, stochastic tasks often include extra protection such as rolling the die in a cup to convince participants that misreporting could not be detected.

The main appeal of stochastic tasks lies in its simplicity—a die roll or a coin flip takes almost no time, and little administration effort is required. Although built-in anonymization precludes measuring honesty at the individual level, group-level dishonesty can be easily inferred by comparing the aggregate reported outcome to the expected outcome of the random variable, and no control group is required. The die-roll experiment is also particularly popular among experimental economists (Gächter & Schulz, 2016; Hilbig & Thielmann, 2017; Rosenbaum et al., 2014) because it does not require deception, which is banned from experimental economics (Jacobsen & Fosgaard, 2017).

Social Tasks

In their review, Jacobsen and Fosgaard (2017) define social tasks as "those that involve more than one person (not counting the experimenter), which means that either the pay-off to the individual depends on another person, or the task involves a social component that might influence behavior." Gneezy's sender-receiver game (2005) is the most prominent example in this category. In this game, participants are randomly paired and assigned the role of a sender or a receiver. Both parties are informed that the total monetary reward is split between the sender and receiver according to one of two distributions that either favor the sender or the receiver, but only the sender is informed of the two distributions and their labels (option A or option B), which are kept from the receiver. The sender's task is to send a message to the receiver indicating which option is favorable to her, and the receiver makes the final decision on the option to be implemented. Gneezy shows empirically that most people expect the receiver to trust the sender's message and choose the recommended option. This pattern gives the sender the opportunity to take advantage of her superior information by lying to the receiver, and dishonesty at the individual level can be measured by comparing the sender's recommendation to the actual monetary distributions.

Instructed Intention Tasks

Cognitive psychologists are interested in the "lie effect," the phenomenon that an untruthful answer usually comes with longer response time and lower accuracy (Gozli, 2019; Suchotzki et al., 2017; Verschuere et al., 2018). This effect is typically studied in the lab by instructing participants to lie, and the instructed intention task (Foerster et al., 2018) is illustrative of how this works. In experiments based on instructed intention tasks, experimenters first obtain a number of facts about participants' recent actions by either asking them to answer some factual questions (e.g., did you watch TV today?) or asking them to perform certain actions (e.g., sending an email). Following this fact-collection stage, participants are then instructed to either "tell the truth" or "lie" about those facts with yes/no answers. (That is, when instructed to "lie," the participants have to change the answer from what they provided the first time or deny having performed the instructed action.) Participants' error rate and response time, while participants enact these instructions, are taken to reflect the cognitive efforts required by the task. As such, increased error rates and response time for "lie" responses are interpreted as the extra cognitive effort required in lying.

Following the brief descriptions above, the next three sections discuss some observations about the social contexts of these experimental paradigms. The focus is on how dishonesty is operationalized in these studies, and we ask if from the participants' viewpoint what they see is really similar to what we mean by dishonesty in everyday language.

Dishonesty with and Without Deception

One issue emerging from the previous overview of major experimental paradigms is that dishonesty seems to mean different things in different studies. In this light, it is also interesting to note that published experimental papers rarely define dishonesty. A possible explanation for this is that dishonesty, despite its prevalence and significance in society, is difficult to define in a way that can withstand serious logical inquisition (Lackey, 2013; Mahon, 2008; Meibauer, 2018).

Cognizant of this challenge, here we steer away from philosophical debates and begin with Coleman and Kay's prototypical lie (1981), where a person knowingly makes a false statement with the intention to deceive an addressee. Although Coleman and Kay focus on gradients along three dimensions (falsehood, the speaker's belief about the falsehood, and intention to deceive), for our purpose of comparing and contrasting experimental paradigms, we combine the first two conditions into one criterion "knowingly making a false statement" because no experimental studies on dishonesty are interested in people making a false statement without knowing it to be false. This results in two criteria: knowingly making a false statement and intentional deception. While all experimental paradigms examined in the previous section meet the first criterion, the instructed intention task does not meet the intentional deception criterion.

As the name "instructed intention task" suggests, in such experiments, participants do knowingly make a false statement, but they do so on the instruction of the experimenter rather than their own volition. In instructed intention experiments, there is no stake for the participants; they "lie" not for making gains or avoiding losses—unless we count participants' eagerness to impress the experimenter with their ability to follow instructions. Although this is a social factor, pleasing an experimenter is a completely different behavior from misrepresenting the state of affairs or hiding an undesirable fact.

Another issue is that many of these questions that can arise from such research methods (e.g., probability of choosing "lie" across groups or the time it takes to press a "yes" response untruthfully) are quite trivial, such as "did you go down a staircase?" The problem is that such activities may be too trivial to register and people may not be able to correctly recall the answer when suddenly asked about it. Perceiving the question as trivial, some may just give a rushed answer and repeat the same answer when asked about it the second time around. When this is the case, the experimental task is reduced to a recall test of whether one answered yes or no the first time. Even if participants have no trouble recalling the correct answer, this kind of task is essentially a purely logical task of repeating or flipping a fact/statement, and while the results can inform us about the dominance of accessibility of the factual statement, it says nothing about the common-sense notion of lying—which invokes morality and has social consequences (Gozli, 2019).

Last, as instructed intention tasks usually measure dishonesty with responses to close-ended questions as opposed to open-ended questions, they often lack the richness of questions arising with respect to dishonest behaviors in everyday life. That is, the truth/lie dichotomy is quite clear-cut in experimental designs, overlooking the ambiguities of everyday scenarios and the option of making up alternative facts when responding to open-ended questions.

Simulating Dishonesty in a Lab

As dishonesty is by nature a covert behavior, it is hard to observe it without interfering with it. Attempts to achieve this dual goal in the lab can easily threaten the external validity of the experiment—measurement in the lab is usually achieved through participants' cooperation with experimental procedures, but just as in the case of instructed intention tasks, asking participants to behave dishonestly creates a social context totally different from what we ordinarily mean by dishonesty. Through this lens, the biggest difference among major experimental paradigms appears to be how they address the difficult balance between inducing dishonesty and making the dishonest behavior observable. While performance misreporting tasks, stochastic tasks, and social tasks all induce dishonesty with monetary incentives, they provide participants with different levels of protection for their individual identity, which plays an important role in inducing dishonesty. Guaranteed anonymity is the standard practice for protecting individual identity in the lab, but it may not be as straightforward as it seems at first glance.

Let us start with the widely used matrix problem task. Recall that this design enables detection of dishonesty at the group level by comparing the average reported numbers of matrix problems solved in two conditions. Unlike the control condition where cheating is made impossible by verification, the recycle condition allows for cheating as explained in Mazar et al. (2008):

> … at the end of the four-minute matrix task, participants indicated the total number of correctly solved matrices on the answer sheet and then tore out the original test sheet from the booklet and placed it in their belongings (to recycle later), thus providing them with an opportunity to cheat (p. 636).

Since the original instructions are not published with the paper, we have to rely on this short description to reconstruct what the instructions for the recycle group may look like. This can introduce uncertainties, but one thing for sure is that for this treatment to be effective, people in the recycle group have to know *in advance* the exact answer submission and payment procedures. Specifically, before submitting their answers, participants must be informed that they will report the number of correct answers on an answer sheet separate from the test sheet and that they need to tear the original test sheet from the booklet and "recycle it." This can easily raise suspicion, as normally for a math task like this there is no need to bother with a separate answer sheet, not to mention tearing the original test sheet from the booklet for recycling. Although from the short description it is hard to know exactly how the recycling works and how elaborate the original instructions were, the great lengths the experimenter went to ensure anonymity are likely to accentuate the unnatural lab context and the social contract between the experimenter and the subjects (Böhme, 2016; Gozli, 2017). Adding the widely known fact that deception is commonly used in social psychology experiments, it is not implausible that some participants would start wondering or even figure out the purpose of the procedures.

With this design, Mazar et al. (2008) find that people do cheat—but only moderately. They explain this pattern with their proposed "self-concept maintenance theory," which states that "people behave dishonestly enough to profit but honestly enough to delude themselves of their own integrity. A little bit of dishonesty gives a taste of profit without spoiling a positive self-view." While this explanation is very plausible, we could also turn our attention to the experimental context and ask: if people are sensitive enough to calculate the right balance between self-concept and profit, why would they not be sensitive enough to notice that they are lab subjects under study in an unnatural situation? Would the fact that they are in a lab not make them more self-conscious and accentuate their self-concept? These possibilities are also consistent with their findings.

This kind of awkward questions is typical of experiments based on performance misreporting tasks. For example, in Gino and Ariely (2012, p. 449), where the matrix task is followed by a knowledge quiz that also offers monetary incentives for performance, participants indicate their answers according to the following description:

> The experimenter told them to circle their answers on their question sheet and explained that they would transfer their answers to a bubble sheet after finishing. When participants finished the quiz, the experimenter told them that, by mistake, she had photocopied bubble sheets that already had the correct answers lightly marked on them. She then asked the participants to use these pre-marked bubble sheets, recycle the test sheets with their original responses, and submit the bubble sheets for payment.

For a participant, this probably seems quite peculiar after the matrix task. First an experimenter brings the wrong bubble sheets and tells participants that those pre-marked answers are the correct answers, and then they are asked to recycle the test sheets with their original responses. Peculiarity aside, from the point of view of participants, the experiment probably seems to be quite poorly executed for a study measuring their general knowledge. This, of course, can help people rationalize dishonest behavior, but we then need to take this context into consideration when interpreting the results.

As discussed in the previous section, stochastic tasks based on random events draw on natural probabilistic distribution of random events for detection of dishonesty. With a die and a cup, measurement can be done in a matter of minutes. Simplicity is the method's beauty but at the same time also its potential weakness. The physical and transient nature of the task and its separability from the reporting mechanism help participants infer that reported outcomes cannot easily be verified. However, exactly because the task is so simple and making money out of it is so easy that its purpose is rather suspicious, as admitted in the original paper (Fischbacher & Föllmi-Heusi, 2013, p. 529):

> In order to make the experiment as plausible as possible, we told the subjects that the reason for rolling the die was to determine the payoff for filling in a questionnaire. It is clearly not very plausible to pay subjects differently for doing exactly the same task. Still, it is more plausible to let them roll the die in order to determine a payoff for doing something instead of just letting them roll the die and paying them without any explanation.

To what extent this cover story prevents potential suspicion is uncertain. The purpose of this simple experiment could seem transparent to some participants.[1] For participants who have guessed at the moral focus of the experiment, the experiment becomes more than just about dishonesty, and it is also about what a subject ought to do in an experiment. If a participant's only goal in participating

[1] Ting and Fitzgerald (2020, p. 336) provides such an example.

in the experiment is to make as much money as possible, her behavior will reflect only dishonesty. However, there could easily be other superordinate goals that allow for alternative subordinate goals and ambiguous interpretations (Gozli, 2017, 2019). A participant who thinks the researchers want to see dishonesty may reason that by cheating she not only makes money for herself but also helps the researchers. In this case, her cheating reflects both dishonesty and her desire to cooperate. Contrarily, she may find the experimenter manipulative and, therefore, act defiantly, seeing herself as the agent who brings fairness to the situation. In a large participant pool, there is no guarantee that these different factors cancel out each other, and the observed aggregate behavior is likely a mix of all these factors.

Fischbacher and Foellmi-Heusi also address the issue of anonymity. They made it as obvious as possible to participants that there is no way the experimenters could learn about what number a participant actually roll. This is done by encouraging the participants to roll the die as many times as they want (on the pretext of testing if the die is loaded) so that the first roll (which is the one that counted) can be erased completely. They also acknowledge that although actual outcomes cannot be observed, the reported outcomes may serve as a signal of potential dishonesty. Since participants expect dishonest people to inflate their outcomes (p. 541), reporting large numbers (4 and 5) can make one self-conscious. To address this issue, the authors implemented a double-anonymous condition where reported outcomes cannot be traced back to individual participants. Their paper so meticulously addresses the anonymity issue to the point that they even considered the possibility that participants might be concerned that the payoff they claimed could be inferred by the sound of taking coins out from an envelope. In addressing these concerns, they design a set of elaborate procedures that participants go through: (1) from a box presented by the experimenter, they take an unmarked envelope (containing another unmarked envelope and the maximum payoff, which was five coins in their study), (2) take out from the first envelope the second envelope and the coins they reportedly earned (leaving the remaining coins in the second envelope), (3) seal the envelope with the remaining coins in it, and (4) anonymously deposit it in a box by the door. Conceivably these elaborate procedures probably take longer than the actual die-rolling task, which, while making it "as obvious as possible that we had no chance to trace back any decisions on the individual level" (p. 531), can also make the moral focus of the experiment front and center. This would not be a problem if participants' only consideration is money, but in the presence of other motives, it could threaten the external validity of the experiment.

Of course, heightened self-concept and increased self-consciousness can also arise in many difficult real-life decisions where one is trading off honesty with personal gains. Yet if it is the elaborate anonymization procedures that bring the lab context to the forefront and make people self-conscious, what the experiments show probably does not reflect situations where people are desensitized to dishonesty and see dishonesty as part of the business, which characterize many of the commonly cited high-profile dishonesty cases (Bazerman & Tenbrunsel, 2012; Sezer et al., 2015).

Harm and Victim Identity

In the performance misreporting and stochastic tasks, dishonesty harms an abstract victim (the experimental budget). In contrast, social tasks feature very concrete victims (usually another participant). The effect of this aspect of victim identity is examined in a recent meta-analysis (Köbis et al., 2019), which finds that if the victim is an abstract entity people's more intuitive response tilts toward dishonesty (but this effect disappears with a concrete victim).

Two points can be made about the role of victim identity in dishonesty experiments. First, it is not easy to simulate naturally occurring dishonesty in an environment closely associated with the notion of lab animals—simple dishonesty experiments tend to involve a tradeoff between anonymity and simulating social harm in the lab. As previously discussed, anonymity is elaborately highlighted in performance misreporting and stochastic tasks, which accentuates the lab context and the abstractness of any perceived victims (one could reason that budget must not be an issue since a researcher concerned with budget and data quality would probably be more careful about potential cheating). On the other hand, by pairing participants into potential cheater-victim pairs, social tasks foreground the victim, which arguably increases the moral stake and realism. This comes with a cost, though, as it is likely that, with another person's interest at stake, people become more alert about the judgment of the experimenter and the fact that they are being watched. Balancing these two concerns is not an easy task. Can we simulate social harm under the condition of anonymity? It is possible to imagine increasing the group size in social tasks and thus obfuscating the identity of players, but the fact that everyone still has to enter a choice probably does not go far enough to alleviate this concern. Eventually, to more realistically simulate dishonesty as a social construct in the lab, deception and further obfuscation may be necessary.

Second, for real-life dishonesty of consequences, the victim is often neither as specific as another peer nor as vague as the experimental budget. An athlete deciding whether to dope is weighing between personal gains against a number of things: potential punishment if she gets caught, the meaning of a victory won by doping, the community of people who love the sport, competitors who choose to play by rules, competitors who choose not to play by rules, etc. A bank employee deciding whether to sell a fraudulent financial product to customers not only sees the would-be victims but also at stake are important values such as "just doing the job right" and "public interest" (Heumann et al., 2013). These system- or community-wide values involve another layer of formal/explicit rules beyond the basic principle of honesty, and the associated cost of violations probably cannot be summarily captured by an experimental budget.

Hierarchy of Rules and Norms

As social beings, humans constantly operate in multiple social spaces and hierarchies that impose various rules and norms on members. On the one hand, these rules and norms help maintain order by coordinating people with diverging goals; on the other, rules and norms also function as resources people rely on to navigate through those spaces and hierarchies, and in this process, the way people draw on rules/norms reshapes rules/norms and their roles. Take jaywalking as an example. School children are taught to walk through pedestrian crosses only on the green light, and this rule is supposed to be followed regardless of how much traffic there is. However, as people socialize into bigger social contexts, we pick up social cues from others which often run counter to the rules on the book. While the green light means "go" for people all around the world, it is not hard to find places (usually major intersections in crowded cities) where jaywalking is the norm. Newcomers and children often take up the norm quickly and their perception of the rule of crossing changes in the process, and this reinforces the jaywalking norm in those places. Contrarily, there are also places where the rule of crossing on only green light has far more force and jaywalkers stand out as rule-breakers. In short, rules/norms and group members' behavior mutually shape each other (Cicourel, 1974; Garfinkel & Sacks, 1970). We do not usually pay much attention to such dynamic adjustment to and reshaping of rules/norms because there are so many of them, and we rely on "autopilot" for basic rules to economize our cognitive resources. These basic rules become salient only when things do not work as what we are used to and/or take for granted, such as when we make a faux pas at a dinner with foreign guests.

Paying attention to all the ways people can violate all the implicit rules and social norms but do not, we can appreciate the incredible achievement of children picking up linguistic cues and learning to speak and that of people using language and nonverbal cues to figure out how to follow rules and pass as a member of a group. As people go through this learning process, decisions that originally took cognitive effort, such as turning away and covering up one's mouth and nose when sneezing, become automatic and slowly recede into the background. For most of us, it would feel very unnatural and nerve-wrecking if we were asked to sneeze right into someone's face! Rules like this (that we automatically follow) can only be unlearned. With repetition, people can learn to suppress learned automatic responses with new behaviors. This example not only illustrates the social nature of rules but also brings out the dynamic aspect of rule-following.

Most people belong to multiple social categories and are ingrained with the rules and norms pertinent to those categories, which are not always compatible. In almost every decision-making situation, though, some categories and the associated rules and norms take precedence over others. Which rules/norms dominate reflects people's overall priorities regarding their roles and goals, and these have been called the superordinate goals (Gozli, 2019). For example, imagine an experiment where participants are instructed to inhale pepper and then sneeze into the experimenter's face. In this case, subjects face two goals that are likely to be at odds with each other—to be a "good" person versus a "good" participant in the experiment. Even required by the experimenter, most subjects probably will not be able to follow this instruction successfully on (at least) the first few tries. Why is this instruction difficult to follow? If you are reading this, chances are you were taught that coughing and sneezing spread germs and diseases and that covering up and turning away when sneezing are not only part of hygiene but also a gesture of concern for those around us, "good manners" based on which people are judged. For some, probably many, the idea of sneezing into another's face may be so disgusting and so damaging to one's self-image that it is easier to drop out of the experiment instead.

All things considered, despite participants' desire to be cooperative "good subjects," their ability to follow those instructions will inevitably be compromised by ingrained social norms and their desire to be "good" people who are considerate, thoughtful, and well-mannered. In other words, observed subject behaviors reflect combined effects of the experimenter's instructions, and the social norm subjects bring with them into the lab. This is also true for other experiments, especially when the instructed task goes against the wider social norms relevant to the context (Mazar et al., 2008, p. 640). If social norms can be seen as instructions people pick up and internalize, then the experiment reviewed above essentially pitch one set of new instruction against other sets of instructions that subjects have internalized long before coming into the lab. In this case, the internalized norms will prevent experimenters from observing people's default response to the newly introduced rule. Although it is definitely possible to design a context- and value-free rule as an experimental task and observe people's default response to it, whether findings based on such socially neutral rules can readily apply to situations with real social and personal stakes is a question that must be answered first.

As implied in the jaywalking example, social norms also have a dynamic aspect that is often neglected in experimental studies on dishonesty. As social constructs' rules do not enforce themselves, an effective rule requires an enforcement infrastructure on which a common understanding of the status of the rule is hinged. But here is the tricky issue that common understanding of the status of the rule is part of the enforcement infrastructure and they mutually mould each other. For example, if a group or an organization has a bad track record of enforcing rules, members will take new rules lightly, which makes them less effective. This in turn reinforces the bad track record, encourages dishonesty, and forms a feedback loop (Ting, 2020) and a culture/climate of rule-breaking. Such dynamics are important in groups and organizations, but they are rarely touched upon in the experimental literature on dishonesty which often cites organizational misconducts and illegal behaviors as opening hooks. The studies reviewed above do cover social factors, but they cover only those reflected in the values participants bring with them into the lab and pays insufficient attention to the dynamic social context according to which the participant views the situation. We cannot understand dishonesty as rule-breaking in organizations without considering the enforcement infrastructure as part and parcel of the puzzle.

Before closing, a practical point and emphasis is in order. As Pfister (this volume) points out, experimental psychology develops through identifying key concepts in incremental steps of theorizing and empiricizing—such incremental approach is integral to the study of subjects as complex as the human mind as there are always additional factors beyond our theoretical model and methods. However, the point here goes beyond the realization that there are many factors influencing dishonest and rule-breaking behavior. Rather, what I want to emphasize in this chapter is that the categories of dishonest and rule-breaking behavior are themselves varied and context-dependent; thus, they ought to be carefully distinguished. Put another way, experimental operationalizations of dishonesty and rule-breaking should account for the difference between the general normative expectations and what the experiment rule aims to simulate, which, while presenting significant methodological challenges, is key to expanding our knowledge of dishonesty as a social construct.

Conclusion

This article analyzes the social aspect of dominant paradigms of the experimental literature on dishonesty and discusses its methodological implications. While the significance of dishonesty in human society is closely linked to its social nature, under closer examination the range of social contexts covered by dominant experimental paradigms appears to be quite limited due to definitional issues and potential lab artifacts. Striving to induce dishonest behavior in the lab, making the behavior easy to identify and measure, researchers often lose sight of the social context of dishonesty and the fact that dishonest behavior is meaningful because of its place in a social context. This is unfortunate, because the social and context-dependent nature of dishonesty is, in the first place, what makes it an interesting and challenging research topic.

References

Bazerman, M., & Tenbrunsel, A. (2012). *Blind spots: Why we fail to do what's right and what to do about it*. Princeton University Press.

Böhme, J. (2016). 'Doing' laboratory experiments: An ethnomethodological study of the performative practice in behavioral economic research. In I. Boldyrev & E. Svetlova (Eds.), *Enacting dismal science: Perspectives from social economics* (pp. 87–108). Palgrave Macmillan.

Bucciol, A., & Piovesan, M. (2011). Luck or cheating? A field experiment on honesty with children. *Journal of Economic Psychology, 32*(1), 73–78. https://doi.org/10.1016/j.joep.2010.12.001

Cicourel, A. V. (1974). *Cognitive sociology: Language and meaning in social interaction*. The Free Press.

Coleman, L., & Kay, P. (1981). Prototype semantics: The English word lie. *Language, 57*(1), 26–44. https://doi.org/10.1353/lan.1981.0002

Dyck, I. J., Morse, A., & Zingales, L. (2013). How pervasive is corporate fraud?. *Rotman School of Management Working Paper*, (2222608).

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise—An experimental study on cheating. *Journal of the European Economic Association, 11*(3), 525–547. https://doi.org/10.1111/jeea.12014
Foerster, A., Wirth, R., Berghoefer, F. L., Kunde, W., & Pfister, R. (2018). Capacity limitations of dishonesty. *Journal of Experimental Psychology: General, 148*(6), 943–961. https://doi.org/10.1037/xge0000510
Gächter, S., & Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature, 531*, 496–499. https://doi.org/10.1038/nature17160
Garfinkel, H., & Sacks, H. (1970). On formal structures of practical actions. In J. C. McKinney & E. A. Tiryakian (Eds.), *Theoretical sociology: Perspectives and development*. Appleton-Century-Crofts.
Gino, F., & Ariely, D. (2012). The dark side of creativity: Original thinkers can be more dishonest. *Journal of Personality and Social Psychology, 102*(3), 445–459.
Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior the effect of one bad apple on the barrel. *Psychological Science, 20*(3), 393–398. https://doi.org/10.1111/j.1467-9280.2009.02306.x
Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review, 95*(1), 384–394. https://doi.org/10.1257/0002828053828662
Gozli, D. G. (2017). Behaviour versus performance: The veiled commitment of experimental psychology. *Theory & Psychology, 27*(6), 741–758. https://doi.org/10.1177/0959354317728130
Gozli, D. G. (2019). *Experimental psychology and human agency*. Springer. https://doi.org/10.1007/978-3-030-20422-8
Heumann, M., Friedes, A., Cassak, L., Wright, W., & Joshi, E. (2013). The world of whistleblowing. *Public Integrity, 16*(1), 25–52. https://doi.org/10.2753/pin1099-9922160102
Hilbig, B. E., & Thielmann, I. (2017). Does everyone have a price? On the role of payoff magnitude for ethical decision making. *Cognition, 163*, 15–25. https://doi.org/10.1016/j.cognition.2017.02.011
Jacobsen, C., & Fosgaard, T. (2017). Why do we lie? A practical guide to the dishonesty literature. *Journal of Economic Surveys, 32*(2), 357–387. https://doi.org/10.1111/joes.12204
Köbis, N. C., Verschuere, B., Bereby-Meyer, Y., Rand, D., & Shalvi, S. (2019). Intuitive honesty versus dishonesty: Meta-analytic evidence. *Perspectives on Psychological Science, 14*(5), 778–796. https://doi.org/10.1177/1745691619851778
Lackey, J. (2013). Lies and deception: An unhappy divorce. *Analysis, 73*(2), 236–248. https://doi.org/10.1093/analys/ant006
Mahon, J. E. (2008). The definition of lying and deception. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of philosophy*. https://plato.stanford.edu/archives/win2016/entries/lying-definition/
Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*(6), 633–644. https://doi.org/10.1509/jmkr.45.6.633
Meibauer, J. (2018). The linguistics of lying. *Annual Review of Linguistics, 4*(1), 357–375.
Pfister, R. (this volume). Operationalization and generalization in experimental psychology: A plea for bold claims. In D. Gozli & J. Valsiner (Eds.), *Experimental psychology: Ambitions and possibilities*. Springer.
Rosenbaum, S. M., Billinger, S., & Stieglitz, N. (2014). Let's be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology, 45*, 181–196.
Sezer, O., Gino, F., & Bazerman, M. H. (2015). Ethical blind spots: Explaining unintentional unethical behavior. *Current Opinion in Psychology, 6*, 77–81. https://doi.org/10.1016/j.copsyc.2015.03.030
Shalvi, S. (2016). Behavioural economics: Corruption corrupts. *Nature, 531*(7595), 456–457. https://doi.org/10.1038/nature17307
Stapel, D. (2016). *Faking science: A true story of academic fraud* (N. J. L. Brown, Trans.). https://errorstatistics.files.wordpress.com/2014/12/fakingscience-20141214.pdf

Suchotzki, K., Verschuere, B., Van Bockstaele, B., BenShakhar, G., & Crombez, G. (2017). Lying takes time: A meta-analysis on reaction time measures of deception. *Psychological Bulletin, 143*, 428–453. https://doi.org/10.1037/bul0000087

Ting, C. (2020). The feedback loop of rule-breaking: Experimental evidence. *The Social Science Journal.* https://doi.org/10.1016/j.soscij.2018.11.004

Ting, C., & Fitzgerald, R. (2020). The work to make an experiment work. *International Journal of Social Research Methodology, 23*(3), 329–345. https://doi.org/10.1080/13645579.2019.1694621

Verschuere, B., Köbis, N. C., Bereby-Meyer, Y., Rand, D., & Shalvi, S. (2018). Taxing the brain to uncover lying? Metaanalyzing the effect of imposing cognitive load on the reaction-time costs of lying. *Journal of Applied Memory & Cognition, 7*, 462–469.

Vincent, L., Kj, E., & Ja, G. (2013). Stretching the moral gray zone: Positive affect, moral disengagement, and dishonesty. *Psychological Science, 24*(4), 595–599. https://doi.org/10.1177/0956797612458806

Yaniv, G., Tobol, Y., & Siniver, E. (2019). Self-portrayed honesty and behavioral dishonesty. *Ethics & Behavior*, 1–11. https://doi.org/10.1080/10508422.2019.1678162

Chapter 6
What Is a Task and How Do You Know If You Have One or More?

Eliot Hazeltine, Tobin Dykstra, and Eric Schumacher

Introduction

Understanding how the brain uses incoming sensory information to activate motor systems to produce goal-based behavior is a fundamental question in psychology and neuroscience. Not only are the links between the events in the environment and the desired actions potentially arbitrary (i.e., any stimulus can signal that any response should be made), but they must change according to the current context and the needs of the individual. Moreover, the environment does not consist of a single stimulus but rather presents a constantly changing torrent of objects and events, each of which may lead to multiple candidate actions. How does our brain navigate this sea of drives and affordances to chart a desirable course?

To develop rigorous theories for how we perform coherent behaviors in complex environments, psychologists and neuroscientists have proposed a range of accounts with a common theme. The overarching idea is that stimulus-response (SR) associations are activated by the environment and control processes are engaged so that only one goal drives behavior at a time. That is, theories of voluntary behavior differ along multiple dimensions, including how control is implemented (see, e.g., Badre et al., 2021; Braver, 2012; Cookson et al., 2020; Duncan, 2013; Grant et al., 2020; Hazeltine et al., 2011a; Koch et al., 2018; Logan, 2002; Weissman et al., 2014) and how control processes are organized (see., e.g., Badre & D'Esposito, 2009; Badre & Nee, 2018; Courtney et al., 2007; Dosenbach et al., 2008; Fuster, 2008; Koechlin & Summerfield, 2007; MacDonald et al., 2000; Petrides, 2006; Sakai, 2008), but they share the notion that control processes govern which SR associations become most

E. Hazeltine (✉) · T. Dykstra
Department of Psychological and Brain Sciences, University of Iowa, Iowa City, IA, USA
e-mail: eliot-hazeltine@uiowa.edu

E. Schumacher
Department of Psychology, Georgia Institute of Technology, Atlanta, GA, USA

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9_6

active and ultimately drive behavior. The persistent reliance on SR associations may stem from the fact that the connection between stimulus codes, consisting of diverse representations of environmental events, and response codes, consisting of motor states, is mysterious, so the necessary computations to move from one to another is difficult to specify (Hommel et al., 2001; Prinz, 1990). The concept of an SR association provides a convenient shortcut for tackling this problem.

Psychological theories have applied this approach in a variety of ways. Early accounts proposed that a unitary central executive monitors the activation of SR associations and allows the most appropriate one to access motor structures (e.g., Norman & Shallice, 1986; Shallice, 1982). More contemporary theories have fractionated the central executive (e.g., Miyake et al., 2000; Monsell & Driver, 2000), but the basic division of labor has remained the same: A set of control processes enables some SR associations to win out over others and drive behavior. For example, the conflict adaptation model (Botvinick et al., 2001; see also Cohen et al., 1990) proposes that the coactivation of competing responses activates attentional systems that bias input, allowing task-relevant information to more strongly activate the relevant responses. Thus, according to this account, control is implemented through attentional processes modulating the activation of sensory representations (Desimone & Duncan, 1995). In this model, to diminish the activation of inappropriate SR associations, the corresponding stimuli are inhibited.

Some recent behavioral evidence suggests that the control processes mediating the effects explained by the conflict adaptation model involve more than just input attention (e.g., Grant et al., 2020; Hazeltine, et al., 2011b). In response, some modifications of the model have incorporated learning – that is, changes in the strengths of SR associations – rather than changes in the activation of stimulus representations to account for the dynamic control of behavior (e.g., Schmidt et al., 2016; Verguts & Notebaert, 2008). However, the basic idea remains that control processes external to the SR pathway modulate the activation of responses by stimuli.

Neuroscience accounts of cognitive control also largely rely on the notion of SR associations. Miller and Cohen's (2001) model of prefrontal cortex (PFC) function, for example, assumes the PFC essentially provides a set of intervening links between stimuli and responses, so contextual information can guide the activation from the stimulus to alternative responses that are more appropriate in a particular setting. While this approach adds intervening links between stimulus representations and response representations, it is consistent with the notion that behavior is driven by SR associations. Representations of context can bias which SR association is most active, but the contextual information, in conjunction with the stimulus information, still activates responses in a feedforward way. What has changed is that the input driving the selection of the response is now more complex, incorporating multiple aspects of the environment (e.g., context) or even information that is not present in the environment (e.g., the contents of working memory).

The idea that behavior is driven by SR associations is also at the heart of several popular theories of PFC function. These theories propose that the control of behavior is achieved through the coordinated activity of hierarchical modules in PFC that act on different levels of information that together determine which SR association guides behavior (for reviews, see Badre, 2008; Badre & D'Esposito, 2009; Badre & Nee, 2018). The exact nature of the hierarchy differs across theories, but they share the idea that associations between stimuli and responses are mediated by the most caudal regions of PFC. Which particular SR association drives behavior depends on goal-related or other contextual information mediated by modules in more rostral PFC regions. The more rostral in the hierarchy, the more abstract is the representation, and the less dependent it is on the current stimulus input.

In sum, the development of complex models of voluntary behavior has fractionated control processes so that a homuncular mechanism (i.e., a control process that has access to all relevant information and "decides" what to do) is no longer necessary for the selection of appropriate actions. However, in part to maintain computational tractability, these approaches continue to rely on the notion of the SR association as the basic unit driving goal-directed behavior.

Addressing the Limitations of SR Associations

Despite the widespread reliance on SR associations to explain how we behave, they do not provide an adequate framework to account for a range of voluntary behaviors (see Hazeltine & Schumacher, 2016). In fact, the SR association account does not even adequately explain phenomena from the behaviorist tradition from which the idea emerged. For example, Rescorla (1988a, b) argued that conditioning is better characterized as learning the relationship between the external environment and an animal's internal representation of that environment. Similarly, Tolman (1932) argued that animals create an internal representation of their world as they explore it. These, and other examples from behavioral psychology (c.f., Hazeltine & Schumacher, 2016), demonstrate that even non-human behavior involves more than simple SR associations. Indeed, an internal representation of the world (i.e., the animal's *task*) is fundamental.

In cognitive psychology, early evidence that complex mental representations guide behavior comes from Bartlett (1932) (and subsequently by Brewer & Treyens, 1981; Gozli, 2019; Tolman, 1948), who showed that the way we organize the relationship between learned information and our existing knowledge guides what we remember and how we remember it. For example, how witnesses represent an incident affects what and how they remember (Tuckey & Brewer, 2003). In addition to the effect of mental representations on memory, mental representations of our goals may also guide how we attend to and respond to the world.

In the area of cognitive psychology investigating human performance, the specific mental representations and processes required to perform a task are often called a *task set* (for reviews, see Monsell, 2003; Sakai, 2008). These representations are often explicitly hierarchal, combining different levels of our proximal and distal goals and additional information about the nature and organization of the task (e.g., Gozli, 2019; Schumacher & Hazeltine, 2016). The way we represent a task has behavioral consequences – from task-switching effects (discussed in more detail below) to guiding attention (Dreisbach, 2012) and to how and when we remain focused on an external task or allow our minds to wander (Bezdek et al., 2019; Murray et al., 2020).

To characterize the various approaches to understanding how complex representations can impinge on control processes and guide behavior, Badre et al. (2021) draw a distinction between modulatory accounts and transmissive accounts. Modulatory accounts assume that control processes monitor and adjust the activation of SR associations depending on independent representations of context and goal states (e.g., Botvinick et al., 2001; E. K. Miller & Cohen, 2001; Shallice, 1982). Thus, the SR associations are represented separately from the information that determines their appropriateness. Transmissive accounts, on the other hand, hold that SR associations are part of complex representations that include context and goal states (e.g., Duncan, 2013; Hazeltine, Lightman, et al., 2011; Hommel et al., 2001; Hommel et al., 2004; Kikumoto & Mayr, 2020; Schumacher & Hazeltine, 2016; see below). In this way, behavior is guided by the complex representations of actions rather than separate control processes that influence their activation.

We argue that, at least regarding voluntary behavior, the transmissive approach appears to have greater explanatory power because tasks are not simply collections of SR associations but are structured to include elaborate relationships that are not obviously related to the current context or goals. Consider, for example, when one reaches for a coffee mug to clean it rather than to fill it with coffee, which may be the more frequent behavior associated with the mug. In such cases, there are sometimes “action slips” (Norman, 1981) in which the presently undesired (and usually more frequent) action is performed with the object, suggesting that inhibitory processes must suppress actions that are inappropriate for the current circumstance. However, action slips typically occur when the individual is initiating an action (e.g., reaching for the mug), not in midstream (e.g., scrubbing the mug), suggesting that once the action is embedded in an ongoing task context, control is more stable. This is consistent with the proposal that the surrounding, related actions activate the current, appropriate response through the associations that have formed as part of the task representation.

A groundbreaking example of a transmissive account by Hommel and colleagues (Hommel, 2004; Hommel et al., 2001) proposed that voluntary actions are coded as *event files*. Event files are representations that bind stimulus features with the response features along with the current environmental context. The empirical foundation for this theory comes from studies showing that behavior is worse when only some task features overlap from one trial to the next (partial overlap) compared to both when all features overlap or non-overlap (Frings et al., 2020; Hommel, 1998). These results suggest that the context in which the stimulus and response appears is also encoded into the action representation (viz., the event file), even though they are extraneous to the actual SR applied to generate the action. In short, in event files, context is intrinsically bound with features of the stimuli and responses. This transmissive approach contrasts with modulatory accounts where contextual information activates control processes that activate or inhibit separate SR associations (as in, Botvinick et al., 2001; Miller & Cohen, 2001). Event file theory has generated a wealth of studies demonstrating that the production of a response leads to the encoding of multiple contextual factors, which, when repeated, may cause the retrieval of the same response, producing facilitation or conflict. However, the account does not directly address how this conflict is resolved and actions are ultimately selected.

Kikumoto and Mayr (2020) provided neuroscientific evidence for event files using EEG. They performed representational similarity analyses on the spectral profiles of EEG data to identify components associated with the stimulus, response, task set, and SR conjunction on a given trial in a task-switching procedure. The component associated specifically with the SR conjunction predicted intertrial variation in reaction time (RT). That is, the stronger this component, the faster the participants responded. This suggests that the strength of the event file or task set, which includes combined representations of the stimuli and responses, mediates performance. In a second experiment, they used tasks with overlapping SR rules. That is, some stimuli in the two tasks required the same response, and some stimuli required different responses. In this way they could distinguish between the effects of SR conjunctions with an integrated rule vs. rule-independent SR conjunctions. Consistent with their first experiment, they found that SR conjunction representations were more predictive than the stimulus and response representations alone. Additionally, the SR conjunction that was also integrated with the task set was more predictive than the task set-independent SR conjunction, suggesting that higher-order information plays a major role throughout action selection. Takacs et al. (2020) reported additional evidence for task set representations in EEG data where an identifiable cluster of activity for a task set remains after factoring out stimulus and response activity. Together these data support the idea that the combined representation of stimuli, responses, context, goals, etc. (i.e., the task set) is maintained by the brain and has consequences for goal-directed behavior.

One limitation with the event file theory is that the contextual information included in the event file is underspecified. Indeed, the original evidence for the theory focused on stimulus and response features – implicitly limiting context to environmental context, although recent formulations have been more inclusive (e.g., Frings et al., 2020). Schumacher and Hazeltine (2016) noted that the focus on stimulus and response information neglects the contribution of representations that include abstract relational information about actions. To emphasize that the organization of behavior is largely imposed by internal representations rather than the environment, Schumacher and Hazeltine proposed the task file hypothesis (see also Bezdek et al., 2019; Cookson et al., 2020; Grant et al., 2020; Hazeltine et al., 2011b; Smith et al., 2020). Like an event file, a task file is a mental representation that binds stimulus and response features with contextual information. However, task files explicitly include goals and motivations. Task files are also explicitly hierarchical, so that perceptual and response information is integrated as competition is resolved across a range of interactive levels (e.g., stimulus features, stimulus affordances, motor codes, intentions, etc.). Thus, resolving competition at a higher level in the hierarchy may alter the nature of the competition at lower levels (see Cookson et al., 2020; Hazeltine et al., 2011b; Smith et al., 2020).

As evidence for task files, Schumacher and Hazeltine point to findings in which the interactions between concurrent and consecutive actions do not appear to be driven by stimulus factors but rather the participants' conceptualizations of the task (e.g., Dreisbach, 2012; Halvorson & Hazeltine, 2015; Hazeltine, 2005; Schumacher & Schwarb, 2009). Note that, as the event file account allows for the integration of more diverse types of information (c.f., Frings et al., 2020), including intentions and goals, it converges with the task file framework. Nonetheless, theories of event files emphasize the role of the various sources of information as retrieval cues (Frings et al., 2020), and, by contrast, theories of task files would assume these sources operate at different levels of a control hierarchy.

Task Switching and Task Representation

The recognition that actions are embedded in larger representations that organize behavior in terms of information beyond what is available in the environment makes strong links with the extensive task-switching literature (c.f., Kiesel et al., 2010). Task switching investigates how the performance of one action affects the performance of an immediately subsequent one, and in many cases, the stimulus is ambiguous as to what action should be performed. In a typical procedure, it is assumed that the experimenter has a priori knowledge of the task structure. Participants are asked to make consecutive actions. Performance is compared when the actions belong to the same task to when the actions belong to different tasks. The general finding is that performance is worse (e.g., RT and error rates increase) when the actions belong to different tasks (i.e., a switch is required) compared to when they belong to the same task – even when no cues are used to generate expectations regarding the upcoming task (Jersild, 1927). In fact, performance costs associated with switching tasks are observed when the switch is determined solely by the participant (Arrington & Logan, 2004; Mittelstädt et al., 2018). Because switch costs are observed when neither the stimulus nor the response repeats (e.g., Mayr et al., 2006; Monsell, 2003; Rogers & Monsell, 1995; Ruthruff et al., 2001), these relationships do not depend solely on stimulus and response overlap. Thus, the standard interpretation of such findings is that SR associations are (somehow) grouped into task sets.

Neuroimaging experiments have largely accepted this interpretation and used task-switching procedures to identify brain regions and networks underlying the switch from one task to another. Meta-analyses of these studies have identified dorsolateral, ventrolateral, and medial PFC, as well as posterior parietal cortex as critical for task switching (Derrfuss et al., 2005; Kim et al., 2012). More recent studies, using pattern classification and functional connectivity analysis techniques, confirm the role that regions in the frontoparietal brain network play in task switching (e.g., Qiao et al., 2017; Qiao et al., 2020; Waskom et al., 2014; Wisniewski et al., 2015; Woolgar et al., 2011). The assumption is that frontoparietal regions encode the currently active SR associations. When a switch occurs, the network is reconfigured to represent the newly relevant set. For example, Qiao et al. (2017) used compound stimuli consisting of an overlapping face and building. Participants responded to the gender of the face or the size of the building on each trial. They found evidence that the representation of one of the tasks increased in frontoparietal cortex as the number of repeat trials increased, suggesting that this network maintains the current task set.

Limits of Task Switching

Yet, despite the wealth of studies examining how individuals switch from one task to another, the understanding of the underlying processes and how these processes are determined by the experimental procedures is weak. It is generally assumed that task sets are loaded into working memory together (e.g., Meiran et al., 2000; Rogers & Monsell, 1995), but this has received little experimental investigation (e.g., Logan, 2004). For example, if task sets are sets of SR associations that are necessarily loaded together in WM, their size should reflect capacity limits of working memory, which to the best of our knowledge has never been directly tested. That is, are tasks divided into separate sets when the number of possible stimuli or responses exceeds the capacity of working memory? If so, on what basis are these divided into sets? There is preliminary evidence that the number of responses, rather than the number of stimuli, provides the primary constraint on task set size (e.g., Wifall et al., 2015). The outsize weight that responses have compared to stimuli on task sets seems inconsistent with the simple idea that sets consist of SR associations. However, much of this experimental space has yet to be investigated. Moreover, we do not know how task boundaries are shaped by instructions, practice, and situational demands.

Not only are the operations associated with task switching not well-defined (see, e.g., Kiesel et al., 2010; Monsell, 2003; Wylie & Allport, 2000), they are likely not uniform across experiments. That is, the performance costs associated with switching tasks are often measured under conditions in which other operations might be affecting RT and accuracy. For example, in many task-switching experiments (e.g., Allport et al., 1994; Barcelo et al., 2006; Hayes et al., 1998; Poljac et al., 2009; Qiao et al., 2017), the switches involve reorienting from one stimulus dimension (e.g., color) to another (e.g., shape). This reorienting process may take time and may not be possible to complete before the onset of the stimulus indicating the response. There are experimental tasks that avoid this attention shift confound by using stimuli for which the relevant feature is identical across tasks. In these tasks (e.g., Kikumoto & Mayr, 2020; Logan & Schneider, 2006; Mayr & Bryck, 2005), a single relevant stimulus attribute is used (e.g., location), and the mappings between the attribute values and responses is changed on switch trials. While this approach eliminates shifts of attention as a possible source of the costs, it is by no means standard in the literature.

Moreover, because the same stimulus is associated with multiple responses, inhibition may be required to resolve the resulting response conflict. This is true in procedures using unidimensional stimuli (e.g., Logan & Schneider, 2006; Mayr & Bryck, 2005) as well as more procedures with multiple relevant stimulus dimensions (Kiesel et al., 2010; Mayr, 2002; Rogers & Monsell, 1995). Response conflict lengthens RT and increases error rates in many experimental paradigms, such as flanker (Eriksen & Eriksen, 1974), Stroop (Stroop, 1935), and Simon (Simon, 1969) tasks. These procedures are generally thought to engage cognitive control processes, but they are not thought to involve switching task sets. Thus, it can be argued that the inhibition of previously relevant mappings is part of the set of processes associated with task switching, but it is present in different degrees across the various procedures used to tap task-switching operations. In fact, when univalent stimuli (i.e., each stimulus associated with a single response) are used, switch costs are much reduced compared to when bivalent stimuli (i.e., each stimulus associated with multiple responses) are used (Rogers & Monsell, 1995), indicating that inhibitory process can play a sizeable role in the magnitude of switch costs. In sum, given the considerable differences in tasks, it is likely that the processes associated with task switching are not homogeneous across studies. In fact, researchers have exploited differences in the various procedures in efforts to isolate these putatively separate components associated with attention and inhibition (e.g., Gopher et al., 2000; Kim et al., 2012; Mayr et al., 2006; Rogers & Monsell, 1995), which raises the question of what components are essential to task switching.

Finally, our understanding of task switch costs is hindered by the fact that there is no independent definition of a task or task boundary (see Gozli, 2019). It is generally assumed that the experimenter controls or knows the task representation and then measures how crossing the boundary between tasks affects performance. The obvious limitation of this approach is that the task representation is not independently measured, so, while observed costs indicate that one set of transitional RTs (i.e., those that putatively cross the task boundary) are generally longer than the other set of transitional RTs (those that do not cross the boundary), there is limited evidence that the task structure consists of distinct sets of related SR associations. That is, because all the possible transitions are lumped into a small number of groups (usually two), the observation of a difference between these two groups is not highly diagnostic of a particular task structure. As the number of possible transitions increases, so does the risk that such differences are taken as support for a task structure that is quite different from what is actually supported by the data (see below). An unbiased approach would examine all possible transitions to provide a data-driven description of the task structure. In this way, one could verify that the task boundary identified by the observed switch cost was in fact the only (or even the dominant) boundary between different sets of responses. This approach is examined here.

Switching Costs May Not Always Reflect Switching Tasks

We conducted an experiment to compare task switch costs to the task structure as measured by a richer characterization of all the transitional RTs (see also Dykstra et al., in prep). To minimize the roles of inhibition and attention, we used univalent stimuli, presented one at a time in the center of the screen. Each response was mapped to a single stimulus, each of which was equally probable on every trial. For one group of participants, all the stimuli were numbers, whereas for the other group, the stimuli indicating left-sided responses were numbers, and the stimuli indicating right-sided response were faces. The question was whether switch costs would be observed under such conditions. Note that previous studies have observed switch costs with univalent stimuli although they are typically smaller than those observed with bivalent stimuli (e.g., Hirsch et al., 2016; Rogers & Monsell, 1995).

To increase the likelihood that some task structure would be imposed (i.e., the SR associations would be divided into task sets), we used a relatively large number of SR alternatives (8) and (in one condition) stimuli belonging to different categories (i.e., some were numbers and some were faces). As noted above, there is a reason to expect larger SR sets will be divided into smaller ones, even if the stimuli are univalent, to accommodate capacity limitations in working memory, although we are not aware of this being directly tested. Moreover, we are aware of no formal account that makes clear predictions about how this collection of SR alternatives will be divided into tasks (i.e., which SR pairs will be grouped together). Our aim is to consider all transitional RTs and thereby provide preliminary data on the role of stimulus properties on the organization of task representations.

Our analytical approach had two parts. First, to examine how stimulus set affected performance (a traditional task-switch cost), we used a two-way mixed design with a number of stimulus sets (1 or 2) as a between-subjects factor and switch (repeat or switch) as a within-subjects factor. The two groups of participants differed in terms whether the eight stimuli all belonged to one set (numbers 1–8) or two distinct sets (4 numbers and 4 black and white images of faces). Switches were defined as trials in which the stimulus on the previous trial indicated a response with the opposite hand as the stimulus on the current trials. For the two-set group, this meant that the stimulus set also switched (number ➔ face or face ➔ number) on consecutive trials.

Second, we evaluated the individual transitional RTs to generate a more complete, less theory-driven depiction of the task structure. Because we were not testing specific hypotheses about how tasks are organized with this alternative approach, we did not perform any inferential statistics on these transitional RTs. Instead, our goal was to compare how switch costs reflect more complete depictions of task structure gleaned from consideration of all transitions. To do this, we normalized RTs for each response and computed the normalized RT as a function of the previous response. Thus, we made no assumptions about which responses belong to the same set but instead used a data-driven approach to assess how responses appear to be grouped together according to the transitional RTs.

Method

Participants responded to a single stimulus appearing on each trial. A stimulus could appear in one of eight locations (four to left of center and four to right). Participants in the one-set group saw only numbers 1–8, each of which indicated one of the eight possible responses, the keys “s,” “d,” “f,” “g,” “h,” “j,” “k,” and “l” on the middle row of the qwerty keyboard. The four leftmost responses were made with the four fingers of the left hand, and the four rightmost responses were made with the four fingers of the right hand. The mapping was compatible so that 1 indicated the leftmost response, “s,” 2 the response second from the left “d,” etc. Participants in the two-set group saw numbers 1–4 and four faces that differed in terms of age (i.e., there was a child, college-aged adult, a middle-aged adult, and an older adult). The numbers were mapped to the leftmost responses so that the exact mappings for these stimuli were the same as in the one-set group. The faces were mapped to the rightmost responses in a compatible way so that the (clearly differentiable) ages were mapped in order from left to right. Pilot work in other studies has indicated that using this mapping is easy for participants and there are compatibility effects when performance on this mapping is compared to performance with other mappings. All stimuli (letters and numbers) were 0.8° visual angle presented in the center of the display.

Each trial began with presentation of a fixation cross in the center of the display for 500 ms. This was immediately followed by the stimulus which remained on screen for 1000 ms. The screen remained blank for 3000 ms. If the response was incorrect, a feedback screen showing the response mapping for all possible stimuli was displayed. If the response was correct, the next trial would begin.

Participants first completed a practice block of 16 trials and then 8 blocks of 32 trials in which each stimulus was presented four times. They then completed 2 blocks of 32 trials in which only a subset of the stimuli was presented. Data from these final two blocks will not be discussed here.[1]

Data Analysis

RTs from the first two blocks and first two trials of each block were not used in any of the analyses. Pilot data indicated that decreases in mean RT were much smaller after the first two blocks, making these data more stable for our transitional analyses. Moreover, we also eliminated all trials with an incorrect response and those immediately following an incorrect response. Error rates were low across all conditions (mean accuracy, 97%) and not analyzed further. RTs less than 200 ms and greater than 2000 ms were eliminated from the analyses.

First, we took the conventional approach and performed a 2 × 2 ANOVA on the RTs with group (one-set vs. two-set) as a between-subjects factor and switch (i.e.,

[1] The blocks were included to examine the separate question of whether the decreases in RT associated with reducing the number of stimulus and response alternatives depended on which stimuli were removed from the set. This question is not related to our focus, which whether conventional measures of task switching capture the structure of task as determined by the complete set of transitional RTs.

whether the response on the current trial required the same hand as the response on the previous trial) as a within-subjects factor.

Second, we evaluated every possible transitional RT excluding exact repetitions. To eliminate differences between the various effects (e.g., participants may be faster responding with their right index fingers than their left little fingers), we computed the mean and standard deviation of the RT for each response for each participant and recoded this as a Z-score. The Z-scores were computed after eliminating stimulus repetition trials from the data set. In this way, we could evaluate how much slower or faster a particular response was for a particular participant given the response of the previous trial. To simplify and increase the number of observations per cell, we ignored the direction of the transition and grouped together pairs of responses regardless of which occurred on the previous trial and which occurred on the present trial. This approach is justified by the strong correlation between opposite direction transitions (e.g., response 3 ➔ response 6 and response 6➔ response 3), $r = 0.90$. Thus, each Z-score represented the relative speed of that particular response for that particular participant.

Because this is presently an exploratory analysis, we attempted to visualize the data by using an open software package called Gephi (gephi.org) that depicts the underlying structure of networks. Each response was given a node, and the connections between the nodes (edges) were given a weight depending on the mean RT Z-score for that transition regardless of direction:

$$weight = e^{-5Z}$$

where Z represents the mean RT score from the particular response transition. The constant 5 was chosen to provide a range of weight strengths (e.g., 0.2–12). These weights were then used to create a force atlas that assumed each node (response) repelled the others with a force dependent on a global parameter but was also attracted to each other node depending on the weight.[2] This caused the nodes to be distributed in two-dimensional space such that nodes with shorter transitional RTs are represented by thicker edges and are closer to each other. That is, if making one response led to making another response on the subsequent trial faster than average (and vice versa), the two are represented close together and connected by a thick line, and if making one response led to the slower production of the other (and vice versa), the two are represented farther apart and connected by a thin line. The goal is to create a depiction of the transitional RTs that allows all of them to be considered simultaneously.

[2] The actual parameters used were as follows: inertia, 0.1; repulsion strength, 20,000.0; attraction strength, 10.0; maximum displacement, 10.0; auto-stabilized function, true; autostab strength, 80.0; autostab sensibility, 0.2; and gravity, 30.0. Only the repulsion strength was changed from the default value.

Results

The ANOVA performed on the RTs revealed no significant effect of group, $F(1,28) = 2.45$, $p = 0.13$, $\eta_p^2 = 0.08$ (one-set, 753 ms; two-set, 817 ms), but the effect of switch was significant, $F(1,28) = 43.75$, $p < 0.0001$, $\eta_p^2 = 0.61$ (repeat, 760 ms; switch, 810 ms) (Fig. 6.1). There was little indication of an interaction, $F(1,28) = 0.08$, $p = 0.78$, $\eta_p^2 = 0.003$ (one-set switch costs, 55 ms; two-set switch costs, 48 ms). In short, the ANOVA indicated that there was a significant cost of switching from the right hand to the left hand or vice versa but that this effect was the same for both groups. The magnitude of this costs was consistent with other studies reporting switch costs with univalent stimuli (e.g., Rogers & Monsell, 1995). Thus, based on the conventional approach, there is evidence for two task sets in both groups (separated by hand) and little evidence that the stimulus set manipulation affected the task representation as switch costs were nearly identical for the two groups. Alternatively, one could argue that costs do not reflect switch costs but instead indicate that the single task was organized by hand (e.g., Adam et al., 2003; Rosenbaum, 1980, 1983). In either case, the conventional analysis indicates that the SR mappings are grouped according to hand.

However, when we plotted the connection strengths of the responses for the two groups (Fig. 6.2), differences in the underlying structure became apparent.[3] For the one-set group (Fig. 6.2, Panel a), the structure was not readily characterized as two sets (clusters) but rather as a single chain that includes all responses. The connections between adjacent fingers, including left index [L4] and right index [R5], were stronger than the other connections, regardless of whether they were on the same or opposite sides. That is, the model recreated the relative positions of the eight

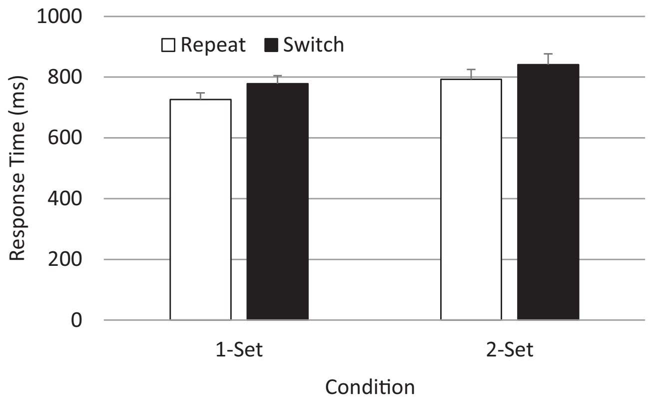

Fig. 6.1 Reaction times and standard errors for the four conditions of the experiment

[3] The Gephi software does not make an identical graph each time it is run. That is, the positions vary from run to run; although with networks this is simple, they are generally similar. The graphs we have chosen are highly typical of those produced by the software. Moreover, our conclusions are based on the strengths of the connection between responses, depicted by the edge thickness. This is a property of the data and does not change across runs.

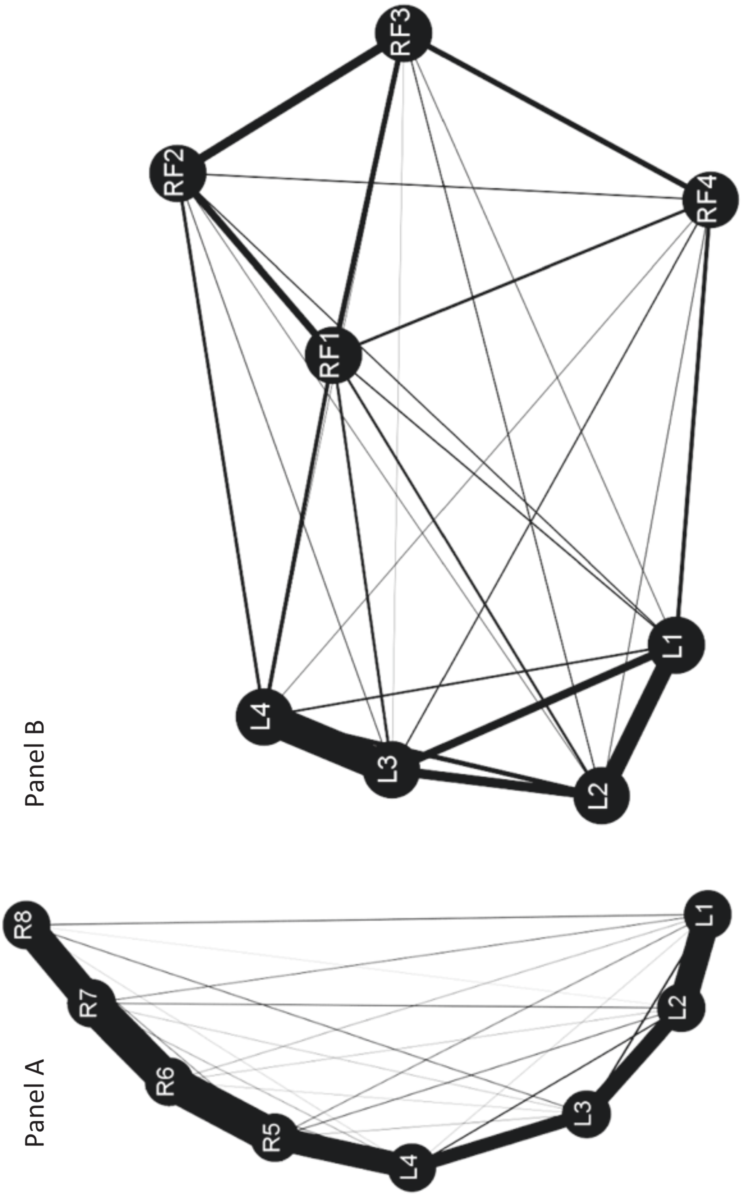

Fig. 6.2 Visual depictions of the structure of the transitional RTs from the two conditions of the experiment. Thickness of the lines and distance between nodes (edges) represent the transitional RT between responses, with thicker edges and shorter distances signifying shorter transitional RTs. Panel (**a**) shows the structure observed when numbers 1–8 were mapped to eight responses, four for the left hand (L1–L4) and four for the right hand (R5–R8). Panel (**b**) shows the structure when the left responses (L) were mapped to the numbers (1–4) and the right responses (R) were mapped to four faces (F1–F4)

responses in physical space given only the transitional RTs. However, because neighboring responses were much more common among responses that also share the same hand, this pattern produced a robust switch cost.

The pattern of transitional RTs produced a different task structure for the two-set group (Fig. 6.2, Panel b). Here, although the switch costs were nearly identical (and numerically smaller), the structure looked more like there are two subtasks, with the left-hand responses in one cluster and the right-hand responses in another. Moreover, while the left-hand responses, which were mapped to numbers, were aligned so that neighboring responses were closely associated with each other, the right-hand responses, which were mapped to faces, formed a quadrilateral. The right-hand responses are represented as quadrilateral because the edges are relatively weak and vary less (c.f. the range of edge thicknesses between left-sided responses and between right-sided responses).

Discussion

Although we designed the tasks to minimize attentional and inhibitory processes, robust switch costs were observed when participants switched from a left-sided response to a right-sided one or vice versa. Critically, the switch costs were nearly identical for the one-set and two-set groups. The traditional interpretation of these switch costs would suggest that both groups divided the tasks into sets based on response hand. That is, based on the switch cost, it appears that the stimuli had little effect on how the groups represented the tasks.

However, when we examined all the transitions between responses, it appeared that the switch costs reflected different factors across the two groups that differed in terms of the stimulus sets. Visualizations of the structure of the transitions suggested that, for the one-set group, the switch cost reflected the short RTs associated with neighboring responses, which were much more frequent for same-side response than for different-side responses (Fig. 6.2, Panel a). In contrast, for the two-set group, the cost appeared to reflect something more closely related to the conventional conceptualization of task sets (Fig. 6.2, Panel b). The left-side and right-side responses formed an approximation of two clusters – one for each task/stimulus set/hand.

Intriguingly, there were also differences among the strengths of edges within the two clusters, particularly the edges between left-hand responses. This provides evidence that the left- and right-side sets were structured differently: the left-side set was organized as a chain, with neighboring responses exclusively showing strong connections, whereas the right-side set was less organized less like a chain. Thus, the left-hand alternatives appear to be organized by the ordinal relationships of the stimuli and/or the relative locations of the responses, not simply as unstructured "set" of SR mappings.

We make no strong claims about the factors that affect the task representation or how tasks are generally represented. The patterns of transitional RTs may reflect a variety of factors, including switching from making the response with one hand to the other, processing different types of visual information, and retrieving mappings from memory. This was an exploratory analysis without a priori hypotheses. Our point is that different tasks that produce near-identical switch costs can have distinct underlying structures when the full set of transitional RTs is considered. The structure observed for the one-set condition does not appear consistent with the encoding of two task sets even though a significant switch cost was observed.

Thus, the present data demonstrate the pitfalls of taking switch costs as indicators of how tasks are represented. Instead, we propose that considering all possible transitions and visualizing the resulting structure may be useful for generating new hypotheses that do not rely on the premise that the experimenter has a priori knowledge of the task representation. With further work, testable hypotheses can be developed that specify the factors that determine the task representation. Consideration of these alternative hypotheses is not clearly motivated by traditional measures of switch costs but may only become apparent when finer-grained analyses of transitional RTs are used. For example, the present data suggest the use of numbers with a compatible SR mapping leads to strongly "linear" (i.e., strong connections between neighboring responses, weak connections elsewhere) representations, whereas other types of stimuli that may be distinguished in terms of non-ordinal relationships may produce different organizational clusters. In this way, evaluating the task structure can provide insight into how the encoding of tasks produces SR compatibility. That is, the connections can reveal interrelationships among items that reflect element-level compatibility (Fitts & Deininger, 1954; Kornblum et al., 1990).

The broader implications of this finding are that, at least when the tasks are sufficiently complex (i.e., have a sufficiently large number of stimuli and/or responses), there are effects on transitional RT (i.e., effects of the previous response on the current one) that are not readily attributable to attention and inhibition but appear to relate to the structure of the task representation. Therefore, caution is recommended when interpreting transitional RT effects, including switching (e.g., Rogers & Monsell, 1995), binding (e.g., Frings et al., 2020; Hommel, 1998), and anatomical effects (e.g., Collins & Frank, 2016).

Given that task representations affect transitional RTs when attentional, inhibitory, and binding demands are minimal, it is likely that they also contribute to RT when they are present. It is unclear how to disentangle the contributions of these various effects. However, it may be prudent to consider how binding or inhibitory effects, for example, are impacted by changes in the number of SR alternatives or other manipulations that affect the structure of a task to argue against alternative explanations. Considering all the possible transitions individually may reveal the factors that have the most salient effects on RT across an array of possible transitions.

Finally, we note that events that do not require responses, such as task cues (e.g., Allport et al., 1994; Kikumoto & Mayr, 2020; Mayr & Bryck, 2005; Qiao et al., 2017) and precues (e.g., Adam et al., 2003; Cookson et al., 2016; Cookson et al., 2020; Miller, 1985; Rosenbaum, 1980), can be evaluated in terms of their effects on specific responses. It is possible that such an approach would reveal that some task cues or precues show variable effectiveness for different responses within the set that they indicate (see Lien et al., 2005). Such variability might reflect the structure of the task representation as particular responses may be more strongly associated with the other members of the cued group.

Summary

Cognitive control is often framed as a process of selecting some SR associations over others, but there is a wealth of evidence indicating that SR associations are not adequate for describing how voluntary behavior is guided by sensory systems (see Hazeltine & Schumacher, 2016). Not understanding how tasks are represented poses a serious obstacle for theories of cognitive control. A better conceptualization of the task representations governed by control processes will help specify how they operate.

The dominant description of task representations is the task set, a collection of SR associations, whose presence is inferred primarily through task-switch costs. However, we argue that this approach has serious limitations that are often ignored. First, task switch costs likely reflect numerous processes, including those relating to attention and inhibition, that vary across experimental procedures and complicate their interpretation. While attention and inhibition, for example, are considered related to cognitive control, their roles in task representation are less clear. The present data indicate that even when all responses are made to univalent stimuli presented alone, structure in the transitional RTs is observed.

Second, the switch cost measure is coarse in that it lumps transitions into a small number of (usually two) categories. As we demonstrate empirically, this procedure can produce misleading results. Observing a performance cost when a putative task boundary is crossed may be too coarse a measure to adequately describe how a task is organized. Unfortunately, alternative organizations that may produce the observed cost are rarely considered.

It may be productive to abandon the notion that task representations consist of packets of SR associations. Instead, we should consider how tasks are structured by evaluating how the performance of different components of task affects others. This can be done without assuming that the task representation relies on grouped SR associations. Each action may be bound to others at different levels of a hierarchical representation (Gozli, 2019; Schumacher & Hazeltine, 2016), which may produce complex effects that are not easily categorized in terms of membership in a task set. Coarse measures of transitional effects such as task switch costs may reify this simplistic task set account and therefore should be used with caution. In short, the observation of task switch cost does not necessarily indicate that behavior is generated by SR association organized into task sets.

References

Adam, J. J., Hommel, B., & Umilta, C. (2003). Preparing for perception and action (I): The role of grouping in the response-cuing paradigm. *Cognitive Psychology, 46*, 302–358.

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance (Vol. attention and performance XV)* (pp. 421–452). Harvard University Press.

Arrington, C. M., & Logan, G. D. (2004). The cost of a voluntary task switch. *Psychological Science, 15*, 610–615. https://doi.org/10.1111/j.0956-7976.2004.00728.x
Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences, 12*(5), 193–200. https://doi.org/10.1016/j.tics.2008.02.004
Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical. *Nature Reviews Neuroscience, 10*, 659–669.
Badre, D., & Nee, D. E. (2018). Frontal cortex and the hierarchical control of behavior. *Trends in Cognitive Sciences, 22*, 170–188. https://doi.org/10.1016/j.tics.2017.11.005
Badre, D., Bhandari, A., Keglovits, H., & Kikumoto, A. (2021). The dimensionality of neural representations for control. *Current Opinion in Behavioral Sciences, 38*, 20–28.
Barcelo, F., Escera, C., Corral, M. J., & Periánez, J. A. (2006). Task switching and novelty processing activate a common neural network for cognitive control. *Journal of Cognitive Neuroscience, 18*, 1734–1748.
Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge University.
Bezdek, M. A., Godwin, C. A., Smith, D. M., Hazeltine, E., & Schumacher, E. H. (2019). How conscious aspects of task representation affect dynamic behavior in complex situations. *Psychology of Consciousness: Theory, Research, and Practice, 6*, 225–241.
Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652.
Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences, 16*, 106–113.
Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology, 13*, 207–230.
Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332–361.
Collins, A. G. E., & Frank, M. J. (2016). Motor demands constrain cognitive rule structures. *PLoS Computational Biology, 12*, e1004785. https://doi.org/10.1371/journal.pcbi.1004785
Cookson, S. L., Hazeltine, E., & Schumacher, E. H. (2016). Neural representations of stimulus-response associations during task preparation. *Brain Research, 1648*, 496–505.
Cookson, S. L., Hazeltine, E., & Schumacher, E. H. (2020). Task structure boundaries affect response preparation. *Psychological Research, 84*, 1610–1621. https://doi.org/10.1007/s00426-019-01171-9
Courtney, S. M., Roth, J. K., & Sala, J. B. (2007). A hierarchical biased competition model of domain-dependent working memory maitenance and executive control. In *Working memory: Behavioural and neural correlates* (pp. 369–383).
Derrfuss, J., Brass, M., Neumann, J., & von Cramon, D. Y. (2005). Involvement of the inferior frontal junction in cognition control: Meta-analyses of switching and Stroop studies. *Human Brain Mapping, 25*, 22–34.
Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.
Dosenbach, N. U. F., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences, 12*(3), 99–105. https://doi.org/10.1016/j.tics.2008.01.001
Dreisbach, G. (2012). Mechanisms of cognitive control: The functional role of task rules. *Current Directions in Psychological Science, 21*, 227–231.
Duncan, J. (2013). The structure of cognition: Attentional episodes in mind and brain. *Neuron, 80*, 35–50.
Dykstra, T., Smith, D. M., Schumacher, E. H., & Hazeltine, E. (in prep). Measuring task structure with transitional response times: Task representations are more than task sets.
Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143–149.
Fitts, P. M., & Deininger, R. L. (1954). S-R compatibility: Correspondence among paired elements within stimulus and response codes. *Journal of Experimental Psychology, 48*, 483–492.

Frings, C., Hommel, B., Koch, I., Rothermund, K., Dignath, D., Giesen, C., & Philipp, A. (2020). Binding and retrieval in action control (BRAC). *Trends in Cognitive Sciences*, 375–387. https://doi.org/10.1016/j.tics.2020.02.004

Fuster, J. M. (2008). *The prefrontal cortex* (4th ed.). Elsevier.

Gopher, D., Armony, L., & Greenshpan, Y. (2000). Switching tasks and attention policies. *Journal of Experimental Psychology: General, 129*, 309–339.

Gozli, D. (2019). *Experimental psychology and human agency*. Springer.

Grant, L. D., Cookson, S. L., & Weissman, D. H. (2020). Task sets serve as boundaries for the congruency sequence effect. *Journal of Experimental Psychology: Human Perception and Performance, 46*, 798–812.

Halvorson, K. M., & Hazeltine, E. (2015). Do small dual-task costs reflect ideomotor compatibility or the absence of crosstalk? *Psychonomic Bulletin & Review, 22*, 1403–1409. https://doi.org/10.3758/s13423-015-0813-8

Hayes, A., Davidson, M., Keele, S. W., & Rafal, R. D. (1998). Toward a functional analysis of the basal ganglia. *Journal of Cognitive Neuroscience, 10*, 178–198.

Hazeltine, E. (2005). Response-response compatibility during bimanual movements: Evidence for the conceptual coding of action. *Psychonomic Bulletin & Review, 12*, 682–688.

Hazeltine, E., & Schumacher, E. H. (2016). In B. Ross (Ed.), *Understanding central processes: The case against simple stimulus-response associations and for complex task representation* (Vol. 64, pp. 195–245). Psychology of Learning and Motivation.

Hazeltine, E., Akçay, Ç., & Mordkoff, J. T. (2011a). Keeping Simon simple: Examining the relationship between sequential modulations and feature repetitions with two stimuli, two locations, and two responses. *Acta Psychologia, 136*, 245–252.

Hazeltine, E., Lightman, E., Schwarb, H., & Schumacher, E. H. (2011b). The boundaries of sequential modulations: Evidence for set-level control. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 1898–1914.

Hirsch, P., Schwarzkopp, T., Declerck, M., Reese, S., & Koch, I. (2016). Age-related differences in task switching and task preparation: Exploring the role of task-set competition. *Acta Psychologia, 170*, 66–71.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus–response episodes. *Visual Cognition, 5*, 183–216.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Science, 8*, 494–500.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC). *Behavioral and Brain Sciences, 24*, 849–878.

Hommel, B., Proctor, R. W., & Vu, K. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological Research, 68*, 1–17.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology, 89*.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching-a review. *Psychological Bulletin, 136*(5), 849–874. https://doi.org/10.1037/a0019842

Kikumoto, A., & Mayr, U. (2020). Conjunctive representations that integrate stimuli, responses, and rules are critical for action selection. *Proceedings of the National Academy of Sciences, 117*, 10603–10608.

Kim, C., Cilles, S. E., Johnson, N. F., & Gold, B. T. (2012). Domain general and domain preferential brain regions associated with different types of task switching: A meta-analysis. *Human Brain Mapping, 33*, 130–142.

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking – An integrative review of dual-task and task-switching research. *Psychological Bulletin, 144*, 557–583. https://doi.org/10.1037/bul0000144

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences, 11*(6), 229–235. https://doi.org/10.1016/j.tics.2007.04.005

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility--A model and taxonomy. *Psychological Review, 97*, 253–270.

Lien, M.-C., Ruthruff, E., Remington, R. W., & Johnston, J. C. (2005). On the limits of advance preparation for a task-switch: Do people prepare all of the task some of the time or some of the task all the time. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 299–315.

Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review, 109*, 376–400.

Logan, G. D. (2004). Working memory, task switching, and executive control in the task span procedure. *Journal of Experimental Psychology: General, 133*, 218–236. https://doi.org/10.1037/0096-3445.133.2.218

Logan, G. D., & Schneider, D. W. (2006). Interpreting instructional cues in task switching procedures: The role of mediator retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*, 347–363. https://doi.org/10.1037/0278-7393.32.3.347

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal cortex and anterior cingulate cortex in cognitive control. *Science, 288*, 1835–1838.

Mayr, U. (2002). Inhibition of action rules. *Psychonomic Bulletin & Review, 9*, 93–99.

Mayr, U., & Bryck, R. L. (2005). Sticky rules: Integration between abstract rules and specific actions. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*, 337–350.

Mayr, U., Diedrichsen, J., Ivry, R. B., & Keele, S. W. (2006). Dissociating task-set selection from task-set inhibition in the prefrontal cortex. *Journal of Cognitive Neuroscience, 18*, 14–21. https://doi.org/10.1162/089892906775250085

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology, 41*, 211–253.

Miller, J. (1985). A hand advantage in preparation of simple keypress responses: Reply to reeve and Proctor (1984). *Journal of Experimental Psychology: Human Perception and Performance, 11*, 221–233.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Mittelstädt, V., Miller, J., & Kiesel, A. (2018). Trading off switch costs and stimulus availability benefits: An investigation of voluntary task-switching behavior in a predictable dynamic multitasking environment. *Memory & Cognition, 46*, 699–715.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*, 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Monsell, S., & Driver, J. (2000). Banishing the control homunculus. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 3–32). MIT Press.

Murray, S., Krasich, K., Schooler, J. W., & Seli, P. (2020). What's in a task? Complications in the study of the task-unrelated-thought variety of mind wandering. *Perspectives on Psychological Science, 15*, 572–588. https://doi.org/10.1177/1745691619897966

Norman, D. A. (1981). Categorization of action slips. *Psychological Review, 88*, 1–15.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1–18). Plenum.

Petrides, M. (2006). The rostro-caudal axis of cognitive control procesing within lateral frontal cortex. In *From monkey brain to human brain: A Fyssen Foundation symposium* (pp. 293–314).

Poljac, E., Koch, I., & Bekkering, H. (2009). Dissociating restart cost and mixing cost in task switching. *Psychological Research, 73*, 407–416.

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167–201). Springer-Verlag.

Qiao, L., Chen, A., & Egner, T. (2017). Dynamic trial-by-trial recoding of task-set representations in the frontoparietal cortex mediates behavioral flexibility. *Journal of Neuroscience, 37*, 11037–11050. https://doi.org/10.1523/JNEUROSCI.0935-17.2017

Qiao, L., Xu, M., Zhang, L., Li, H., & Chen, A. (2020). Flexible adjustment of the effective connectivity between the fronto-parietal and visual regions supports cognitive flexibility. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2020.117158

Rescorla, R. A. (1988a). Behavioral studies of Pavlovian conditioning. *Annual Review of Neuroscience, 11*, 329–352.

Rescorla, R. A. (1988b). Pavlovian conditioning: It's not what you think it is. *American Psychologist, 43*, 151–160.

Rogers, R., & Monsell, S. (1995). The costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: Human Perception and Performance, 124*, 207–231.

Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General, 109*, 444–474.

Rosenbaum, D. A. (1983). The movement precuing technique: Assumptions, applications, and extensions. In R. A. Magill (Ed.), *Memory and control of action* (pp. 230–274). North-Holland Publishing Company.

Ruthruff, E., Remington, R. W., & Johnston, J. C. (2001). Switching between simple cognitive tasks: The interaction of top-down and bottom-up factors. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 1404–1419.

Sakai, K. (2008). Task set and prefrontal cortex. In *Annual review of neuroscience* (Vol. 31, pp. 219–245). Annual Reviews.

Schmidt, J. R., De Houwer, J., & Rothermund, K. (2016). The parallel episodic processing (PEP) model 2.0: A single computational model of stimulus-response binding, contingency learning, power curves, and mixing costs. *Cognitive Psychology, 91*, 82–108.

Schumacher, E. H., & Hazeltine, E. (2016). Hierarchical task representation: Task files and response selection. *Current Directions in Psychological Science, 25*, 449–454. https://doi.org/10.1177/0963721416665085

Schumacher, E. H., & Schwarb, H. (2009). Parallel response selection disrupts sequence learning under dual-task conditions. *Journal of Experimental Psychology: General, 138*, 270–290.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B, 298*, 199–209.

Simon, J. R. (1969). Reactions towards the source of stimulation. *Journal of Experimental Psychology, 81*, 174–176.

Smith, D. M., Dykstra, T., Hazeltine, E., & Schumacher, E. H. (2020). Task representation affects the boundaries of behavioral slowing following and error. *Attention Perception & Psychophysics, 82*, 2315–2326. https://doi.org/10.3758/s13414-020-01985-5

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Takacs, A., Mückschel, M., Roessner, V., & Beste, C. (2020). Decoding stimulus-response representations and their stability using EEG-based multivariate pattern analysis. *Cerebral Cortex Communications*. https://doi.org/10.1093/texcom/tgaa016

Tolman, E. C. (1932). *Purposive behavior in animals and men*. Century.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*, 189–208. Retrieved from https://pdfs.semanticscholar.org/0874/a64d60a23a20303877e23caf8e1d4bb446a4.pdf

Tuckey, M. R., & Brewer, N. (2003). The influence of schemas, stimulus ambiguity, and interview schedule on eyewitness memory over time. *Journal of Experimental Psychology: Applied, 9*, 101.

Verguts, T., & Notebaert, W. (2008). Hebbian learning of cognitive control: Dealing with specific and nonspecific adaptation. *Psychological Review, 115*(2), 518–525. https://doi.org/10.1037/0033-295x.115.2.518

Waskom, M. L., Kumaran, D., Gordon, A. M., Rissman, J., & Wagner, A. D. (2014). Frontoparietal representations of task context support the flexible control of goal-directed cognition. *Journal of Neuroscience, 34*, 10743–10755.

Weissman, D. H., Jiang, J., & Egner, T. (2014). Determinants of congruency sequence effects without learning and memory confounds. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 2022–2037.

Wifall, T., Hazeltine, E., & Mordkoff, J. T. (2015). The roles of stimulus and response uncertainty in forced-choice performance: An amendment of hick/Hyman law. *Psychological Research Psychologische Forschung*. https://doi.org/10.1007/s00426-015-0675-8

Wisniewski, D., Reverberi, C., Tusche, A., & Haynes, J.-D. (2015). The neural representation of voluntary task-set selection in dynamic environments. *Cerebral Cortex, 25*, 4715–4726.

Woolgar, A., Hampshire, A., Thompson, R., & Duncan, J. (2011). Adaptive coding of task-relevant information in human frontoparietal cortex. *Journal of Neuroscience, 31*, 14592–14599.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs". *Psychological Research, 63*, 212–233.

Chapter 7
The Problem of Interpretation in Experimental Research

Davood Gozli

When we design an experiment, we set out to compare a number of conditions with respect to a set of dependent variable(s). The design serves its intended purpose when the conditions are similar except with respect to our independent (manipulated) variables. A well-known error arises when an independent variable is conflated with an unintended change (a confound). In the present chapter, we are interested in a particular type of confound, namely, the meaning participants assign to events in the experiment. By raising the question, “What else could the events mean for the research participants?,” we are raising another closely connected question, “What else can the findings mean?” If the meaning assigned to events changes across experimental conditions, we can no longer assume that independent variables are manipulated within otherwise “controlled” conditions. Recognizing how the meaning of events might have changed for the research participants in turn changes the meaning of the experimental findings, which could deflate or undermine both the rationale and the theoretical significance of the research.

Let us illustrate the main point with a simple example. Imagine that we are interested in the effect of the loudness of task instructions on participants’ performance. We divide participants into two groups, one receives instruction in normal voice and the other receives instruction in a loud voice. Some of the participants in the “loud” condition might interpret the experimenter’s loudness as impatience, rudeness, or negative mood. If that is the case, describing the two conditions in terms of loudness alone would be inadequate. Consequently, any observed difference in performance cannot be attributed to the loudness of instructions alone. To explain a difference between the two conditions, we will have to consider that a loudly delivered instruction might have a different meaning, compared with the same instruction delivered in a normal voice. Such differences in interpretation can produce differences in the perceived context of research participation (Bergner, 2010, 2016). A condition in which participants follow the instructions of a rude researcher represents a qualitatively different context, compared to one where the researcher behaves politely (Johns, 2006).

D. Gozli (✉)
University of Macau, Taipa, Macau S.A.R., China

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9_7

To formulate the problem more generally, researchers might want to test whether there is a relation between two variables (e.g., loudness of the verbal instructions and the speed of task performance) while assuming no change in the kind of processes under investigation. However, differences in the participants' interpretation can undermine that assumption (Toomela, 2008; Valsiner & Brinkmann, 2016). There is no guarantee for participants' perspective to uniformly conform to what is assumed in the experimental design and the categories applied in advance to what happens during the experiment. Across different experimental conditions, events might differ partly in ways that are intended by the experimenters (e.g., loudness of voice) and partly in ways that are unintended (e.g., perceived rudeness or impatience of the experimenter).

In the following sections, I first review several studies that have explicitly addressed the role of meaning and interpretation. The message from these studies is that changing the meaning assigned to a task (e.g., how the procedure is described to the participants), without changing anything else about the experimental situation, can change the results by influencing participants' sensitivity and responsiveness to the situation. Next, I turn to another set of studies that illustrate how people detect normative dimension in a given situation, including the experimental situation. These studies suggest that norms can be detected rapidly, automatically, and without explicit instructions. Participants can move from "X happened" to "X ought to happen" or from "X ought not to happen" to "X is acceptable." For our purpose, it is important that research participants might consider a type of action desirable or acceptable in a situation, without the researchers recognizing it. There might be, in other words, a mismatch between the normative situation perceived by the participants and those presupposed by the researchers. I will next turn to several examples in experimental research that involve neglecting possible changes in meaning. The possible mismatch between how participants experience events, on the one hand, and what researchers believe about the participants' experience, on the other hand, is important with regard to the meaning of the experimental findings. The present argument, therefore, intends to show the problem of interpretation in experimental psychology and the continuing relevance of theoretical psychology in experimental research.

Meaning of Events

In explaining human performance, it might appear that all we need is third-person knowledge of the structure of the task, including what features of the environment people are acting upon and what types of movements are available to them. Third-person knowledge of a task does, under some circumstances, give us some predictive ability. Before beginning to discuss variations in task performance, however, the participants' intention (e.g., agreeing to respond to stimuli according to the instructions), as well as their selective attention and interpretation of the situation, is assumed to (a) mirror the researchers' instructions and (b) be fixed throughout the experimental session and across conditions. Such assumed transparency enables researchers to bracket out the participants' interpretation and construct third-person descriptions that appear complete and independent of the participants' interpretation (Mammen, 2017). In the present section, I review research from several fields in cognitive-experimental psychology, all sharing a common theme: participants' understanding of a task (first-person perspective) can change how the task is performed and, consequently, what factors that can further influence performance, even in the absence of any overt change (third-person description) in the situation and task features.

In simple stimulus-response (S-R) tasks, people are on average faster with congruent S-R arrangements (left and right keys paired, respectively, with left and right stimuli) than with incongruent arrangements (left and right keys paired, respectively, with right and left stimuli). A variant of this phenomenon is the effect of the relation between an irrelevant spatial feature of the stimulus and the response (Simon, 1990). For example, imagine a task in which people are instructed to use left/right keys to respond to high−/low-pitched tone. The tone could be emitted either from the left or the right side. The pitch of the tone is relevant to the task, and its location is irrelevant. Nevertheless, on average, responses are faster when the tone and response are congruent in their location, compared to when they are incongruent. This is known as the Simon effect (Hommel, 2011).

In a landmark study, Hommel (1993) reported that the Simon effect can be reversed just by changing the instructions delivered to the participants. In his experiment, participants responded to low−/high-pitched tones using left/right keys. Each key was connected to a light on the opposite side, such that pressing the right key would illuminate a light on the left-hand side and pressing the left key would illuminate a light on the right-hand side. What was special about this task was the ambiguity of the response location. How do you respond to a low-pitched tone? Do you press the *key* on the *left*-hand side or turn on the *light* on the *right*-hand side? Both descriptions were available to the participants. Hommel (1993) divided the participants into two groups. For the first group, the instructions emphasized the location of the keys ("press the left key if you hear low tone"). For the second group, the instructions emphasized the location of the lights that turned on with key-press responses ("turn on the light on the right side if you hear the low tone"). Despite identical physical properties of the two conditions, the Simon effect was reversed across the two groups. Thus, the intended response location, i.e., how participants conceived of the responses, made a qualitative difference in the interaction between stimulus and response features.

Another example of how interpretation can change a behavioral effect was found in the joint version of the Simon task. Imagine the Simon task performed by two participants, sitting side by side, such that each is responsible for one response (Sebanz et al., 2003). For instance, two participants respond to the onset of a single red/green visual stimulus that could appear on the left/right side of the screen. We instruct the participant sitting on the left side to respond to the green stimuli, ignoring stimulus location, while instructing the participant on the right-hand side to respond to the red stimuli, again ignoring stimulus location. In this setup, when a stimulus appears on the left side, regardless of its color, the participant sitting on the left tends to be (on average) faster than the participants sitting on the right side. Likewise, in response to a stimulus on the right side, the participant sitting on the right would be, on average, relatively faster. This *joint* Simon effect suggests that two actors take into account, in addition to what they are instructed to do, the role of their co-actor.

Hommel et al. (2009) introduced an additional manipulation in a joint Simon task. Before the co-actors began performing the task together, the authors induced either a cooperative or a competitive relation between them. In the cooperative condition, the authors found a joint Simon effect, suggesting that participants took each other's role into account. By contrast, in the competitive condition, the authors did not find a joint Simon effect. These findings suggest that changing the way two people think about each other can affect the organization of a shared task and, subsequently, how the shared task is susceptible to further manipulations, even if the overt physical structure of the task remains the same.

With a task that is performed by one person, the presence of another person who observes task performance can change the meaning of the task, even when all other aspects are the task and instruction are kept the same. Sartori et al. (2009) asked participants to perform a series of manual actions (reaching and lifting) with some objects on a table. In the "individual" (control) condition, participants performed the actions alone according to the instructions. When they were being observed by a fellow participant, however, their movements differed in subtle ways. The researchers reasoned that the communicative intention—showing one's movement to the other person—changed performance characteristics. A follow-up experiment confirmed the role of communicative intention, as opposed to the mere presence of another person, by testing the effect of a blindfolded person. As predicted, the blindfolded person did not cause changes in movement characteristics, compared with the "individual" condition (Sartori et al., 2009; see also Quesque et al., 2013, for a related demonstration in experimental economics, and see Dana et al., 2007).

The next example is from a task-switching study. Experiments that investigate task switching require participants to learn and prepare for two distinct tasks, performing only one task per trial (e.g., Task 1, judging a number's parity; Task 2, judging if the number is smaller/larger than 5). A cue might determine which task is performed at any given time (e.g., the color of the number). The typical finding is a performance cost (slower response time and reduced accuracy), when the task switches from trial n to trial $n + 1$, compared with when the task repeats. But does the switch cost depend on participants' understanding that they are, indeed, performing two distinct tasks? To answer this question, Dreisbach et al. (2007) provided different instructions for the same task (involving eight stimuli and two responses). One group of participants was instructed to perform a single task involving eight alphanumerical items mapped, arbitrarily, onto two responses. For the second group, the task was described, less arbitrarily, in terms of two sub-tasks: a "number task" and a "letter task" (each sub-task associated with four stimuli and two responses). Dreisbach et al. found a robust switch cost (when the stimulus switched from letter to number and vice versa) in the latter group and found no switch cost in the former group. In short, the presence of a switch cost, which is itself indicative of how a task is organized, depended on participants' understanding that there were two distinct tasks (Gozli, 2019, Chap. 5).

Another example comes from a visual-search study. Huffman et al. (2017) used a visual search, in which participants looked for a non-salient target (a green circle among green squares). A salient distractor could also be present on the display (a red square among the green items), which was supposed to be ignored. In one condition, the location of the salient distractor changed randomly from trial to trial, while in another condition, it was predictable, moving in a clockwise pattern. First, Huffman et al. found that the salient distractor interfered with performance more when its location changed predictably, compared to when it changed randomly. More relevant for our purpose, the researchers compared participants who noticed the predictability of the distractor location with those who did not notice it. The former group was presumably more likely to keep track and *actively* ignore the distractor, but it was for this group that the distractor caused the largest interference. Therefore, for the same task, with the same search arrays and the same instructions, participants' understanding that a salient distractor is predictable and should be ignored increased the cost of the distractor on performance. One might argue that participants' awareness of the predictable distractor resulted from their noticing the higher-performance cost of the distractor and not the other way around. Regardless, what is important in the present context is the association between, on the one hand, different meanings assigned to the distractor (predictable vs. unpredictable) and, on the other hand, different performance costs of the distractor. Without inquiring about the participants' points of view and asking whether they were keeping track of the distractor, we could only attribute variations in performance to external factors.

As a final example for this section, we can turn to a study on the effect of perceived effort and commitment of a partner in shared task. Székely and Michael (2018) measured participants' commitment to a game, defined as the time taken before the participant quits a round of the game. The authors used a computer-mediated two-party game, which became increasingly boring over time. Participants were playing with a "partner" that was, in fact, a computer algorithm. At the beginning of each round, the "partner" had to unlock the round by solving a CAPTCHA problem that was either easy or difficult. The participant could either end the round, by pressing the space bar, or wait for their partner to solve the problem. Participants showed more commitment to the game as a result of the perceived effort of their partner. This was observed only when they believed the partner was another person (Experiment 1) and not when they believed the partner was a computer program (Experiment 2) or when they played the game alone (Experiment 3). Therefore, even when the superficial characteristics of a game remain the same, participants' decisions change based on the meaning they attribute to events.

Detection and Adoption of Norms

Meaning of events and situations can vary across many different dimensions, and central among them are those groups under the "normative" category (Brinkmann, 2010). Norms feature in our experience not as isolated and detached individuals but as members of communities and groups (Searle, 1995). A norm cannot remain a norm if everyone around you, or everyone in your group, violates it. Thus, the way we think about a norm is sensitive to how others—particularly others in our group—regard the norm. In reference to the meaning of actions, the normative meaning can change with a change in the social context (Bergner, 2016), particularly in response to how other people evaluate the given action.

If someone breaks a norm (or a rule) repeatedly, then their evaluation of the norm or their self-judgment might change in order to rationally adjust to their own behavior (Festinger, 1964). Imagine a person violating a norm for the first time (at time t_1), then for the second time (t_2), and so on until the tenth time (t_{10}). Among other considerations, we ought to consider whether and how the meaning of the norm changes in the perspective of this person. Particularly when breaking the norm is associated with a positive outcome, and no negative outcome, it is possible that the violation is regarded as more acceptable at t_{10} than at t_1.

Our general sensitivity to norms has been demonstrated in developmental studies with children. A study with 3-year-olds suggested that when children observe someone play with a toy for the first time, even once, they interpret the use in a normative sense ("one ought to play with the toy in this manner"; Schmidt et al., 2016). This claim was based on the observation that when the children later see someone else play with the toy in a different way, they object and try to enforce the way in which the toy "ought to" be used.

Efficient adoption of norms in children has been linked to the phenomenon of over-imitation. In imitating others, human children imitate both causally relevant and causally irrelevant steps of procedure, compared to non-human primates who tend to imitate only the causally relevant steps (Whiten et al., 2009). Kenward et al. (2010) asked whether over-imitation in human children should be attributed to norm learning or to a distorted causal learning. They designed an apparatus through which children were instructed to retrieve objects. Children were then asked to provide verbal explanations for their understanding of the causal mechanisms. The apparatus was in a transparent case, and the instructions to reach the objects included both causally relevant (moving the objects with a stick) and irrelevant actions (inserting the stick into an empty compartment). Although the children's action included the irrelevant action, their explanations suggested that they performed the unnecessary action out of norm learning and not for misunderstanding the causal mechanism.

The label "unnecessary action" (i.e., instrumentally superfluous) should be used with caution. Picking up and conforming to seemingly unnecessary actions play a crucial role in communication and in sustaining our social-cultural reality (Toomela, 2016). A polite gesture could appear unnecessary while being *not* unnecessary (such double-negations are discussed by Engelsted, 2017), if it signifies something about the relationship, its cultural embeddedness, its history, and its anticipated future. Thus, our capacity to detect norms or rules rapidly and flexibly should be viewed in light of the complex and dynamic nature of our social reality (Mammen, 2008). Indeed, the presence of another person can change the likelihood of rule-breaking, depending on their stance on (the meaning they assigned to) rule-breaking. Simons-Morton et al. (2014) studied rule violations in young males' driving with a driving simulator with or without a passenger (confederate). They compared traffic rule violations when participants drove (a) alone, (b) with passenger who was accepting of risk, and (c) in the company of a passenger who was aversive to risk. They found that the presence of a risk-accepting passenger can increase the likelihood of committing traffic rule violation. These findings demonstrate the social nature of rules and norms and the role of others' perspective in determining what one ought to do.

The influence of others can be exerted indirectly, based on the observable effects of their actions. It might be possible, for instance, that observing the violation of one norm could promote the violation of other norms. Keizer et al. (2008) described such phenomena as "cross-norm inhibition" (p. 1683). In their first study, Keizer et al. staged different conditions in public bicycle-parking areas. In their so-called "disorder" condition, they covered the wall with graffiti right next to a clearly visible sign prohibiting graffiti. In the "order" condition, they kept the wall clean. Moreover, flyers were attached to the handlebars of the parked bicycles, and the owners had to remove it before riding the bicycle. The question was whether people litter when they remove the flyers attached to their bicycles, particularly in the presence of an already existing violation. The authors found that, when seeing rule violation in one domain (graffiti against the rules), people become more likely to violate a rule in another domain (littering).

Neglecting Meaning

In this section, I turn to several studies that seem to have neglected the role of interpretation and meaning. Explanations are offered with reference to situational factors, but not with reference to the possible differences in subjective meaning assigned to those factors. Researchers might only be interested in how participants respond overtly to stimuli and how those responses change with changes in situational factors, without concern for the meaning assigned to those events. Careful experimental designs and instructions, which negotiate and clarify the meaning of events prior to collection of data, are attempts to side-step to the issue of meaning (Wachtel, 1973). Nevertheless, it is possible that the intended manipulations result in unintended changes in participants' interpretations, with some degree of individual differences. The following studies are selected from a diverse set of topics related to decision-making, rule violation, second-language effects, meditation, and cooperative/punitive behavior. The studies share in common an insistence on describing the experimental manipulation and their effects from a third-person viewpoint, neglecting the possibility that the (first-person) meaning assigned by the participants might be a confounding factor and perhaps the primary explanation of the findings. Taking changes in meaning into account, as we shall see, results in a view of the research findings that fundamentally diverges from the one provided by the researchers.

The experimental-cognitive studies of rule violation have insisted on adopting a third-person perspective on participants' behavior, which requires fixing the meaning of "rule" and "rule-breaking" across different conditions, ignoring possible variations of meaning in the participants' perspective. For instance, research initiated by Pfister and colleagues examines the potential costs of rule violation on performance (Pfister et al., 2016a, b). More generally, this research aims to offer description of rule-following and rule-breaking action in terms of relatively low-level, sensorimotor characteristics of actions, although this requires specifying the task requirement at the higher level of abstraction (i.e., determining what counts as rule violation). Researchers have found, using target-directed movement tasks, that hand movements that violate a rule tend to be on average slower, both in their initiation and in their completion, and their trajectory tends to deviate from a straight path, compared to movements that conform to the rule (Pfister et al., 2016b; Wirth et al., 2016).

Wirth et al. (2018a) found that the costs associated with rule violation are eliminated if the rule violation (a) is performed frequently and (b) has been committed recently. The authors assumed that increasing the frequency of rule-violation trials and their recency does not change the meaning of the behavior with respect to the rules. This assumption would be inconsistent with the research on norms, which suggests we adopt norms rapidly and flexibly in response to the changing contexts and in response to changes in our own behaviors (Ting, 2018). If we accept that the meaning of rule (violation) changed across the conditions, then recency and frequency are confounded with meaning. Rather than stating that participants are now committing the same action ("rule-breaking") with more efficiency, we might have to state that participants are performing a different type of action ("breaking a relatively strict rule" vs. "breaking a nominal and flexible rule"). The meaning assigned to an action might change during the experiment, as an unintended outcome of experimental manipulations.

Another example highlights how tasks might be differently interpreted by different groups of participants. Kozasa et al. (2012) aimed to test the effect of long-term meditation practice on attention. They compared the performance of regular meditators and non-meditators in a Stroop task (Stroop, 1935). In a Stroop task, participants are instructed to report the color (ink) of words, one word at a time, while ignoring the meaning of the words. For each color-word stimulus, the color and the meaning of the word could be congruent (RED typed in red), incongruent (GREEN typed in red), or unrelated (SHELF typed in red). Compared to neutral trials, performance is usually better on congruent trials and worse on incongruent trials. Kozasa et al. (2012, p. 745) reasoned that the Stroop task requires "attention and impulse control" and asked whether these abilities are superior in regular meditators. Using brain imaging, the authors found increased activity on incongruent trials relative to congruent trials but only for non-meditators. The increased brain activity, associated with incongruent color-word stimuli, was not observed in regular meditators.

How do we know that the meditators and non-meditators had the same understanding of the Stroop task? Recall that Huffman et al. (2017) showed how an active approach to ignoring a distractor might increase the cost of the distractor. Could a similar and costly strategy have been adopted by the non-meditators in the Kozasa et al.'s (2012) study? If we assume that there is only one way to understand and engage with this task, or at least both groups of participants understood the task in the same manner (deploying the same set of cognitive capacities), then we could conclude that meditators were relatively more efficient than the non-meditators, in exercising a set of common capacities. If, however, the meditators took a different approach to the Stroop task and, in effect, performed a different task, then we cannot conclude that the meditators had a quantifiable increase with regard to the same capacities. For example, the regular meditators might have taken a more passive approach to the irrelevant word meaning—letting go of the irrelevant stimulus feature rather than actively suppressing it—making it unnecessary to exert cognitive control on monitoring color-word conflict. In other words, rather than enhancing the ability to cognitive control (in response to interference), meditation might promote an understanding of task performance that reduces the necessity for cognitive control (by *letting go* of the source of interference).

In social and cooperative situations, different interpretations of actions can similarly result in different outcomes. Six-year-olds have been found to intervene in an unfair interaction, even when that is costly for themselves (McAuliffe et al., 2015). When observing another child unfairly and selfishly dividing candies with a partner, children demonstrate a tendency to intervene; they refuse the allocation, thereby punishing the "selfish" allocator. We might describe the perspective of the child who can intervene as a detached third party, and this characterization might be reasonable in certain conditions. However, some manipulations might change the child's perspective. Consider, for instance, that the intervention itself can become costly (i.e., intervention could cost the child a candy). Here, the intended manipulations are described as the *fairness* of the allocation decision and the *cost* of intervention, though the two factors might interact and change the meaning assigned to the act of intervention. Choosing a costly intervention might involve the child's standpoint having shifted from a neutral third party to an "ally of the underdog," given that both are placed in positions of disadvantage. If so, what happens in the experiment might not fit the neat and clear categories imposed by the experimental design. While experimenters claim to establish a causal link between variables, the cost of intervention, and the probability of intervention, they end up instead constructing different dramaturgical scenarios (narratives) in the different conditions (Harré, 1993). If we acknowledge the differences between conditions as differences between the dramaturgical scenarios, we will not assume the same variables are at play across the conditions.

Turning to a line of research quite different from what I have discussed above, though it helps further illustrate my main point, let us now consider the so-called second-language effect on cheating behavior. The background for this line of research are studies on cheating, in which participants roll a die and report whether the outcome matched a pre-specified target number (in which case they would receive monetary reward). The die is rolled inside a cup, and no one other than the participant can see the outcome. There is, therefore, an incentive to cheat, and participants can cheat without being detected, although the rate of cheating can be estimated at the group level (Hilbig & Thielmann, 2017). Interestingly, Bereby-Meyer et al. (2018) reported that the rate of cheating is lower when people use their second language (L2) during the study session, compared to when they use their native language (L1). The authors interpreted this effect in terms of the different efficiency of information processing across different languages. Presuming that we are less efficient with L2, communicating in L2 would prolong decision-making, increasing the probability that the relatively slow and rational modes of thought dominate our decision. Alternatively, presuming that using L2 and cheating are both associated with extra cognitive effort, participants might decide to behave honestly merely to avoid the additional effort (Pfister et al. 2016a, b). These interpretations both treat cheating as the same type of action (qualitatively equivalent) across the two language conditions.

It is possible, however, that different languages evoke different norms, leading participants to assign different meanings to a dishonest expression depending on the medium of expression. Communicating in L2, particularly for novices, might be associated with a range of relationships and sociocultural positions (teachers, fellow language students, being in the out-group, etc.), distinct from those that are dominantly associated with L1. It is possible that communicating in L2 evokes situations in which dishonesty is either less accessible or evaluated more negatively. Wirth et al. (2018b) have argued that rule-breaking sensitizes participants to the concept of authority. This phenomenon might itself depend on the medium of communication. In particular, the concept of authority might be more accessible, and more easily evoked, when using L2. If so, then a dishonest expression in L1 and L2 ought to be considered as qualitatively different actions, by virtue of evoking different meanings (Bergner, 2016; Gozli, 2017). This interpretation stands in contrast to the approach that compares the probability of cheating, *as the same type of action*, across different conditions (for a review, see Gozli, 2019).

The second-language effect has been investigated on other forms of decision-making bias. The studies on the so-called framing effect show that, for the same decision, people can become more or less tolerant of risk, depending on whether their choice is presented with respect to its potential loss or its potential gains (Tversky & Kahneman, 1981). Recent studies have shown the framing effects can decrease if the choices are presented in L2, compared to when they are presented in L1 (Costa et al., 2014; Keysar et al., 2012). Keysar et al. (2012) found no evidence of a framing effect in L2. They interpreted the results in terms of a cognitive and emotional distance that comes with using L2, which in turn could result in more rational responses. Similarly, Costa et al. (2014) presented participants with moral dilemmas and found that a higher number of participants responded in a utilitarian ("rational") manner, e.g., saving five people by killing one person, when the dilemma is presented in L2.

Again, it is possible that certain heuristics—or decision "shortcuts"—are less accessible in L2. Embodied cognition perspectives are relevant for these lines of research (Barsalou, 1999; Gallese & Lakoff, 2005). The meaning of words is grounded in perceptual-motor experience and bodily affective sensations. Concepts expressed in L1 could be associated with stronger embodied correlates compared to L2. An embodied feeling of loss/gain, and its impact on decision-making, might be more easily evoked in L1; similarly, a utilitarian response to a moral dilemma might evoke more aversive connotations, or embodied meaning, when considered in L1, compared to when it is considered in L2.

A recent study by Korn et al. (2018) fits with the idea that the effect of L2 on the framing effect might be due to differences in meaning, rather than differences in cognitive effort. Korn et al. attempted to generalize the foreign-language effects in terms of the more abstract construct, such as cognitive fluency or cognitive effort. To do so, instead of presenting problems in L1 and L2, they presented them either in an easy-to-read font (fluent, low effort) or in a hard-to-read font (dysfluent, high effort). In their first two experiments, recruiting 158 and 271 participants, the authors found no effect of font type. In an online version of the experiment, recruiting a larger sample of 732, they found a small effect of font type in the expected direction: the framing effect was smaller with the hard-to-read font. The difficulty in obtaining the effect, as reported by Korn et al. (2018), lends support to my alternative account of the second-language effect. Beyond variations in cognitive fluency or efficiency, the differences between L1 and L2 might be related to variations in meaning and interpretation.

A final example comes from research comparing human-human and human-computer interaction. Tenbrink et al. (2010) compared how people gave route instructions to (1) another person and (2) to a computer program. Instructions varied more widely in the former condition, compared to the latter. In other words, when humans addressed a computer, they remained within a narrower range of expressions. While the intended manipulation in the study was the addressee type (two categories: human vs. computer program), the unintended manipulation (meaning) might have been the perceived linguistic competence of the addressee. Assuming that a higher level of linguistic competence in an addressee might lead us to use language more freely and flexibly, consequently introducing more variety in our expressions. If this explanation holds, we would expect to see a similar pattern of findings if we compare how people communicate to computer programs of different degrees of sophistication. Moreover, in human-human interactions, we would also expect differences in how people communicate to native speakers of a language and novice speakers. It would, therefore, be worthwhile to explore variations in how people interpret the context and their addressees.

Conclusion

The aim of the present chapter was to highlight the role of interpretation (on the part of research participants) in experimental research (Gozli, 2019; Toomela, 2008; Valsiner & Brinkmann, 2016). Aiming to discover causal relations among variables of interest, experimental researchers attempt to isolate variables in relatively simple laboratory settings (Gozli & Deng, 2018), in which events and behaviors are described with respect to a set of already determined categories (Gozli, 2017; Mammen, 2017). When the effect of rule-breaking is concerned, for example, researchers aim to keep all else equal across conditions, except for whether a given action reflects “conforming to” or “violating” a rule; similarly, when the effect of long-term meditation is concerned, we aim to keep all else equal, except for the prior meditation experience of the participants. Accordingly, we assume that the meaning of events does not change systematically—from the perspective of the research participants—as we introduce other manipulations. We might implicitly assume that rule-breaking preserves its meaning within the context of the task, regardless of the frequency and recency of rule-breaking. Similarly, we might assume that the meaning assigned to a distracting item (i.e., how one ought to think, and what one ought to do, about the distractor) is the same in the meditator and non-meditator groups.

My argument consisted of three stages. First, I reviewed several examples from experimental psychology, where task interpretation was explicitly manipulated—by changing the task instructions—and was found to impact performance in the task. These changes do not require changing the physical features of the task. Second, I discussed a particular dimension of meaning, namely, the normative dimension, within which actions could be identified as “good,” “bad,” “unacceptable,” “desirable,” and so forth. Research has demonstrated our ability to detect norms rapidly and flexibly, our sensitivity to the presence of others, and our sensitivity to the perspective of others in evaluating our own actions. Third, I reviewed several different lines of research (on rule-breaking, meditation, intervention by children, and second-language effects), where there are clear possibilities for the intended manipulations to be confounded with changes in participants’ interpretation of the events. If my criticisms are valid, it would mean that the experiments did not isolate the variables of interest (e.g., meditation ➔ distractor suppression) within controlled processes (i.e., one and the same task). Rather, different conditions correspond to qualitatively different tasks, which could not be neatly compared only with respect to the intended manipulations and measures.

Disentangling meaning of an event from experimental manipulations is not straightforward (Bergner, 2016). Meaning is not an isolable part of the experimental setting, but has to do with how the setting is framed. When the meaning of a given task/event changes from the point of view of the research participants, it could mean that (a) what the participants view as potentially available is different, including potential social incentives and prohibitions; (b) what the participants regard as acceptable, good, bad, and so forth, in the experimental context might be different; and (c) the presumed purpose of the experimental session might change. All these might happen despite the researchers’ belief in the validity of their prior categories of description and evaluation.

Once we recognize that our intended manipulations can change the meaning of events from the perspective of research participants, we will allow ourselves to describe differences between conditions in qualitative terms, in terms of tasks that differ in kind, and with respect to different normative standards. What else could change, for instance, when a problem is presented in L2, rather than L1? Might long-term meditators think differently about the task, which subsequently renders them differently prepared for the events in the experiment? Might a child interpret intervention differently, once we associate the act of intervention with personal cost? Pursuing such questions requires including the participants' perspective, rather than an insistence on a uniform, third-person description. Social situations, including experimental sessions that include researchers and participants, are usually open to multiple descriptions and transformation, which is to say they have depth (Mammen, 2017, 2019). What actually happens, therefore, as a result of our experimental manipulations might not fit within the pre-specified categories of description and evaluation.

Acknowledgments I would like to thank Carol Ting, Heath Matheson, and Kieran O'Doherty for reading earlier drafts of this chapter and offering helpful comments.

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577–660.

Bereby-Meyer, Y., Hayakawa, S., Shalvi, S., Corey, J., Costa, A., & Keysar, B. (2018). *Honesty speaks a second language*. Topics in Cognitive Science. Online first.

Bergner, R. M. (2010). What is descriptive psychology? An introduction. In K. Davis, F. Lubuguin, & W. Schwartz (Eds.), *Advances in descriptive psychology* (Vol. 9, pp. 325–360). Descriptive Psychology Press.

Bergner, R. M. (2016). What is behavior? And why is it not reducible to biological states of affairs? *Journal of Theoretical and Philosophical Psychology, 36*(1), 41–55.

Brinkmann, S. (2010). *Psychology as a moral science: Perspectives on normativity*. Springer.

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014). Your morals depend on language. *PLoS One, 9*(4), e94842.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory, 33*(1), 67–80.

Dreisbach, G., Goschke, T., & Haider, H. (2007). The role of task rules and stimulus–response mappings in the task switching paradigm. *Psychological Research, 71*(4), 383–392.

Engelsted, N. (2017). *Catching up with Aristotle: A journey in quest of general psychology*. Springer.

Festinger, L. (1964). *Conflict, decision, and dissonance*. Stanford University Press.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22*(3–4), 455–479.

Gozli, D. (2017). Behaviour versus performance: The veiled commitment of experimental psychology. *Theory & Psychology, 27*, 741–758.

Gozli, D. (2019). *Experimental psychology and human agency*. Springer.

Gozli, D., & Deng, W. (2018). Building blocks of psychology: On remaking the unkept promises of early schools. *Integrative Psychological and Behavioral Science, 52*, 1–24.

Harré, R. (1993). *Social being* (2nd ed.). Wiley.

Hilbig, B. E., & Thielmann, I. (2017). Does everyone have a price? On the role of payoff magnitude for ethical decision making. *Cognition, 163*, 15–25.

Hommel, B. (1993). Inverting the Simon effect by intention: Determinants of direction and extent of effects of irrelevant spatial information. *Psychological Research, 55*, 270–279.

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychologica, 136*(2), 189–202.

Hommel, B., Colzato, L. S., & van den Wildenberg, W. P. M. (2009). How social are task representations? *Psychological Science, 20*, 794–798.

Huffman, G., Rajsic, J., & Pratt, J. (2017). *Ironic capture: Top-down expectations exacerbate distraction in visual search*. Psychological Research.

Johns, G. (2006). The essential impact of context on organizational behaviour. *Academy of Management Review, 31*, 386–408.

Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science, 322*(5908), 1681–1685.

Kenward, B., Karlsson, M., & Persson, J. (2010). Over-imitation is better explained by norm learning than by distorted causal learning. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20101399.

Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science, 23*(6), 661–668.

Korn, C. W., Ries, J., Schalk, L., Oganian, Y., & Saalbach, H. (2018). A hard-to-read font reduces the framing effect in a large sample. Psychonomic Bulletin & Review, 25, 696–703.

Kozasa, E. H., Sato, J. R., Lacerda, S. S., Barreiros, M. A., Radvany, J., Russell, T. A., et al. (2012). Meditation training increases brain efficiency in an attention task. *NeuroImage, 59*, 745–749.

Mammen, J. (2008). What is a concept. *Journal of Anthropological Psychology, 19*(2), 25–27.

Mammen, J. (2017). *A new logical foundation for psychology*. Springer.

Mammen, J. (2019). *A grammar of praxis: An exposé of "a new logical Foundation for Psychology", a few additions, and replies to Alaric Kohler and Alexander Poddiakov*. Integrative Psychological and Behavioral Science. Online first.

McAuliffe, K., Jordan, J. J., & Warneken, F. (2015). Costly third-party punishment in young children. *Cognition, 134*, 1–10.

Pfister, R., Wirth, R., Schwarz, K. A., Foerster, A., Steinhauser, M., & Kunde, W. (2016a). The electrophysiological signature of deliberate rule violations. *Psychophysiology, 53*, 1870–1877.

Pfister, R., Wirth, R., Schwarz, K., Steinhauser, M., & Kunde, W. (2016b). Burdens of non-conformity: Motor execution reveals cognitive conflict during deliberate rule violations. *Cognition, 147*, 93–99.

Quesque, F., Lewkowicz, D., Delevoye-Turrell, Y. N., & Coello, Y. (2013). Effects of social intention on movement kinematics in cooperative actions. *Frontiers in Neurorobotics, 7*, 14.

Sartori, L., Becchio, C., Bara, B. G., & Castiello, U. (2009). Does the intention to communicate affect action kinematics? *Consciousness and Cognition, 18*, 766–772.

Schmidt, M. F. H., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young children see a single action and infer a social norm: Promiscuous normativity in 3-year-olds. *Psychological Science, 27*, 1360–1370.

Searle, J. R. (1995). *The construction of social reality*. Simon & Schuster.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition, 88*(3), B11–B21.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective* (pp. 31–86). Elsevier.

Simons-Morton, B. G., Bingham, R. C., Falk, E. B., Li, K., Pradhan, A. K., Ouimet, M., Almani, F., & Shope, J. T. (2014). Experimental effects of injunctive norms on simulated risky driving among teenage males. *Health Psychology, 33*(7), 616–627. https://doi.org/10.1037/a0034837

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662.

Székely, M., & Michael, J. (2018). Investing in commitment: Persistence in a joint action is enhanced by the perception of a partner's effort. *Cognition, 174*, 37–42.

Tenbrink, T., Ross, R. J., Thomas, K. E., Dethlefs, N., & Andonova, E. (2010). Route instructions in map-based human–human and human–computer dialogue: A comparative analysis. *Journal of Visual Languages & Computing, 21*(5), 292–309.

Ting, C. (2018). *The feedback loop of rule-breaking: Experimental evidence*. The Social Science Journal. Online first.

Toomela, A. (2008). Variables in psychology: A critique of quantitative psychology. *Integrative Psychological and Behavioral Science, 42*(3), 245–265.

Toomela, A. (2016). What are higher psychological functions? *Integrative Psychological and Behavioral Science, 50*(1), 91–121.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.
Valsiner, J., & Brinkmann, S. (2016). Beyond the "variables": Developing metalanguage for psychology. In S. H. Klempe & R. Smith (Eds.), *Centrality of history for theory construction in psychology* (pp. 75–90). Springer.
Wachtel, P. L. (1973). Psychodynamics, behavior therapy, and the implacable experimenter: An inquiry into the consistency of personality. *Journal of Abnormal Psychology, 82*(2), 324–334.
Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1528), 2417–2428.
Wirth, R., Pfister, R., Foerster, A., Huestegge, L., & Kunde, W. (2016). Pushing the rules: Effects and aftereffects of deliberate rule violations. *Psychological Research, 80*(5), 838–852.
Wirth, R., Foerster, A., Herbort, O., Kunde, W., & Pfister, R. (2018a). This is how to be a rule breaker. *Advances in Cognitive Psychology, 14*(1), 21–37.
Wirth, R., Foerster, A., Rendel, H., Kunde, W., & Pfister, R. (2018b). Rule-violations sensitise towards negative and authority-related stimuli. *Cognition and Emotion, 32*(3), 480–493.

Chapter 8
Methodology of Science: Different Kinds of Questions Require Different Methods

Aaro Toomela

This chapter is not about *methods* of psychology in particular or of science in general. Rather, it is about *methodology* of science. This topic is rather unusual in psychology today. To set the stage for the following discussion, I think it is useful first to take a look into the notion of methodology.

There Is Methodology and There Is Methodology

Methodology is an uncommon topic in and for the psychology today. Superficially, the opposite may seem to be true: there are many books written on what is called "methodology" of psychology or social-behavioural sciences. But let us see what "methodology" is according to such books. I bring just one example from an introductory book on experimental psychology:

> We will study **methodology**, the scientific techniques used to collect and evaluate psychological **data** (the facts and figures gathered in research studies). (Myers & Hansen, 2002, p. 3, emphasis in original)

So, methodology is the study of scientific techniques, the ways data are collected and evaluated. If this is methodology, then what is method? This term was used already by ancient Greeks. In Greek, *methodos* (μέθοδος [< μετά, ὁδός]) was understood as:

> following after, pursuit [...] — hence, II. pursuit of knowledge, investigation [...] 2. mode of prosecuting such inquiry, method, system (Liddell et al., 1940b, p. 1091)

A. Toomela (✉)
Tallinn University, School of Natural Sciences and Health, Tallinn, Estonia
e-mail: aaro@tlu.ee

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9_8

Method, thus, is a mode of prosecuting an investigation. Methodology, as can be seen from the first quote above, seems to be the same, it is knowledge of scientific techniques, the modes of prosecuting investigations. So, method and methodology seem to refer to the same thing for psychology and related sciences today. And it is so, in many textbooks on research methods in psychology, the methods are described either with no mentioning of methodology or calling such recipe books "methodology". Methodology of mainstream psychology today has turned into a kind of toolbox, a predetermined set of techniques (e.g. Toomela, 2009a; Valsiner, 2017).

Maybe this is not a problem? Maybe scientists have developed wonderful methods that allow to answer all the important questions, and if some new method will be created, such as a new statistical data analysis technique, then it will be just added to the recipe book of scientific methods? There are several – and very fundamental – problems with such a recipe book approach, however.

First of all, scientific methods are not just neutral, in terms of a theory, intermediate steps between a hypothesis and scientific knowledge. Even if to separate methods from theories, the methods may become a ground to develop new theories. In other words, theories about whatever is studied, mind, for instance, will be shaped according to the tools accepted by a scientific community, the accepted methods of research (Gigerenzer, 1991, 1992).

If theories are shaped by research methods, as Gigerenzer has convincingly demonstrated, then science eventually ceases to be able to create new and more advanced understanding of the world. Russian philosopher of science, Kohanovski, made a relevant observation:

> Every method turns out to be ineffective and even useless if it is not used as a "guiding thread" in a scientific or other form of activity, but as a ready-made template for reshaping facts. The main purpose of any method is – on the basis of corresponding principles (requirements, prescriptions, etc.), to ensure the successful solution of certain cognitive and practical problems, an increase in knowledge, the optimal functioning and development of certain objects. (Kohanovski et al., 2003, p. 302)

So, methods are used to solve certain (scientific) problems, they are not solutions in themselves and they, in themselves, are also not sources of novel questions, novel perspectives on the studied things and phenomena. It is actually quite absurd to study methods as things in themselves in order to create theories about whatever it is that is intended to be understood better and not to study the thing or phenomenon itself. Study of microscope itself, outside of the context where it is used, does not tell anything about what, say, bacteria are. And if we turn microscopes to the sky, we will see something, but that something would definitely add nothing to understanding astronomy. The same way, we can endlessly study statistical distributions and sophisticated ways to manipulate numbers, and yet all that knowledge does not contain anything about the essence or functioning of the human mind.

The science develops not through studying methods, and yet it is novel methods that allow to construct novel knowledge, as the outstanding Russian physicist, Pyotr Kapitsa, has observed:

> As it is known, the development of science consists in finding new natural phenomena and in the discovery of those laws which they obey. Most often this is achieved due to the fact that new methods of research are found. (Kapitsa, 1977, p. 249)

Methods, thus, are important, but the science advances because new methods are created and not because already existing methods are turned into theories. So far, I have discussed methods as something separate from a theory the methods are used to advance. However, such separation, so common in mainstream psychology today, is clearly wrong. Prigogine has made an interesting observation:

> We come to problems where methodology cannot be separated from the question of the nature of the object investigated. (Prigogine & Stenger, 1984, p. 204)

Prigogine, thus, suggests that methodology is somehow related to the theory of the studied thing or phenomenon – because in science, as a rule, theories are created to define what might be the nature of the studied objects. Kohanovski has been more specific on this issue:

> Thus, theory and method are simultaneously identical and different. Their similarity lies in the fact that they are interconnected, and in their unity there is an analogue, a reflection of reality. Being united in their interaction, theory and method are not rigidly separated from each other and at the same time are not directly one and the same. [...] The main differences between theory and method are as follows: a) theory is the result of previous activity, method is the starting point and prerequisite for subsequent activity; b) the main functions of the theory are explanation and prediction (with the aim of finding truth, laws, reasons, etc.), the method is the regulation and orientation of activity; c) theory – a system of ideal images that reflect the essence, laws of the object, method – a system of regulations, rules, prescriptions that act as a tool for further cognition and change of reality; d) the theory is aimed at solving the problem – answering the question, what is the given subject; the method – at identifying the ways and mechanisms of its research and transformation. (Kohanovski et al., 2003, p. 305)

Indeed, a method for studying something and the theory about that same thing cannot be separated in principle. This is because methods must allow to study what is intended to study; methods are essentially procedures that transform directly non-observable events into observable (see below a few more thoughts on the idea of directly non-observable and its relation to science). Therefore, all methods are based on a theory of how the studied thing or phenomenon interacts with its environment so that the consequences of that interaction become observable. In that sense, theory of a method is always part of the theory of what is studied using the method. And yet, method is not a theory, method must be based on theory, but in itself it is a way of cognizing a technique for study.

The concept of methodology must be put into this context. Above I showed that methodology may mean the study of methods. But the term methodology has two different meanings:

> The concept of "methodology" has two main meanings: a system of certain methods and techniques used in a particular field of activity (in science, politics, art, etc.); teaching about this system, general theory of method, theory in action. (Kohanovski et al., 2003, p. 300; see for a similar distinction, Yedronova & Obcharov, 2013)

Study of methods in the first sense of methodology is obviously necessary. No research can succeed, or appreciate its own achievements, without knowing how to apply methods, and this is exactly what methodology in this sense aims to achieve. It is also the only meaning that methodology has in mainstream psychology today.

However, there is far more important question to ask – whether a certain method should be used at all. This question cannot be answered by studying methods alone; it can be only answered if a method is studied in the context of the theory about what is studied. It must be demonstrated that the selected methods allow to answer the research questions; it must be demonstrated how properties of the studied thing can, theoretically, become observable through the use of the method. This kind of methodology is practically nonexistent in psychology and related sciences today. The situation is quite curious in these fields of knowledge. There is a small group of scholars, who ask methodological questions and demonstrate that methods used in psychology today do not in principle allow to answer the questions asked (e.g. Gozli, 2019; Lindstad et al., 2020; Logie, 2018; Michell, 2004, 2012a, b; Molenaar, 2004a, b, 2007, 2008; Smedslund, 1988; Toomela, 2007a, 2008a, 2009a, b, 2010a, d, 2011, 2014a, 2019; Toomela & Valsiner, 2010; Valsiner, 2017). The majority of researchers, however, do not even ask such questions. And this is despite the fact that there is no theoretical-methodological justification that methods used in psychology are valid, that they allow to study what is supposed to be studied and that they allow to achieve understanding of what is studied.

In the following discussion, I use the term "methodology" to signify the general theory of method and not a toolbox kind of methodology, a list and description of the application of particular ways to conduct studies.

Some Definitions

Discussing methodology means asking what kinds of methods could be used in science (of psyche) and what kinds of questions the selected methods allow us to answer. The possible area of methodological studies would be very wide. Methodology is relevant both generally and about each and every study planned to conduct in particular. I think that methodology can be constructed hierarchically: general principles can be formulated that apply to all lower levels of analysis until the single studies. Even though mainstream psychology today does not bother itself with such questions, methodology of science at the general level of analysis has been discussed by many scholars over many centuries, beginning with Greek philosophers more than two millennia ago. Several interesting works – all, of course, very distant from the mainstream – have added new perspectives to methodological discussions recently (e.g. Branco & Valsiner, 1997; Gozli, 2019; Haig, 2014; Valsiner, 2017).

One way to proceed would be to make a summary of the state of the general methodology today and proceed to lower levels of analysis by applying the principles to particular fields of studies. But it seems to me that there are aspects of methodology that can be developed further at the general level. If I am correct – the reader will decide whether it is so after reading this chapter – then it is not time yet to discuss particulars without further elaborating general principles that might apply to all particular studies. Generally, different methods suit to answer different questions. As far as I know, there is one perspective from which methodology of science is not discussed in sufficient details – there are different *general kinds* of scientific questions. Each kind of questions requires specific kind of methods to answer them. I am going to distinguish such kinds of scientific questions and propose which kind of methods is suitable to answer which kind of questions.

Before going into particulars, some phenomena must be defined. I am going to discuss methods of science. Science, however, is defined in several different ways. It is necessary to define what it is I call science and to discuss why science should be distinguished as a special way of knowledge-seeking. Next, what scientific explanation is must be defined. The reason why it is necessary is clear also: there are different theories as to what comprises scientific explanation. Depending on how scientific explanation is understood, different kinds of questions are considered to be scientific. If questions are different, then the ways to answer them, the methods, must be different too. Further, it is not self-evident why science needs method(s) at all. These questions, I think, must be answered also before going into methodology.

What Is Science?

Science has been defined in many different ways. I do not see a reason to delve deep into this topic by providing different definitions, comparing them and either selecting one or ending up with understanding that a definition with some novel aspects is needed for providing a coherent account of the relations between science, its methods and methodology. Instead, I propose a definition that, as far as I know, has a novel (or at least uncommon) aspect I find very important to take into account for better understanding the human pursuit of understanding the world as it is. So, the phenomenon I am discussing, science, I define as follows:

> *Science is a body of knowledge about structural-systemic causes of (mostly nonsensory) things and phenomena, which is constructed with methodologically grounded methods.*

Next, I discuss shortly each of the components of the definition.

Science Is Knowledge

First, science is *knowledge*. Here I follow the long tradition to define science, *scientia*, as "a clear and certain knowledge of any thing" (Chambers, 1728b, p. S32). Roots of using the term science to signify knowledge go back to Greek philosophy, where, among others, Plato and Aristotle used a term *episteme* (ἐπιστήμη) to signify (scientific) knowledge or science (Liddell et al., 1940a, p. 660).

Knowledge of Causes

Historical tradition, however, is not the only, or even the main, reason for defining science as a kind of knowledge. Science could also be defined as an activity, pursuit for knowledge.[1] Any purposeful activity, including pursuit of knowledge, can be successful only if its aim is clearly defined. If we have not explicitly defined, what we want to achieve through sciencing,[2] there is no way to define methods as ways towards achieving what we aim at and there is no way to realize, whether expected result was achieved. The aim of sciencing is scientific *knowledge*. This is why I define science as a body of knowledge.

Not any kind of knowledge, however, can be considered scientific. In the definition of science I proposed that scientific knowledge is defined as *knowledge about [...] causes*. This idea also has been proposed at the dawn of modern science by Aristotle:

> We suppose ourselves to possess unqualified scientific knowledge of a thing, as opposed to knowing it in the accidental way in which the sophist knows, when we think that we know

---

[1] As I mentioned, there are many definitions of "science". For example, science has been defined recently as "the pursuit and application of knowledge and understanding of the natural and social world following a systematic methodology based on evidence" (Science Council, 2009). According to this definition, science is an activity of pursuit and application – two very different kinds of activities that require very different principles – but it is not clear whether the pursued and applied knowledge is of a special kind or just any knowledge that emerges following "systematic methodology", which, according to them, includes objective observation (measurement and data), evidence, experiment and/or observation as benchmarks for testing hypotheses, induction, repetition, critical analysis and verification and testing. Leaving aside that components of methodology in this list are clearly overlapping, which makes the list hard to understand and also hard to apply, it is obvious that methodology is understood by them in the "toolbox of methods" way (as I discussed above). I suggest no science as an activity can provide satisfactory knowledge and understanding of whatever is studied relying only on toolbox methodology. In addition, in this definition there is no restriction on the nature of knowledge that is pursued and applied; anything that emerges when toolbox methods are used counts as scientifically acceptable knowledge. Toolboxes, however, allow to create a lot of meaningless "knowledge", as mainstream psychology (as well as other social-behavioural sciences) clearly demonstrates. In science, as I understand it, methods need true methodology, and both of them are constrained by what can be considered to be scientific knowledge.

[2] Anthropologist Leslie Alvin White suggested that the word "science" can be used both as a noun and as a verb: "Science is not merely a collection of facts and formulas. It is pre-eminently a way of dealing with experience. The word may be appropriately used as a verb: one *sciences*, i.e., deals with experience according to certain assumptions and with certain techniques" (White, 1949, p. 3). I partly disagree with White: it is not sufficient to define sciencing as just an activity that is conducted "according to certain assumptions and with certain techniques". Assumptions and techniques are needed, but these must be selected on the ground of certain principles. And these principles follow from what is expected to achieve by sciencing, i.e. by what is considered to be scientific knowledge that emerges in sciencing. So, if to define sciencing as a pursuit of scientific knowledge, "science" as a verb is fully meaningful and can be coherently interpreted. This, of course, requires definition of "scientific knowledge" – this I have already done in other publications (e.g. Toomela, 2019), and I am going to remind this definition also later in this chapter.

> the cause on which the fact depends, as the cause of that fact and of no other, and, further, that the fact could not be other than it is. (Aristotle, 1941b, 71b9–12, p. 111)[3]

It might seem to be obvious to many that all kinds of knowledge should not be considered scientific. I suggested in Footnote 1 that it is not sufficient to define scientific knowledge only by the methods through which it is constructed. However, large amount of knowledge is called scientific only for this reason. It is certainly so that certain methods are absolutely necessary to construct scientific knowledge, but not all knowledge achieved by using such methods is necessarily scientific. For example, testing of a priori and necessarily true propositions empirically provides no useful knowledge; such studies are pseudoempirical, and results of them are essentially known before the study is conducted. Pseudoempirical studies in psychology are very common, for instance, the studies where questionnaires are used or studies of developmental sequences often belong to this category (Smedslund, 1991). Knowledge achieved through pseudoempirical studies is not scientific.

It is also important that scientific knowledge has been defined differently; not all definitions define scientific knowledge as knowledge about causes of things or phenomena. For example, sometimes scientific knowledge is related to prediction or foresight (cf. Chang, 2017). There are two reasons why I think prediction is necessarily characteristic of scientific knowledge, but not all predictive knowledge is science. Both reasons can be found in the following example. For more than a year, every day around noon, I have been putting seeds in a birdhouse in my garden to feed wild birds. The birds learnt in a few days that food is available in this birdhouse. Moreover, in a week or two, they also learnt *when* the food becomes available. So they are waiting for food to come every day I go out. They can predict the food coming with quite a high accuracy.

I think the first reason why scientific knowledge should not be constrained to prediction must be obvious already: prediction needs no special method of discovery. *Every* living creature is able to predict some future states of their environment; otherwise staying alive would be impossible (Anokhin, 1974, 1978; Toomela, 2020a).

Second, in principle two kinds of predictive knowledge can be distinguished: one includes causal knowledge and the other does not. Clearly it is possible to predict future events without knowing causal relationships between a cause and the predicted effect. Birds in my example have no idea why the food appears every day – it is just a fact of their life. Limits of such predictive knowledge are, I suppose, obvious also: there is no way to understand the situation if the – very highly

[3] It has been hard to translate Aristotle's surviving works. This passage is not an exception; there are different translations, both similar to the one I provided (e.g. Angioni, 2016, p. 140) and quite different. In one of the most highly regarded translation, there is no mentioning of knowledge about causes in the same passage; instead of "cause", the term "explanation" is used (Aristotle, 1984c, p. 115). Whatever can be the "correct" translation in this particular case, the version I provided not only makes the best sense in the context of my discussion, but it also fits very well with Aristotelian theory of knowledge, understanding and explanation in general as well as with his theory of causality in particular (see Toomela, 2019, for a thorough discussion of Aristotelian epistemology).

reliably![4] – predictable event does not appear one day. It has happened a few times that I have been late with bringing food. The birds just disappear if the food is not there in the expected time. They are not looking through the windows to see, whether I am at home – perhaps I am sleeping (as it has happened), and the food will appear later. Or perhaps I am travelling and the food will not "come" before I will be back from travel. Or perhaps I forgot to put the food, and the situation would be solved by reminding me that the birds did not get the food – just knocking on the window of the room where I am located could bring a solution. But the birds do nothing like that; they just fly away if the food does not appear in about half an hour after the expected time.

So, prediction based on causal knowledge is considerably more effective. But it is also much harder to obtain. In many cases, as I am going to show below, causal knowledge can be constructed only if justified methods of knowledge construction have been used and causal theories are created. In other words, in many cases, causes can be discovered only by sciencing.

Different Theories of Causality

The next concept in the definition of science is that of *structural-systemic*[5] causality: by my definition only a special kind of causal knowledge, *knowledge about structural-systemic causes*, is scientific. This topic I have discussed in many more details elsewhere (Toomela, 2019). So I mention here just the most important ideas related to this issue. First, there are several different definitions of the term "cause" (e.g. Chambers, 1728a, pp. 175–176). Today, causality is predominantly understood as a relationship between cause and effect where a cause is an event that somehow produces an effect, another event. Without causes the effects would not come into existence, whereas the causes themselves can exist without effects. In such

[4] During the last year, the prediction that the food will be available every day about noon was about 95% correct. And the prediction that there will be new amount of food every day appearing during a daytime was 100% correct.

[5] I call this theory structural-systemic to avoid confusion with terms. The problem is that both terms, "structure" and "system", have been defined in several different ways. The former, for instance, is sometimes equalled to atomism, and both terms have been defined also as mathematical concepts. So, mathematical structure is a set of abstract entities and relations between them, and system can be a set of interrelated variables. Mathematical expressions can be composed of abstract entities or variables. But mathematics can describe only very limited aspects of the world. Among other limitations, mathematics cannot describe, what a thing *is*. And this is one of the most important aims of sciencing – to answer the "what is?" questions. Genes are not composed of abstract entities or variables, they are composed of nucleotides; organisms are composed of organs and societies of living creatures. No mathematics is needed to understand how gene grounds protein synthesis; the same applies to most of the world what we do understand. If to exaggerate a little, I would say that mathematics comes when there is no understanding achieved yet. At least in biology, psychology and social sciences. So, I use the term structural-systemic to stress that one specific theory is signified, the one I define also below.

cause-effect relationships, it is assumed that causes, after having their effect, remain what they were before creating an effect. This kind of causality is usually called "efficient". Another often mentioned theory of causality, which usually is not followed, was formulated by Aristotle, who distinguished four complementary causes: material[6] (material from which something is coming into being, also parts of a thing), formal (the whole, synthesis or essence of a thing), efficient (that from which the change begins) and final (for the sake of which a thing is). Aristotle also explained that the last three kinds of causes often coincide (see also Aristotle, 1941a, 1984a, 1984b).

If there is more than one theory of causality, it should be explicitly justified why one should be preferred over the other(s). It is interesting, *why* Aristotle's complex theory was replaced with a primitive theory where only efficient causality is retained. I have identified only two justified explanations why to prefer the more primitive theory. One was given by Descartes; he demonstrated that only efficient causality should be considered because there is omnipotent and omnipresent God. The other was David Hume, who suggested that other kinds of causes would not be knowable to humans. Many scholars today rely on primitive theory of causality – and, together with it, on a highly limited understanding of the essence of scientific knowledge – without realizing that the reasons behind such a limited view should not be acceptable for sciences today (for an extensive discussion, see Toomela, 2019).

Aristotle's theory of causality should not be accepted without modification. But only slight changes are needed to arrive at the understanding that comprises, in my opinion, the essence of scientific explanation. I proposed that scientific understanding is achieved, when three questions are simultaneously – and coherently – answered: (1) what are the parts or elements of the phenomenon or thing under study?, (2) in which specific relations these parts are in the studied whole? and (3) what is the whole that emerged in the synthesis, what qualities characterize it?

I admit that claiming one kind of knowledge – structural-systemic – being somehow more advanced or better than other kinds must be well justified. I have provided several reasons to prefer this epistemology elsewhere (Toomela, 2019), but perhaps one reason could be especially pointed out. Structural-systemic theory is the only kind of theory that contains knowledge about how to make a thing or phenomenon under study and also to understand what kind of knowledge to look for if expected state of affairs does not appear – if a thing is not functioning as expected or a phenomenon does not appear as expected.

Furthermore, a lot of scientific knowledge in physics and biology, I even dare to suggest that all knowledge where a thing or phenomenon is really explained, is exactly structural-systemic. Mathematical formulas explain nothing, they are just exact descriptions – nothing happens or is in the form it is *because* it behaves according to a certain formula. But knowing the elements and relations of a whole is an explanation. Take, for example, synthetic biology, where it is understood that

[6] The names for different causes – material, formal, efficient and final – were created later; Aristotle did not use these terms.

"you can only understand things if you can make them" (Gross, 2011, p. R614; see also Ruiz-Mirazo & Moreno, 2013). We can make something only when we have an idea what elements should be put in which relations in order to get that something we want to understand. It was the structural-systemic understanding of genes that allowed to create an artificial genome for *Escherichia coli* (Fredens et al., 2019), and it is the same kind of knowledge, which growing will ground one day a synthesis of a living creature (cf. Powell, 2018). And it is also structural-systemic knowledge that is necessary to understand function in structural biology – "structure is function" is an unofficial motto of this field of science (Callaway, 2015, 2020). For example, in order to understand the mechanism of touch, it is necessary to reveal a 3D shape of a certain protein; this shape is determined by the atoms and their spatial relationships (cf. Dance, 2020; Mccleskey, 2019).

Structural-systemic theories can be found also in psychology. Vygotsky-Luria's theory of brain-psyche relationships (Luria, 1969, 1973, 2002; Vygotsky, 1960, 1982c) is not only the best theory to understand elements of psyche; it is also a very practical (and efficient!) theory about how to restore higher psychical functions after brain damage (Luria, 1947, 1948; Tsvetkova, 1985). I think also that the structural-systemic theory is the only way to understand how and why language is the ("material") cause of the uniqueness of the human mind as Vygotsky proposed (Toomela, 2020a) and how and by which mechanisms the mind develops over a hierarchical series of stages (Toomela, 2017; see also Toomela, 2003, for a very short early account of the same theory in English).

Knowledge About Nonsensory World

The next aspect of science as I understand it refers to the fact that science is *mostly about nonsensory things and phenomena*. Indeed, humankind managed without science through most of its history. And when science emerged, it emerged to explain the world that is not available for the senses, the nonsensory world. Today, the situation is especially clear – the majority of sciences study the world that is either too small, too big, too distant to be grasped by the senses or not available for the senses at all. The world available for the senses cannot be separated from the nonsensory world; so science also deals with the things and phenomena available for the senses. But the necessity for science emerges exactly because of the partial nonoverlap and noncontiguity of these two worlds. Karl Marx has expressed this idea very well:

> [...] all science would be superfluous if the form of appearance of things directly coincided with their essence. (Marx, 1981, p. 956)

And, in another work where he discussed the nature of profits:

> If you cannot explain profit upon this supposition, you cannot explain it at all. This seems paradox and contrary to everyday observation. It is also paradox that the earth moves round the sun, and that water consists of two highly inflammable gases. Scientific truth is always paradox, if judged by everyday experience, which catches only the delusive appearance of things. (Marx, 1985, p. 127)

Thus, here is the reason why understanding the world available for senses is impossible without some understanding of the nonsensory reality of the same world: things and phenomena that are identical to the senses may be different in aspects that are not available for the senses and vice versa. Things and phenomena that are different for the senses may share nonsensory qualities.

Scientific Knowledge Is Constructed

Not only science but also knowledge has to be constructed. World can be experienced only through the senses. This fact is of utmost importance for understanding science and sciencing. Namely, the sensory organs transform a very limited number of qualities in the physical world into neural signals. They also segregate these physical qualities into separate channels. In this way, the experienced organized world becomes a set of independent sensory attributes where the organization of the sensed world is lost. This organization can be recreated in the process of mental development. Newborns and also children in their first months of life do not distinguish even objects in their world; distinction of objects from their ground needs to be learnt (see Toomela, 2017, for a detailed theory of mental development). It follows that all organization of the sensory world is constantly recreated in the continuous flow of sensory experiences in the interaction with learnt experiences.

Another important fact is that thinking, i.e. internal organization of experiences (Vygotsky, 1926), can be conducted in qualitatively different ways. For instance, purely sensory-based forms of thinking do not allow even to realize that there is a nonsensory world – this becomes possible only with the emergence of the complex forms of semiotically mediated thinking, where language became a part of psychic processes (Toomela, 2017, 2020a). Even more developed forms of semiotically mediated thinking are needed to construct valid and reliable knowledge about the nonsensory world (see, on the development of scientific thinking, Toomela, 2008b, 2010b, 2015).

This is why understanding science as constructed knowledge is so important: if there are qualitatively different forms of thinking, those forms that are appropriate for science must be consciously distinguished from other forms of thinking. In that sense the choice of the form of thinking for scientific conclusions is part of the scientific method – and the theory of thinking development is in this context part of methodology.

Science Is Based on Method

Science, a body of knowledge, is constructed with certain *methods*. Everybody, every single being with psyche,[7] knows something. So, knowledge construction is a common phenomenon in the living world, even the simplest beings with psyche; for instance, bees (Solvi et al., 2020) or nematodes (Ardiel & Rankin, 2010) are able to construct knowledge. Then why worry about methods and constrain scientific knowledge only to that achieved by certain ways? Already Aristotle made an interesting remark:

> It is difficult to be aware of whether one knows or not. For it is difficult to be aware of whether we know from the principles of a thing or not – and that is what knowing is. (Aristotle, 1984c, 76a26–28, p. 124)

So – it is not always easy to be aware of how we know, do we know something because it follows from certain principles – this is scientific knowledge (*episteme*), or we actually do not know, we may have just an opinion (*doxa*) or credible/reputable opinion (*endoxa*) (see also Aristotle, 1984d, 100a20–100b25, p. 167). In other words, we must know how our knowledge is justified – otherwise it is not science. Still, why methods are needed for science? The answer was given already above, with a quote by Marx: the form of appearance of things does not directly coincide with their essence. More specifically, as was noted already long before Marx, the problem lies in the noncorrespondence of sense-based appearances and the underlying nonsensory realities. Isaac Watts, probably the first scholar to use the term "science" in the meaning that is close to that of today (Barnhart, 1988), observed:

> There are several Things that make it very necessary that our Reason should have some Assistance in the Exercise or *Use* of it. The first is, the *Depth and Difficulty of many Truths*, and the *Weakness of our Reason to see far into Things at once, and penetrate to the Bottom of them*. It was a Saying among the Antients, *Veritas in Puteo*, Truth lyes in a Well [...] Another Thing that makes it necessary for our Reason to have some Assistance given it, is the *Disguise and false Colours in which many Things appear to us in this present imperfect State*: There are a thousand Things which are not in reality what they appear to be, and that both in the *natural* and the *moral* World: So the *Sun* appears to be flat as a Plate of Silver, and to be less than twelve Inches in Diameter; the *Moon* appears to be as big as the *Sun*, and the *Rainbow* appears to be a large substantial Arch in the Sky; all which are in reality gross Falshoods. [...] Besides, our *reasoning Powers* need some Assistance, because they are so frail and fallible in the present State; we are imposed upon *at home* as well as *abroad*; we are deceived by our *Senses*, by our *Imaginations*, by our *Passions* and *Appetites*; by the *Authority* of Men, by *Education* and *Custom*, &c. and we are led into frequent Errors, by judging according to these false and flattering Principles, rather than according to the Nature of Things. (Watts, 1726, pp. 2–3)

[7] "Psyche is a specifically organized form of living matter. Its purposeful behaviour in anticipating environmental changes that are harmful for itself as a whole is based on individual experience" (Toomela, 2020a). Individual experiences are about the environment; individual experiences become knowledge when they are organized in thought and stored by memory processes.

So, our reason is not reliable, our senses are deceptive and authority of men often leads us astray. The same problems were recognized together with a solution – scientific method – by one of the founders of modern science, Francis Bacon, already a century before Watts:

> [...] we place the foundations of the science deeper and lay them lower [...] than men have ever done before, subjecting them to examination, while *ordinary logic accepts them on the basis of others' belief*. For logicians borrow (if I may put it in this way) the principles of the sciences from the particular sciences themselves; then they pay respect to the first notions of the mind; *finally they are happy with the immediate perceptions of the healthy senses.* [...] As for the first notions of intellect: not one of the things which the intellect has accumulated by itself escapes our suspicion, and we do not confirm them without submitting them to a new trial and a verdict given in accordance with it. Furthermore, we have many ways of scrutinizing the information of the senses themselves. For the senses often deceive [...] *The senses are defective in two ways: they may fail us altogether or they may deceive.* [...] So to meet these defects, we have sought and gathered [...] assistants to the senses, so as to provide substitutes in the case of total failure and correction in the case of distortion. We do this not so much with instruments as with experiments. [...] we do not rely very much on the immediate and proper perception of the senses, but *we bring the matter to the point that the senses judge only of the experiment, the experiment judges of the thing.* (Bacon, 2000, pp. 17–18, my emphasis)

Altogether, the reasons why method is absolutely necessary for science are clear. On the one hand, method is needed to overcome limits of the senses, which occasionally fail us altogether and occasionally "just" deceive. Nonsensory world is not fully expressed in direct sensory-based perception. On the other hand, the ways we think can be inappropriate for constructing scientific knowledge also. We must be able to become aware whether the knowledge we have achieved is properly justified or not.

Scientific Methods Require Methodology

Finally, I propose that science is based only on methodologically justified methods. It does not follow that knowledge achieved through nonjustified methods is necessarily nonscientific in content. Such knowledge is just not sufficiently grounded with arguments from all relevant perspectives and therefore also unreliable. Methodology would not be necessary if there were only one general kind of questions to be answered and only one general kind of scientific methods that can be used to answer these questions. However, there are, as I am going to show below, different kinds of questions that must be answered on the way to construct science, and each of these kinds of questions also requires a different kind of methods. If there are different methods, then it is absolutely necessary to assess methodologically every particular method that is intended to use in research. Otherwise, there is a risk that the selected methods do not allow to answer the question raised.

Two Kinds of Methodological Questions

Methodology can be approached from two perspectives. These perspectives were well distinguished by Vygotsky:

> We see, in this way, that scientific study is simultaneously a study of a fact and a study of the way of cognition of the fact; in other words – we see that methodological work is carried through in the science itself, as much as it moves forward or comprehends its conclusions. (Vygotsky, 1982a, p. 368)

So, on the one hand, there is a way to study a fact, and on the other hand, there is a way of the cognition of the fact. Method always includes both aspects. I think it is feasible to look into this distinction a little deeper. In the beginning of this chapter, I mentioned that methodology – a theory of method – cannot be separated from the theory of what is studied. So, from that perspective, methodology is part of specific theories of what is studied – methodologies must be different for physics, for biology, for psychology, etc. Furthermore, methodologies must be even more specific, down to the particular things and phenomena that are attempted to understand scientifically. This is so because the method of study is a way to arrive from what is studied to manifestation, to some event that can be sensed, be it some reading of a sensor, observed behaviour or any other directly observable event. Thus, method can lead to understanding only when it is understood how the studied thing is going to come into relationships with study conditions, conditions which are created according to the method.

Methodology, however, can also be approached from the opposite direction, from the observer's-scientist's perspective. This is so because the way the scientist interprets the studies is part of the method and, therefore, also belongs to the realm of methodology. Indeed, this was the reason why already Francis Bacon discussed the same topic in his work on methodology of science, *The New Organon*:

> There are four kinds of *illusions* [AT: *"idols"*] which block men's minds. [...] The *idols of the tribe* are founded in human nature itself and in the very tribe or race of mankind. The assertion that the human senses are the measure of things is false; to the contrary, all perceptions, both of sense and mind, are relative to man, not to the universe. [...] The *idols of the cave* are the illusions of the individual man. [...] each man has a kind of individual cave or cavern which fragments and distorts the light of nature. [...] There are also *illusions* which seem to arise by agreement and from men's association with each other, which we call *idols of the marketplace;* we take the name from human exchange and community. Men associate through talk; and words are chosen to suit the understanding of the common people. [...] Plainly words do violence to the understanding, and confuse everything; and betray men into countless empty disputes and fictions. [...] Finally there are the *illusions* which have made their homes in men's minds from the various dogmas of different philosophies, and even from mistaken rules of demonstration. These I call *idols of the theatre*, for all the philosophies that men have learned or devised are, in our opinion, so many plays produced and performed which have created false and fictitious worlds. (Bacon, 2000, Bk I: XXXIX-XLIV; pp. 40–42)

For Bacon, methods were needed exactly to overcome limits, the "idols" of the human mind. So, methodology must take into account also possible limitations of a scientist. In the following discussion, I am approaching methodology from the latter, the scientist's perspective. This perspective allows to approach methodology at the level that is universal for all fields of science as the obstacles on the way to construct science are the same independently of what is studied, from the quantum particles to complex social-cultural phenomena.

Methodology Today and the Role of a Question in Sciencing

I am not going to discuss the state of methodology today in details not only because this topic would require much more than what can be done in a book chapter but also because I do not think it is necessary. I propose a look to science and its methodology that has novel aspects. As all elements of a system change when they are synthesized into a higher-order whole, the ideas that I am taking from the past – and there are many of them – are not the same as they would be in another theoretical context. So, I discuss only a few main points of disagreement and then proceed to methodological questions in the context of the theory of science as I defined it.

I think it is important to pay attention to the role of the question in sciencing. Science can begin only with a question. And it can proceed only with asking further and further questions. Here some scholars may have some doubt – it might not be so obvious that science begins with a question. This issue is directly related to the essence of scientific observation. We can learn that observation is, for instance, "the systematic noting and recording of events" (Myers & Hansen, 2002, p. 15). In order to avoid confusion in interpreting this definition, the same authors also make clear what they mean by "systematic"; it signifies a certain procedure that must be followed in observation: "once the researcher has devised a system for observing, the same system must be applied consistently to each observation" (ibid., p. 15). All this may seem to be coherent and sufficient for sciencing. Yet I did not find in the referred book (and several other books on the methods or "methodology" of psychology or social sciences) any mentioning that observations are always selective – it is not a minor issue, and it is a central idea to understand any observation. In this book, closest to the idea I have in mind was a statement about challenges related to "naturalistic observation" where the authors mentioned: "Deciding who and when to observe and what to record and analyse draws heavily on both the researcher's judgment and observational skills" (ibid., p. 64). I would think that this statement applies to any researcher. Or at least it should apply. Even more, it is known already a long time that:

> The belief that science proceeds from observation to theory is still so widely and so firmly held that my denial of it is often met with incredulity. [...] the fact that we can start with pure observations alone, without anything in the nature of a theory, is absurd [...] Twenty-five years ago I tried to bring home the same point to a group of physics students in Vienna by beginning a lecture with the following instruction: 'Take pencil and paper; carefully observe, and write down what you have observed!' They asked, of course, *what* I wanted them to observe. Clearly the instruction, 'Observe!' is absurd. [...] Observation is always

> selective. It needs a chosen object, a definite task, an interest, a point of view, a problem. (Popper, 1994, p. 61)

I have nothing to add, Popper made his point clearly and convincingly. Observation is necessarily selective because there are endlessly (literally!) many aspects of the world that can be observed. Therefore, we always select what we observe, the only question can be whether we are, or we are not, aware of the reasons why we observe that particular something and not anything else. But there is always a reason, a justification. The fact that Popper had to prove this obvious – after reading him – fact about observation brings one of the reasons why it is indeed "difficult to be aware of whether one knows or not" (see the quote from Aristotle, above).

Even though Popper's argument is convincingly grounded, it is still often ignored in methodology of science today. There seem to be four most prominent theories of scientific method today: inductive method, hypothetico-deductive method, Bayesian hypothesis testing and inference to the best explanation (e.g. Haig, 2014, pp. 5–11). In the referred book, Haig adds to them another, what he calls abductive theory of method. However, if Popper was correct – and I do not see any way to prove the contrary – then inductive theory is simply wrong, and most common today hypothetico-deductive method must be highly questionable.

Why Pure Induction Is Impossible

If observation is based on a theory – that must (!) exist before any observation is possible – then there can be no theory that emerges purely unidirectionally from induction, from bottom-up. So there can be no pure induction. But there is an opposite problem. If all observation begins with some "theory", then there must be some "first theory" or, rather, a set of "first theories"; there must exist some mechanism that underlies the selection of aspects of experiences that should be observed. I think it is also absurd to assume that theory about (the aspects of) the world can exist before any experience of it. This topic is too complex to elaborate here, so I propose only very short description of the solution to this seeming problem. I propose that the ability to observe is an emergent biotically grounded property of an organism; this property emerges in the process of development. One of the basic laws of development of all living matter is the law of differentiation, formulated by Karl Ernst von Baer:

> [...] if to look at the course of the formation [AT: of the embryo], then first of all what catches the eye is that here from the homogeneous, the general, gradually emerges the heterogeneous, the particular. (Baer, 1950, Scholion IIIa, p. 225)

So, all development begins from undifferentiated state (see, for detailed theory of psychic development, Toomela, 2017). The same, I suggest, applies to the relationship between a "theory" and "observation". In the beginning of development, there is a unitary phenomenon where what can be "observed" a "theory" is unequivocally related to what is actually "observed" if the "observable" appears in the environment of an organism. All living organisms relate to their environment with a system of receptors; the existence of a receptor is essentially a "prototheory", and it constrains what can be sensed; and "proto-observation", activation of a receptor takes place when corresponding to a receptor physical or chemical event becomes into contact with a receptor – in this way *Umwelt*, the world-as-sensed (von Uexküll, 1926, esp. pp. 126–127), of an organism emerges (see, for more details on *Umwelt*, Toomela, 2020a, Chap. 2).

Now, when psyche emerges, novel experiences can be constructed by an organism. First such novel experiences are constrained by biotic processes that determine the limits of sensation; the more the psyche develops, the more biotic constraints are overcome; and observation will be increasingly guided and constrained by thinking and by stored knowledge (see, for stages of psychic development, Toomela, 2017, and for early version of this theory in English, Toomela, 2000, 2003).

So, prototheories develop into theories over the course of psychic development. Sciencing, pursuit for knowledge about the nonsensory world, emerges at the highest stages of psychic development. Here knowledge construction becomes an aim in itself. The general principle of differentiation of theories and observations applies also in this special case. The first observations about the world that are going to ground science are not differentiated from the theories underlying them: in the beginning of construction of science, an observer is not aware about the theory that underlies his/her observations. Awareness emerges when a scientist begins to reflect on the results of observations.[8]

Why Hypothetico-Deductive Method Can Be Highly Fallible

If all observations, including those that are scientific, begin with some kind of theory, then it can be conjectured that it is the theory underlying the research question that constrains the nature of the result of observation, knowledge constructed on the basis of it. If the theory and, following from it, the question are ill-defined, then the answer might be meaningless. In the beginning of the history of science, all questions, independently of the area, must have been ill-defined in retrospect. The nature of the problem with defining good scientific questions is that science proceeds when nonsensory aspects of the studied things are understood. Yet in the beginning of studies, all theories, or, more correctly, pre-theories, must have been based only on knowledge that could be constructed without method – nonscientific knowledge that is based fully on information directly available for senses.

Here lies also a reason why hypothetico-deductive method – which from the second half of the nineteenth century became the most popular method of science

[8] I think this explains why pure induction may seem possible for many scholars. If the first observations that ground future explicit theories are not available for self-reflection, i.e. there is no awareness of them, then introspection inevitably leads to an illusion that the first observations were not based on any theory.

(Laudan, 1981) – can be fully misleading. If hypotheses are not theoretically well-developed, the studies will not lead to science. In physics and in biology today, well-defined questions are common if not even the only kind of questions asked. Psychology, however, is still in a sorry state today[9] – most of the questions remain tied to appearances. For instance, it is assumed that if a behaviour is identical in appearance then underlying it psychic processes must be identical also. This assumption, however, is clearly wrong: externally similar behaviours can rely on different psychic processes in different individuals as well as in case of the same individual in different times. It is also incorrectly believed that the structure of a task corresponds directly to the structure of psychic processes that underlie performance on the task. I bring just one example: it is believed that there can be tests for different cognitive processes, such as memory, attention, perception, thinking, etc. This, however, is impossible. No memory task can be solved without perception, attention and other psychic processes; the same is true about performance on all other psychological tasks: performance relies always on psyche operating as a whole.

Altogether, hypothetico-deductive method can be useful only when hypotheses to be studied are theoretically well-grounded. In the beginning of the development of any science, the hypotheses must be ill-defined. With the development of understanding, hypotheses worthy to study will be discovered, and the hypothetico-deductive method can become productive. In psychology today, majority of hypothetico-deductive studies do not lead to better understanding because the hypotheses are constructed following erroneous assumptions about the nature of the psyche.

Why Bayesian (and Haig's Abductive Theory of) Method Is Useless for Psychology

In Bayesian method scientific hypotheses and theory choices are based on statistical analysis of probabilities. It is a method where mathematical analysis is supposed to be suitable for deciding which theory or which hypothesis is more acceptable. Haig's abductive theory also relies heavily on using statistical data analysis methods, such as exploratory factor analysis.

However, not only statistical analyses but all mathematical approaches to *discover* novel aspects of the studied nonsensory reality can work only under very special constraint that occasionally applies in physics, but not in other sciences: Information that is encoded into variables must be unequivocally interpretable. In psychology, where the studied phenomenon, the psyche, manifests only in behaviour, there is no way to achieve such encoding of the observations. Externally

[9] Psychology is a strange science where more advanced theories and scientific approaches were gradually replaced with less developed ones somewhere in the middle of the twentieth century (cf., e.g. Toomela, 2007a, b, 2010c, 2012, 2016a, 2019, 2020a, b).

identical behaviours can be based on different psychic processes and vice versa. Therefore, after the behaviours are encoded – obviously similar behaviours placed into one and the same category and dissimilar ones into other categories – the studied phenomenon is already lost (Toomela, 2008a); the discovery became already impossible in the encoding phase. Yet there are more substantial problems with mathematics that constrain its use to very selected and highly constrained phases of the scientific discovery (see, on these problems and constraints, Toomela, 2010d, 2011).

It is sufficient to bring one fundamental shortcoming of mathematics that makes its use in scientific discovery extremely limited and definitely only secondary. Poincare made the point very clear:

> Mathematicians do not study objects, but the relations between objects; to them it is a matter of indifference if these objects are replaced by others, provided that the relations do not change. Matter does not engage their attention, they are interested by form alone. (Poincare, 1905, p. 20)

And this is the problem: mathematics cannot reveal what the studied thing or phenomenon *is*. It cannot even model the essence of a thing because replacing objects in mathematical formulas leads to no consequences if the relations remain the same. And here three fundamental limitations of any mathematics are hidden. First, we understand the world scientifically exactly when we have understood what is the studied thing and not before. Genetics became an entirely different science after it was discovered, what a gene is. Biology is scientific, because it has revealed what (mostly nonsensory!) things and phenomena are: genes, cells, viruses, organs, synaptic transmissions, etc.. Physics has defined elementary particles, atoms, molecules, fields, etc. Chemistry is only about what things are and how one thing can become into another. Many of those things can lose their identity and become "the same" in mathematical formulas. Many different natural phenomena have, for instance, fractal features. Mathematically they become the same, and nonmathematical language is needed to keep the difference.

The second problem with mathematics is the nature of modelled relations between objects – mathematical relations are imposed on natural phenomena. But in nature one and the same thing can enter into very many qualitatively (!) different relationships; from the standpoint of structural-systemic theory, it is also known that the qualities of one and the same thing change, depending on its relationships. I am not the same as a son (of my parents), a father (of my daughter), a friend (of my friends), a husband (of my wife), a prey (of a mosquito), a teacher (of my students), a head of a committee (of a group of people), etc. And a friend as a head of a committee is not the same as the same person in sauna. Or let us take a little more complex example. Would I be the same depending on whether I put my T-shirt on me as it is usually done or inside out? The T-shirt would keep my body temperature the same in both ways. I would be the same in this sense. But in social situations, it may not be the same. Nobody would comment when I am wearing my T-shirt in the socially common way – Hey, you did put your T-shirt on as other people do! But they would comment, at least in their minds, if they notice the inside out T-shirt.

And if I would have always worn my T-shirts inside out, some would notice and comment when I would wear it in a usual way. Mathematics is not needed – it is actually impotent – to understand such situations where the only change is in the quality of relationships between two things. In different relations I think differently, behave differently – I *am* different; qualitatively different.[10]

Third, mathematics cannot model discontinuity, the emergence of something that does not exist before, the synthesis of elements into a novel whole. In processes of emergence, understanding is achieved when it can be defined, what the novel whole *is*, emergence is discontinuous, nothing like the whole exists in or of the parts before the synthesis. Such a discontinuity cannot be modelled because, again, mathematics is indifferent to the objects, whose relations are mathematically described.

Altogether, mathematics is not *the* tool of scientific discovery. Even worse, mathematics is not *the* tool for describing what has been discovered also. Due to indifference to objects, which relations are modelled mathematically, mathematics can be useful only in modelling situations where qualitative differences of things participating in phenomena can be ignored. And due to the highly limited number of relationships between things that are described mathematically, mathematics makes sense only in modelling relationships between things and phenomena if such relationships do exist between the things in modelled phenomena. No mathematical method – or even all of them together – can be *the* methods of science (some areas of physics excluded, where the modelled relations correspond to the relations between studied things and things can be unequivocally encoded into mathematical symbols).

Why Inference to the Best Explanation Is Problematic

If there are different theories about a phenomenon studied, then it can be asked, whether they have the same explanatory power. If not, then a theory that explains the best should be selected. Obviously, such selection makes sense only when the criteria of what counts as an explanation are well selected. If it is accepted that scientific understanding is achieved when the structure of studied things or phenomena is revealed, then the criteria would become well-defined: the best explanation – which also means the best theory – is that which distinguishes parts of the studied whole, relationships between those parts and qualities of the whole.

It is important that the only way to answer these three structural-systemic questions is the study of development, the emergence and ceasing to be of what is studied (Toomela, 2009a). Therefore, a structural-systemic science also includes

[10] Here is an “easy” way to prove that I am wrong. It is sufficient if a mathematician would build a mathematical model of what I am – including qualitative variability of me in different physical, biotic and social relationships – so that it is possible to take the model and tell that it is a model of me and nobody else. If such a model can be created, then my critique of mathematics is wrong.

understanding of coming into being and ceasing to exist. Otherwise, it would have been impossible to answer the first three questions of structural-systemic science. This fact is extremely important because from it the criterion of truth can be deduced: When we know the necessary elements, the necessary kinds of relationships between elements and the way they become together, we can *make* the thing we study. If, following a theory, we succeed in creating the studied thing or phenomenon, the theory corresponds to reality, i.e. it is true (Toomela, 2016b; see more on this topic below). In many areas science has achieved true understanding: based on theories of physics, atoms can be created; based on chemistry, a myriad of different kinds of molecules can be created; based on biology, genes and neuromediators can be created; based on (Vygotsky-Luria's neuro-)psychology, higher psychological functions can be created, etc.

Now it becomes also clear why inference to the best explanation is problematic from the perspective of the structural-systemic science. As every single thing and phenomenon is composed of certain elements in certain relationships, then there is also only one single way to make each of those things. Theory is correct when we can make the thing we study on the basis of the theory. If we do not succeed in making what we study, the theory is not correct. The problem emerges when there are different unsuccessful theories. There is just no way to know, which one is closer to the truth, which one needs to be modified the least in order to arrive at a successful synthesis of the studied whole. This can be known only post hoc, after the true theory is created and tested.

Still, in many cases, inference to the best explanation might be necessary. For instance, there are things and phenomena that cannot be created by humans for purely technical reasons. Theories of stars, galaxies and other very big and/or distant things cannot be tested even though the theories of them can be correct. This human limitation can be partially overcome by "experiments" of nature. In principle, to test a theory, it is not necessary that a thing or phenomenon is fully created by a scientist. If theoretically posited processes of emergence and destruction can be observed in nature, that can be sufficient to prove a theory if all important aspects of the process can be followed: the elements that become parts, the kind of emerging relations between the parts and the qualities of the emerged whole. Yet sometimes nature does not make needed experiments; in that case, other criteria are needed to select between theories. The same applies to situations where theories are in the making and different possible directions of theory development are recognized.

I acknowledge that this topic may need further elaboration. Yet I am not going to discuss this issue further as it does not add to the main line of argument developed in this chapter. Instead, I go into the topic of this chapter – methods and methodology of sciencing.

Basic Kinds of Scientific Questions and the Methods Corresponding to Them

Now what is searched for in sciencing is defined – the science, a special kind of knowledge. It is clear also that usual divisions of methods into inductive, deductive and other kinds might be questionable. Knowledge can emerge only in the process of answering implicit or explicit questions. In this final part of the chapter, I am going to organize general methods of science according to the general kinds of questions that need to be asked and answered in the process of theory development.

In the beginning of studying a novel thing or phenomenon, it is not possible to immediately ask the questions that ground structural-systemic understanding. The essence of questions has to change in the process of sciencing. This is so because sciencing begins in an undifferentiated state of knowledge, where implicit theory cannot be distinguished from the methods of study. In the process of differentiation, novel kinds of questions emerge, and these require novel methods to answer them. This is why methodology can – and perhaps should – be approached from the perspective of questions.

In the following, I propose a hierarchical sequence of relationships between the kinds of questions and corresponding to the questions methods to answer them.[11] In principle, attempts to answer a question with appropriate methods lead not only to answers but, more importantly, to novel questions. These, in turn, require novel methods to answer them. Such question-method hierarchical loops end with construction of science, with structural-systemic understanding of what was studied.

The First Questions That May Lead to Sciencing

Science requires understanding of nonsensory world, and this, in turn, requires justified methods. Sciencing, however, would be impossible without an idea that nonsensory world exists at all. But how can sciencing begin after the idea of nonsensory world is accepted? The answer was already given, again, by Aristotle:

> When the objects of an inquiry, in any department, have principles, causes, or elements, it is through acquaintance with these that knowledge and understanding is attained. [...] The natural way of doing this is to start from the things which are more knowable and clear to us and proceed towards those which are clearer and more knowable by nature [...] So we must follow this method and advance from what is more obscure by nature, but clearer to us, towards what is more clear and more knowable by nature.
>
> Now what is to us plain and clear at first is rather confused masses, the elements and principles of which become known to us later by analysis. *Thus we must advance from universals to particulars; for it is a whole that is more knowable in sense-perception,* and a universal is a kind of whole, comprehending many things within it, like parts. (Aristotle, 1984b, 184a10–184b10, p. 315, my emphasis)

[11] In the following I am going to develop further ideas I have proposed in an earlier publication (Toomela, 2016b).

The answer is, indeed, obvious (after reading Aristotle, at least): we can begin only from what we know. And we know first only what is knowable on the basis of direct sense-perception. We know wholes – things and phenomena – that seem to be distinct one from another. At the same time, these distinct wholes can be grouped into categories according to similarities between the category members. Category members can be either identical to senses – the situation that would be rare in everyday life or distinct only in properties that are considered irrelevant. Novel knowledge that can be constructed concerns category membership. I propose that it is perceived contradictions in category membership that grounds the development of science.

In order to discover novel questions, the world must be actively studied. There is no sciencing yet as the knowledge that is looked for concerns only directly observable world. At this stage of knowledge-seeking, a question – that emerges on the basis of some "theory" – is not differentiated from ways of knowledge-seeking. What is observed is what is known about the world and what is known is observed. Therefore, the way to discover the first question that will lead to sciencing can be only nonmethodic *everyday observation*. But, as Popper showed, no observation can be theoryless. Indeed, the first scientific questions, I speculate, can emerge from "theories" that underlie everyday behaviour. When certain things belong to one category, it is expected that the category members behave in certain aspects in the same way. This is a "theory". If behaviour of such category members is observed and some repeated exceptions to the expected behaviours emerge, then naturally a question also emerges: *How something that is expected to behave in a certain way as a member of a category behaves differently?* The opposite can also happen. Members of different categories are expected to behave differently. When observing members of different categories, it may happen that behavioural similarities are observed. Such unexpected similarities may also be turned into a question: *How members of different categories can behave similarly?*

I will give examples for each kind of question and corresponding to it method of study. These examples I derive from Vygotsky-Luria's school of thought as this is the only theoretical school in psychology I am aware of that has consistently asked all kinds of scientific questions and used methods that correspond to them.

For example, Vygotsky discussed the possible differences between adults from different cultures:

> The behaviour of the modern cultural man is not only the product of the biological evolution, not only the result of the development in the childhood, but also the product of historical development. In the process of the development of the humankind not only the external relationships between people, not only the relationship between humankind and nature changed and developed, but also human him/herself changed, his/her own nature changed. (Vygotsky & Luria, 1930, p. 57)

So, Vygotsky observed that people from different cultures are different. Psychology, anthropology and other related sciences were developed at his time already to the level where it was not necessary for him to rely on personal everyday observations. These observations were available from written sources. Based on such – and more advanced forms of observations I am going to discuss next – Vygotsky concluded that people, even though representatives of the same species and in that sense members of the same category, behave remarkably differently. He went further and suggested that such differences are not only external; external differences in human behavioural patterns and ways how humans in different cultures relate to the nature are based on nonsensory differences in the human nature itself. Thus, Vygotsky already formulated the next question that leads to the differentiation of sciencing from other forms of knowledge construction.

Attempts to Answer the First Questions Can Lead to the Next: Is There a Nonsensory Cause?

With the emergence of the first question about unexpected behavioural differences or similarities, there is still no necessary question about nonsensory reality. In principle it is also possible that unexpected behavioural differences or similarities can be distinguished on the basis of directly observable properties that were considered to be unimportant before.

If it is decided to answer the first question, the way of the knowledge-seeking changes. Instead of (or, rather, in addition to) everyday observations, a kind of observation that I call *directed observation* emerges. Depending on the question – whether it is about unexpected differences or about unexpected similarities – the behaviour of the members of the same category or the members of different categories will be selectively observed. Such observations can have three outcomes. First, no answer is found and further observation is abandoned. Second, some directly observable property is discovered that can be used to recategorize the observed things or phenomena so that their behaviours become predictable regarding the novel aspect that was discovered with the first question. Third, no directly observable property is discovered so that a new question emerges: *Is there some nonsensory reason that could explain the unexpected behaviours?* Attempts to answer this question lead to the emergence of sciencing because now the use of a scientific method becomes necessary.

The passage from Vygotsky quoted at the end of the previous section already introduced the question of possible nonsensory differences in human nature that may underlie observed behavioural differences between biologically similar people. This hypothesis emerged on the basis of the directed observations by many anthropologists, who had described manifest cultural differences in great details.

It is interesting that Vygotsky's hypothesis is rejected today. All cultures are considered to be at the same level of development, and the cultural differences are attributed only to differences in external (!) conditions of life (e.g. Cole, 1996; Tulviste, 1988). Such conclusions must follow inevitably – and not because there is evidence based on studies with methodologically grounded methods to reject Vygotsky's hypothesis. Such conclusions were reached because certain questions were *not* asked. Without relevant questions, naturally, no relevant studies have been conducted as well. Vygotsky, however, asked and answered also the following questions.

How to Distinguish the Indistinguishable and How to Unite the Ununitable?

It is not possible to answer the question of hypothetical nonsensory reasons of unexpected observations in one step. To reveal nature's hidden powers and secrets, to use Hume's terms, it is necessary to discover ways of selecting what should be studied. It means that it is necessary to find a way to categorize things that cannot be distinguished on the basis of direct perception into different categories or to find a way to categorize things that seem to be different in appearances into one category. It is so because knowledge can emerge only in the process of comparison. Different things can be compared or one and the same thing in different times or contexts. And there is no point to comparing identical things or a thing with itself if it remains unchanged because only sameness could be discovered in this way.

So, to begin with studies of nonsensory differences, it must be first established whether things that appear identical to the senses could be distinguished or things that appear different to the senses could be considered identical in some nonsensory aspect. This can be done with *constrained observations*. These are observations where things are studied in different preselected contexts or situations.

Justification for this method follows from the structural-systemic principles: Qualities[12] of the whole are determined by the parts and relations between the parts of that whole. So, if some parts or relations are different, then the wholes must be different also. In case of complex wholes, a change in a part or some specific relationship between parts does not lead to the change of all the qualities of the whole; only some qualities change. This is why a thing can remain the same for the senses and yet be different in some quality that was not initially manifest. When the context of a thing is systematically changed, it can be observed whether the studied whole comes into relationship with something else in that context. Emergence of the relationship is manifested in emergence of a higher-order whole with novel qualities. In an externally identical context, if one thing remains identical to the senses as it enters a relationship and another does not, it can be conjectured that the things are different in some nonsensory aspect. It is also possible that both things form a relationship – but if the things are actually different, then the emergent whole will also be different.

It might seem impossible to plan constrained observations because the number of contexts where a studied thing could be observed is unlimited. A solution to this problem can come from answers to the earlier questions: it is already clear in what particular aspect to search for either differences or similarities. This knowledge can be used to constrain the study contexts.

Another restriction to constrained observations may emerge if creation of needed study situations is not possible for technical, ethical or other reasons. In that case

[12] "Quality is the potential of a structure to become into relationship with another structure" (Toomela, 2014b, p. 283).

spontaneously emerging situations may be specifically looked for, and constrained observations can be conducted in such situations.

Constrained observations have been and are very common in psychology. Vygotsky, for instance, hypothesized that the essence of the change in human nature that underlies the distinction of what he called "primitive" and "modern cultural man", respectively, is related to the way language is used by individuals in different cultures. So, together with his colleagues, he created several tests and tasks that required following explicit verbal instructions and giving explicit verbal responses. Among such tests were defining concepts, categorization of words, solving syllogisms, etc. (e.g. Luria, 1974, 1979). As expected, people expressed consistent qualitative differences in performance of such tests.

This kind of data is not sufficient for explanatory theory because the methods do not allow to distinguish possible nonsensory differences that may underlie externally similar behaviours and vice versa – the method allows only to distinguish externally undistinguishable (in this case, adult human) individuals in terms of subgroups. Despite arriving at conclusions that opposed Vygotsky about cultural differences, Cole and other authors have found similar subgroups to those discovered by Vygotsky's group. It is noteworthy that they also demonstrated that the same people that seem to be different while performing some tasks may still behave similarly while performing other tasks that are in several respects similar to the distinguishing tasks (e.g. Bernardo, 1998; Cole, 1996; Scribner & Cole, 1981).

So, distinguishing people in terms of groups on the basis of constrained observations is not sufficient. Opposite conclusions regarding the possible nonsensory differences in human nature, for instance, can be achieved with such methods. As the members of the same category – humans – are studied, obviously similarities between category members can be discovered in constrained observations. Therefore, discovering differences among members of the category,

the next question must be asked. This question is more specific: *Is there a nonsensory basis that underlies the differences discovered in constrained observations between members of the same category?*[13] With this question, the direction of further sciencing is highly constrained: studies focus on the hypothetical nonsensory mechanisms that underlie individual differences in *constrained* on the theoretical basis situations.

[13] The same questions must be asked differently when similarities among members of different categories are discovered. In the following discussion, I am focusing on explanation of differences between members of the same category; yet all questions and corresponding methods I am proposing apply equally to sciencing aimed at discovering nonsensory similarities between members of different categories.

There Seems to Be a Nonsensory Difference, How Did It Emerge?

Constrained observations allow to distinguish individuals into subgroups that may emerge on the basis of some mechanisms that are not directly observable. But even if the distinctions made on the basis of constrained observations turn out to be reliable and otherwise similar individuals manifest differences in constrained situations, it is not sufficient to posit one mechanism that explains the observed regularities. People can still be behaving similarly for different causes. One way to proceed would be to continue with novel kinds of constrained observations. But this way would lead to problems. Now a putative nonsensory characteristic is made manifest in a specific situation. Without understanding the reason for such differences, there is no justified ground to select other situational constraints that would concern the same underlying mechanism of the discovered distinction. I think here a new direction of studies is justified. This direction was also proposed by Vygotsky:

> [...] modern psychological type of the European or American [...] Characteristics of this type can be understood by us in no other way but by applying to it the genetic point of view, when we ask from where and how they originated. (Vygotsky & Luria, 1930, p. 57)

Studying development, as was also mentioned earlier in this chapter, is absolutely necessary to achieve structural-systemic understanding of whatever is studied. This is the only way to distinguish elements or parts of wholes. This is so because properties of the parts change when synthesized into a higher-order whole. Thus, no part can be characterized when it is already part of the whole. Parts can be described only when they are not yet parts of the whole or after the whole disintegrates.[14]

However, it is not possible to begin immediately with revealing parts and their relationships in developmental studies because the studied whole is not sufficiently distinguished yet. The nonsensory differences in observed behaviours can be posited with *constrained observations of development*. The structure of the studied thing is a result of development, the result of hierarchical reorganization of a system (see, for detailed study of the mechanisms of development, Toomela, 2017). Development is quite often equifinal: "... in open systems. Here the same final state may be reached from different initial conditions and in different ways. This is what is called equifinality…" (von Bertalanffy, 1968, p. 40).

It is important that, from the perspective of scientific methodology, two kinds of equifinality must be distinguished. In one case the wholes emerge that are composed from the same elements in the same relationships between them. Only the order of the synthesis has been different. In that case the initial states have been different, but the wholes that emerge in the end will be identical. There is also another possibility: different elements and/or different relations between them are

[14] Here the situation is more complex. Often the whole does not disintegrate into the same set of parts it was synthesised from. Therefore methodology of disintegration studies requires description of conditions when conclusions about the parts of the whole can be made.

formed in the process of development. In that case, also, the emerged wholes can be similar – but only in certain aspects; such wholes can never be identical. David Hume's observation is a good example here:

> The bread, which I formerly eat, nourished me; that is, a body of such sensible qualities, was, at that time, endowed with such secret powers: But does it follow, that other bread must also nourish me at another time, and that like sensible qualities must always be attended with like secret powers? The consequence seems nowise necessary. (Hume, 1999, 4:16, p. 114)

So, two pieces of bread may be identical for senses yet different in their nonsensory structure. Constrained observations of development can be used to discover whether externally similar wholes have had similar or different paths of development. In the former case, it can be conjectured that the emergent in the process of the development wholes are identical. In the latter case, it can be suggested that similar to senses wholes *may* be different in nonsensory structure – but not necessarily. It might be that the wholes are identical but just the order in which the synthesis of the wholes took place was different. In addition, constrained observations of development may reveal cases where individuals, who are theoretically identical in the beginning of development, become different in certain aspects when they pass through different paths of development.

Vygotsky's group conducted numerous constrained observations of development. For instance, in studies conducted in Central Asia, adults with different educational backgrounds were compared on the set of similar tests. It turned out that individuals with no formal education seemed to be qualitatively different from individuals who had attended school. It was conjectured that developmental paths of the participants – with or without formal schooling – led to fundamental reorganization of psyche, to the change of the human nature (Luria, 1974). In other studies, the developmental process itself was observed, like in the case of studying concept formation with the method created by Ach and modified by Sakharov. Studies of children, adults and individuals with different pathological conditions revealed qualitative differences in relationships between the elements of concepts that superficially seem identical in adults and in children (Sakharov, 1994; Vygotsky, 1934, esp. Ch. 5). In this way Vygotsky's group demonstrated that there are qualitative differences in thought operations between more and less developed humans.

Constrained observations of development provide a strong ground to suggest the existence of nonsensory structural differences between cases that appear similar and vice versa. The wholes to be studied are definable now and ground emerges to come to the next question: *What is the (nonsensory level) structure of the studied things or phenomena that underlies differences of the wholes*? This question cannot be answered with only one kind of methods.

What Are the Parts?

I would say that only now the sciencing proper begins. Only from this point the structural-systemic science begins to be constructed step by step. So far, the questions answered in studies were about how to distinguish wholes that *may* be different in nonsensory aspects. Now there is sufficient ground to determine what is the studied whole. Only after that it becomes possible to ask – and answer – the questions about the parts of that whole, including what those parts are and in which relationships they must be in order to make up the whole that is studied. Vygotsky was clear in this point – sciencing begins with the *analysis* of the whole:

> Every cultural method of behaviour, even the most complicated, can always be completely analysed into its component nervous and psychic processes, just as every machine, in the last resort, can be reduced to a definite system of natural forces and processes. *Therefore, the first task of scientific investigation, when it deals with some cultural method of behaviour, must be the analysis of that method,* i.e. *its decomposition into component parts*, which are natural psychological processes. (Vygotsky, 1994a, pp. 59–60, my emphasis)

In principle there are two ways to identify the parts of the studied whole. One possibility is to observe the emergence of what is studied, i.e. to observe the process of how the parts come together so that the whole of what we want to understand emerges. I will call this method of sciencing *analytic observation of emergence*. The other possibility is to begin from the whole and disintegrate it into parts. This method I will call *analytic observation of disintegration*. Both of these ways – and the only ones that exist to answer the question about parts – are not easy to apply in sciencing. The reasons why it is so follow from postulates of development discovered by James Mark Baldwin:

> The first or negative postulate: *the logic of genesis is not expressed in convertible propositions*. Genetically, A = (that is, *becomes*, for which the sign ((is now used) B; but it does not follow that B = (becomes, (() A. The second or positive postulate: that series of events only is truly genetic which cannot be constructed before it has happened, and which cannot be exhausted by reading backwards, after it has happened. (Baldwin, 1906, p. 21)

Let us take first a look at problems related to analytic observation of emergence. As Baldwin noticed with his second postulate, the result of true genetic (it means developmental) sequence cannot be known before the development has taken place, i.e. the novel whole has been already synthesised. It must be so also from the structural-systemic perspective: the whole has properties none of its parts have; so the whole that emerges in the synthesis of parts is different from all its parts. Therefore, the study of parts before the synthesis cannot lead to understanding of the properties of the whole. Methodologically it follows that it is impossible to discover what is part of the whole by studying potential parts before they are synthesised into the whole. Therefore, it is also not possible to begin analysis from the parts before the development has taken place because it cannot be known, which are the parts, which future synthesis must be observed.

However, as Baldwin's first postulate of development posits, analytic observation of disintegration is also problematic: the whole does not necessarily disintegrate into parts it was composed of because the (structural-systemic) causes of disintegration have been different from the causes of synthesis. Yet there is no other way to discover parts than observing how the structure is put together or how it disintegrates. It follows that the two methods of analytic observation must be complemented with some additional technique that allows to bypass the limitations of these inevitable methods. I think this technique, proposed by Aristotle, is the same that must be applied in the beginning of sciencing: *Thus we must advance from universals to particulars* (Aristotle, 1984b, 184a23–24, p. 315, my emphasis; see also above). In other words, here again the principle of differentiation becomes relevant. Scientific analysis into elements must begin from general distinctions and proceed towards particulars. Vygotsky applied exactly this method to analysis:

> Usually the two lines of psychological development (the natural and the cultural) merge into each other in such a way that it is difficult to distinguish them and follow the course of each of them separately. In case of sudden retardation of any one of these two lines, they become more or less obviously disconnected as, for example, in the case of different primitiveness. (Vygotsky, 1994a, p. 59)

I admit I may overinterpret this particular quote, but Vygotsky and his group definitely proceeded in this way: they first distinguished principal elements of the human psyche – "natural" and "cultural" in Vygotsky's terms – and then proceeded analytically towards more and more detailed levels of analysis of each of them. They used analytic observations of development, like in studies of so-called double stimulation described by Vygotsky as follows:

> […] the 'functional method of double stimulation', the essence of which may be reduced to the organization of the child's behaviour by the aid of two series of stimuli, each of which has a distinct 'functional importance' in behaviour. At the same time the *conditio* sine qua non of the solution of the task set the child is the 'instrumental use' of one series of stimuli, i.e. its utilization as an auxiliary means for carrying out any given psychological operation. (Vygotsky, 1994a, p. 69; see also Leontiev, 1931; Luria, 1928; Sakharov, 1994)

According to this method, a novel element, the linguistic sign, was introduced into the environment of the problem solving, and it was observed, whether and (if yes, then) how this novel element is related to the change in the structure of the solution. If addition of the element leads to qualitative changes, it can be conjectured that the element is part of the studied structure.

Vygotsky's group also used analytic observations of disintegration. Among them are, for instance, Vygotsky's observations of the consequences of Parkinson's disease or schizophrenia (Vygotsky, 1982b, 1994b) and, of course, Luria's extensive studies of the consequences of local brain damages (Luria, 1947, 1969).

How the Parts Are Related One to another?

With the emergence of knowledge about possible parts of the studied whole, the next question must be answered. Vygotsky, for instance, formulated the question in this way:

> The second task of scientific investigation is to elucidate the structure of that method. Although each method of cultural behaviour consists, as it is shown by the analysis, of natural psychological processes, yet that method unites them not in a mechanical, but in a structural way. In other words, all processes forming part of that method form a complicated functional and structural unity. (Vygotsky, 1994a, p. 61)

Structure of the whole is not only about the parts but also, and equally importantly, about how the parts are interrelated. The same parts in different relationships make qualitatively different wholes. Therefore, after discovering what the parts of the studied whole might be, it is necessary to answer the question, *How are the parts related to one another in the studied whole?* To answer this question, it is necessary to have an idea of possible parts, because relations have no separate existence; relations require minimally two elements that can be related.

When putative elements are distinguished, then what I have called *evocative experiments*[15] can be used to study both the relationships between the elements and also the elements themselves. There are two principal ways to conduct evocative experiments. One possibility is to attempt to change relationships between elements. Another possibility is to either add or separate elements to or from the studied structure. In both cases it is observed whether the whole changes in the expected way, i.e. whether certain "effect" can be achieved by such experimental manipulations.

Vygotsky's group also used evocative experiments. Here examples of double-stimulation studies are relevant again. Now it may seem that it cannot be so – how the same method can be at the same time constrained observation of development, analytic observation of emergence or disintegration and evocative experiment? Indeed, it cannot. Yet the method of double stimulation can be each of them. It is so because scientific method contains more than study materials and instructions for how to use them. Method always contains, in addition to materials and instructions, a question that is to be answered, corresponding to its response patterns that are observed and procedures of interpreting the responses. Even more, I repeat here, method is part of the theory of the studied system. So, when the theory changes, the method also changes – even when it looks similar in some respects, it is not the same method if the theory has changed. For example, double-stimulation situation can be used just for observing how people behave in such a situation; that would be constrained observation of development. The same test situation becomes analytic observation when the question becomes, how addition or removal of certain parts from the study situation is reflected in the test performance. Finally, by adding or removing parts from the study situation and making predictions about changes in

[15] I have troubles with finding the appropriate term for this kind of experiment. Following common today understanding of the theory of causality, such experiments could be called "causative". However, this primitive theory is concerned only with so-called efficient causality and therefore does not fully cover the complex nature of causality. The kind of experiment I am describing here just "makes things happen" or "calls/evokes 'effect' into being". So I use the term "evocative" but without reference to emotions or feelings. "Evoke" is used more or less with the same meaning in the term "evoked potential".

the behaviour of individuals, the same situation, from a studied person's perspective, becomes an evocative experiment.

Before going further, I would like to mention that a certain form of evocative experiment has become probably the most important form of studies in mainstream psychology today. There is, however, qualitative difference between the structural-systemic and the mainstream psychology (based on efficient causality) evocative experiments. The former allows to proceed towards scientific understanding, knowing elements, their relationships, qualities of the studied whole and the process of emergence and development of that whole. The latter, however, leads to no (structural-systemic) understanding but only to probabilistic prediction: if one event, the "cause", takes place, then the other event, the "effect", might be observed beyond the chance level. Such studies can be useful if some practically applicable knowledge emerges in the study but are useless if the results remain only "theoretical". The problem is that in the efficient-causality experiments the "cause" is not conceptualized as a part of a structure but very vaguely as "circumstances that come before" (e.g. Myers & Hansen, 2002, p. 19). Such circumstances or events as "antecedent conditions" that are supposed to be "causes" place no constraint on how an event is conceptualized, and, therefore, it is not possible to know whether some potential part of a structure was manipulated or an undetermined set of possibly important factors together (see, for more on this, Toomela, 2016b).

Did We Get It Right? Confirmation of Truth

One moment sciencing comes to the point when, hypothetically, all the individually necessary and collectively sufficient elements of the studied structure are found and their relations described. This hypothesis – a question, "*Are all and only elements of the studied structure distinguished and described in correct relationships?*" – must be also answered before it can be suggested that scientific truth has been constructed through sciencing. The topic of the possibility of knowing the truth is too complex to cover in this chapter. Briefly, there can be no absolute criterion of truth, because certain assumptions any science must make cannot be proven in principle. First of all, it is not even possible to prove with certainty that reality external to "me" exists (Toomela, 2019). If we assume, however, that an external reality exists independently of us, that this reality is organized, that this reality is knowable to us in principle and that the external organized reality is only material, then it is possible to define the criterion of the truth – which can be established by the method to answer the final question; I have called this method *constructive experiment*. The idea of this kind of experiment and following from this knowledge of truth was first formulated by Engels. Germ of the idea, however, can be found in Marx's early notes:

> The question whether objective truth can be attributed to human thinking is not a question of theory but is a *practical* question. Man must prove the truth, i.e., the reality and power, the this-worldliness of his thinking in practice. (Marx, 1976, p. 3)

Engels explained how to understand the role of "practice" in attaining truth. He asked:

> Is our thinking capable of knowing the real world? Are we able to produce a correct reflection of reality in our ideas and notions of the real world? (Engels, 1996, p. 18)

He suggested that many philosophers answer this question affirmatively, whereas there are others, among them Hume[16] and Kant, who challenge the possibility of any knowledge, or at least of an exhaustive knowledge, of the world. Engels disagreed:

> The most telling refutation of this as of all other philosophical crotchets is practice, namely, experiment and industry. If we are able to prove the correctness of our understanding of a natural process by making it ourselves, producing it from its preconditions and making it serve our own purposes into the bargain, then it's all over with the Kantian ungraspable "thing-in-itself". The chemical substances produced in the bodies of plants and animals remained such "things-in-themselves" until organic chemistry began to produce them one after another [...]. (ibid., p. 19)

Engels not only showed what kind of experiment should be conducted in order to prove that truth is known, but he also showed what kind of method cannot lead to understanding:

> A striking example of how little induction can claim to be the sole or even the predominant form of scientific discovery occurs in thermodynamics: the steam-engine provided the most striking proof that one can impart heat and obtain mechanical motion. 100,000 steam-engines did not prove this more than one, but only more and more forced the physicists into the necessity of explaining it. […] The empiricism of observation alone can never adequately prove necessity. Post hoc but not *propter hoc*. […] But the proof of necessity lies in human activity, in experiment, in work: if I am able to *make* the post hoc, it becomes identical with the *propter hoc*. (Engels, 1987, pp. 509–510)

Furthermore, Engels also demonstrated how – in my terms structural-systemic – causality can be revealed by "making the *post hoc*" and how such knowledge is supported (!) with situations where the expected result does *not* take place:

> If we bring together in a rifle the priming, the explosive charge, and the bullet and then fire it, we count upon the effect known in advance from previous experience, because we can follow in all its details the whole process of ignition, combustion, explosion by the sudden conversion into gas and pressure of the gas on the bullet. And here the sceptic cannot even say that because of previous experience it does not follow that it will be the same next time. For, as a matter of fact, it does sometimes happen that it is *not* the same, that the priming or the gunpowder fails to work, that the barrel bursts, etc. But it is precisely this which *proves* causality instead of refuting it, because we can find out the cause of each such deviation from the rule by appropriate investigation: chemical decomposition of the priming, dampness, etc., of the gunpowder, defect in the barrel, etc., etc., so that here the test of causality is so to say a *double* one. (ibid., pp. 510–511)

[16] Indeed, Hume believed that humans are not able to know the world fully; nonsensory reality is unknowable according to him. This was also the reason why he suggested that only efficient causality can be known; other aspects of causality are not knowable. Psychology today accepts efficient causality as the only causal knowledge that can and should be ultimately achieved by what they call science. I would say that sciencing proper begins from where modern psychology (both mainstream quantitative and non-mainstream modern qualitative) has achieved its final and highest state of knowledge (cf. Toomela, 2012, 2019).

Gun, a functioning gun, can be put together only when it is known what parts it must (!) have and what kind of relations must (!) be between the parts. This is scientific understanding of the gun. If the gun is not functioning in the expected way, then the study of its parts and relations between the parts allows us to understand why it did not work. Experiments that reveal only efficient causality can never explain situations where the expected "effect" does not follow the assumed "cause". Such knowledge is not scientific by the definition of science I proposed but only pretends to be.

Now it can be said that such structural-systemic knowledge can be found in physics, chemistry and biology – it is a fact that science has been created in these fields – but not in psychology because mind is something so special. But, again, it is a fact that such knowledge has been achieved in psychology too: there is ample evidence of successful constructive experiments in neuropsychological rehabilitation (cf., Luria, 1947, 1948; Tsvetkova, 1985). Numerous successful cases of recovery of lost psychic functions through reorganization of the psychological structure of it in the process of special teaching-learning demonstrate that Vygotsky's school achieved structural-systemic scientific knowledge about several psychic functions. Vygotsky-Luria's theories about the structure of many specific psychic functions are proven to be true. I would say that psyche is not understood scientifically yet; only a fraction of it is understood. But there is sufficient evidence that scientific understanding of psyche is in principle possible.

References

Angioni, L. (2016). Aristotle's definition of scientific knowledge. *History of Philosophy and Logical Analysis, 19*(1), 140–166. https://doi.org/10.30965/26664275-01901010

Anokhin, P. K. (1974). *Biology and neurophysiology of the conditioned reflex and its role in adaptive behavior*. Pergamon Press.

Anokhin, P. K. (1978). Operezhajuscheje otrazhenije deistvitel'nosti. (Anticipating reflection of actuality. In Russian. Originally published in 1962.). In F. V. Konstantinov, B. F. Lomov, & V. B. Schvyrkov (Eds.), *P. K. Anokhin. Izbrannyje trudy. Filosofskije aspekty teorii funktsional'noi sistemy* (pp. 7–26). Nauka.

Ardiel, E. L., & Rankin, C. H. (2010). An elegant mind: Learning and memory in *Caenorhabditis elegans*. *Learning and Memory, 17*, 191–201.

Aristotle. (1941a). On generation and corruption (De generatione et corruptione). In R. McKeon (Ed.), *The basic works of Aristotle* (pp. 467–531). Random House.

Aristotle. (1941b). Posterior analytics. In R. McKeon (Ed.), *The basic works of Aristotle* (pp. 110–186). Random House.

Aristotle. (1984a). Metaphysics. In J. Barnes (Ed.), *The complete works of Aristotle. The revised Oxford translation* (Vol. 2, pp. 1552–1728). Princeton University Press.

Aristotle. (1984b). Physics. In J. Barnes (Ed.), *The complete works of Aristotle. The revised Oxford translation* (Vol. 1, pp. 315–446). Princeton University Press.

Aristotle. (1984c). Posterior analytics. In J. Barnes (Ed.), *The complete works of Aristotle. The revised Oxford translation* (Vol. 1, pp. 114–166). Princeton University Press.

Aristotle. (1984d). Topics. In J. Barnes (Ed.), *The complete works of Aristotle. The revised Oxford translation* (Vol. 1, pp. 167–277). Princeton University Press.

Bacon, F. (2000). The New Organon. (Originally published in 1620). In L. Jardine & M. Silverthorne (Eds.), *Francis Bacon. The New Organon*. Cambridge University Press.

Baer, K. E. v. (1950). Istorija razvitija zhivotnykh. Nabljudenija i razmyshlenija. Tom 1. (In Russian. Originally published in 1828 as Über Entwickelungsgeschichte der Thiere. Beobachtung und Reflexion. Erster Theil.). In E. N. Pavlovskii (Ed.), *K. M. Ber. Istorija razvitija zhivotnykh. Nabljudenija i razmyshlenija. Tom pervyi* (pp. 9–376). Izdatel'tvo Akademii Nauk SSSR.

Baldwin, J. M. (1906). *Thought and things. A study of the development and meaning of thought or genetic logic. Volume I. Functional logic, or genetic theory of knowledge*. Swan Sonneschein &.

Barnhart, R. K. (Ed.). (1988). *Chambers dictionary of etymology*. H. W. Wilson Company.

Bernardo, A. B. I. (1998). *Literacy and the mind. The contexts and cognitive consequences of literacy practice*. UNESCO Institute for Education.

Branco, A. U., & Valsiner, J. (1997). Changing methodologies: A co-constructivist study of goal orientations in social interactions. *Psychology and Developing Societies, 9*(1), 35–64. https://doi.org/10.1177/097133369700900103

Callaway, E. (2015). The revolution will not be crystallized. *Nature, 525*, 172–174. https://doi.org/10.1038/525172a

Callaway, E. (2020). 'It will change everything': AI makes gigantic leap in solving protein structures. *Nature, 588*, 203–204. https://doi.org/10.1038/d41586-020-03348-4

Chambers, E. (1728a). *Cyclopædia, or, An universal dictionary of arts and sciences : containing the definitions of the terms, and accounts of the things signify'd thereby, in the several arts, both liberal and mechanical, and the several sciences, human and divine : the figures, kinds, properties, productions, preparations, and uses, of things natural and artificial : the rise, progress, and state of things ecclesiastical, civil, military, and commercial : with the several systems, sects, opinions, &c : among philosophers, divines, mathematicians, physicians, antiquaries, criticks, &c : the whole intended as a course of antient and modern learning. In two volumes. Volume the first.* Retrieved from http://digicoll.library.wisc.edu/HistSciTech/subcollections/CyclopaediaAbout.html

Chambers, E. (1728b). *Cyclopædia, or, An universal dictionary of arts and sciences : containing the definitions of the terms, and accounts of the things signify'd thereby, in the several arts, both liberal and mechanical, and the several sciences, human and divine : the figures, kinds, properties, productions, preparations, and uses, of things natural and artificial : the rise, progress, and state of things ecclesiastical, civil, military, and commercial : with the several systems, sects, opinions, &c : among philosophers, divines, mathematicians, physicians, antiquaries, criticks, &c : the whole intended as a course of antient and modern learning. In two volumes. Volume the Second.* Retrieved from http://digicoll.library.wisc.edu/HistSciTech/subcollections/CyclopaediaAbout.html

Chang, H. (2017). Foresight in scientific method. In L. W. Sherman & D. A. Feller (Eds.), *Foresight. (Darwin College lectures, volume 27)* (pp. 82–100). Cambridge University Press.

Cole, M. (1996). *Cultural psychology. A once and future discipline*. The Belknap Press of Harvard University Press.

Dance, A. (2020). Feel the force. *Nature, 577*, 158–160. https://doi.org/10.1038/d41586-019-03955-w

Engels, F. (1987). Dialectics of nature. (originally written in 1873-1882). In N. Rudenko & Y. Vorotnikova (Eds.), *Karl Marx, Frederick Engels. Collected works* (Vol. 25, pp. 313–590). International Publishers.

Engels, F. (1996). *Ludwig Feuerbach and the end of classical German philosophy. (Originally published in 1888)*. Foreign Language Press.

Fredens, J., Wang, K., de la Torre, D., Funke, L. F. H., Robertson, W. E., Christova, Y., et al. (2019). Total synthesis of *Escherichia coli* with a recoded genome. *Nature, 569*, 514–518. https://doi.org/10.1038/s41586-019-1192-5

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review, 98*(2), 254–267.

Gigerenzer, G. (1992). Discovery in cognitive psychology: New tools inspire new theories. *Science in Context, 5*(2), 329–350. https://doi.org/10.1017/S0269889700001216

Gozli, D. (2019). *Experimental psychology and human agency*. Springer.

Gross, M. (2011). What exactly is synthetic biology? *Current Biology, 16*, R611–R614. https://doi.org/10.1016/j.cub.2011.08.002

Haig, B. D. (2014). *Investigating the psychological world. Scientific method in the behavioral sciences*. MIT Press.

Hume, D. (1999). An enquiry concerning human understanding. (Originally published in 1748). In T. L. Beauchamp (Ed.), *David Hume. An enquiry concerning human understanding*. Oxford University Press.

Kapitsa, P. L. (1977). Rol' vydajuschegosja uchenogo v razvitii nauki. (Originally a conference presentation, 1971). In P. L. Kapitsa (Ed.), *Eksperiment. Teoriya. Praktika. Stat'i, Vystupleniya. Izdaniye vtoroe* (pp. 248–254). Nauka.

Kohanovski, V. P., Zolotuhina, Y. V., Leshkevich, T. G., & Fathi, T. B. (2003). *Filosofija dlja aspirantov: Uchebnoje posobije*. Feniks.

Laudan, L. (1981). *Science and hypothesis. Historical essays on scientific methodology*. Springer.

Leontiev, A. (1931). *Razvitije pamjati. Eksperimental'noje issledovanije vysshih psikhologicheskih funktsii*. Gosudarstvennoye Uchebno-Pedagogicheskoye Izdatel'stvo.

Liddell, H. G., Scott, R., & Jones, H. S. (1940a). *A Greek-English lexicon. Volume I*. Clarendon Press.

Liddell, H. G., Scott, R., & Jones, H. S. (1940b). *A Greek-English lexicon. Volume II*. Clarendon Press.

Lindstad, T. G., Stanicke, E., & Valsiner, J. (Eds.). (2020). *Respect for thought. Jan Smedslund's legacy for psychology*. Springer.

Logie, R. H. (2018). Human cognition: Common principles and individual variation. *Journal of Applied Research in Memory and Cognition, 7*, 471–486.

Luria, A. R. (1928). The problem of the cultural behavior of the child. *Pedagogical Seminary and Journal of Genetic Psychology, 35*, 493–506.

Luria, A. R. (1947). *Travmaticheskaja afasia. Klinika, semiotika i vosstanovitel'naya terapiya. (Traumatic aphasia. Clinic, semiotics, and rehabilitation.)*. Izdatel'stvo Akademii Meditsinskikh Nauk SSSR.

Luria, A. R. (1948). *Vosstanovlenije funkcii mozga posle vojennoi travmy. (Restoration of brain functions after war trauma. In Russian)*. Izdatel'stvo Akademii Medicinskih Nauk SSSR.

Luria, A. R. (1969). *Vyshije korkovyje funktsii tsheloveka i ikh narushenija pri lokal'nykh porazenijakh mozga. (Higher cortical functions in man and their disturbances in local brain lesions.)* (2nd ed.). Izdatel'stvo Moskovskogo Universiteta.

Luria, A. R. (1973). *Osnovy neiropsikhologii*. Izdatel'stvo MGU.

Luria, A. R. (1974). Ob istoricheskom razvitii poznavatel'nykh processov. *Eksperimental'no-psikhologicheskoje issledovanije.* Nauka.

Luria, A. R. (1979). *Jazyk i soznanije*. Izdatel'stvo Moskovskogo Universiteta.

Luria, A. R. (2002). L. S. Vygotsky and the problem of functional localization. (Originally published in 1966). *Journal of Russian and East European Psychology, 40*(1), 17–25.

Marx, K. (1976). Theses on Feuerbach. (Originally written in 1845). In *Karl Marx, Frederick Engels. Collected works, Volume 5* (pp. 3–5). International Publishers.

Marx, K. (1981). Capital. In *A critique of political economy. Volume three.* Penguin Books.

Marx, K. (1985). Value, price and profit. (Written in 1865). In K. Marx & F. Engels (Eds.), *Karl Marx and Frederick Engels. Collected works. Volume 20. Marx and Engels 1864–68* (pp. 101–149). International Publishers.

Mccleskey, E. W. (2019). A mechanism for touch. *Nature, 573*, 100–200. https://doi.org/10.1038/d41586-019-02454-2

Michell, J. (2004). *Measurement in psychology. Critical history of a methodological concept*. Cambridge University Press.

Michell, J. (2012a). Alfred Binet and the concept of heterogeneous orders. *Frontiers in Psychology, 3*(261), 1–8. https://doi.org/10.3389/fpsyg.2012.00261

Michell, J. (2012b). "The constantly recurring argument": Inferring quantity from order. *Theory and Psychology, 22*(3), 255–271. https://doi.org/10.1177/0959354311434656

Molenaar, P. C. M. (2004a). Forum discussion of the Manifesto's Aggregation Act. *Measurement, 2*(4), 248–254.

Molenaar, P. C. M. (2004b). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement, 2*(4), 201–218.
Molenaar, P. C. M. (2007). Psychological methodology will change profoundly due to the necessity to focus on intra-individual variation: Commentary on Toomela. *Integrative Psychological and Behavioral Science, 41*(1), 35–40.
Molenaar, P. C. M. (2008). Consequences of the ergodic theorems for classical test theory, factor analysis, and the analysis of developmental processes. In S. M. Hofer & D. F. Alwin (Eds.), *Handbook of cognitive aging* (pp. 90–104). Sage.
Myers, A., & Hansen, C. H. (2002). *Experimental psychology* (5th ed.). Wadsworth.
Poincare, H. (1905). *Science and hypothesis*. Walter Scott Publishing.
Popper, K. (1994). *Conjectures and refutations*. Routledge.
Powell, K. (2018). Biology from scratch. *Nature, 563*, 172–175. https://doi.org/10.1038/d41586-018-07289-x
Prigogine, I., & Stenger, I. (1984). *Order out of chaos. Man's new dialogue with nature*. Bantam Books.
Ruiz-Mirazo, K., & Moreno, A. (2013). Synthetic biology: Challenging life in order to grasp, use, or extend it. *Biological Theory, 8*(4), 376–382. https://doi.org/10.1007/s13752-013-0129-8
Sakharov, L. (1994). Methods for investigating concepts. (originally published in 1930). In R. Van der Veer & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 73–98). Blackwell.
Science Council. (2009). *Our definition of science*. Retrieved from https://sciencecouncil.org/about-science/our-definition-of-science/
Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Harvard University Press.
Smedslund, J. (1988). *Psycho-Logic*. Springer.
Smedslund, J. (1991). The pseudoempirical in psychology and the case for psychologic. *Psychological Inquiry, 2*(4), 325–338. https://doi.org/10.1207/s15327965pli0204_1
Solvi, C., Al-Kudhairy, S. G., & Chittka, L. (2020). Bumble bees display cross-modal object recognition between visual and tactile senses. *Science, 367*, 910–912. https://doi.org/10.1126/science.aay8064
Toomela, A. (2000). Stages of mental development: Where to look? *Trames: Journal of the Humanities and Social Sciences, 4*(1), 21–52.
Toomela, A. (2003). Development of symbol meaning and the emergence of the semiotically mediated mind. In A. Toomela (Ed.), *Cultural guidance in the development of the human mind* (pp. 163–209). Ablex Publishing.
Toomela, A. (2007a). Culture of science: Strange history of the methodological thinking in psychology. *Integrative Psychological and Behavioral Science, 41*(1), 6–20. https://doi.org/10.1007/s12124-007-9004-0
Toomela, A. (2007b). History of methodology in psychology: Starting point, not the goal. *Integrative Psychological and Behavioral Science, 41*(1), 75–82. https://doi.org/10.1007/s12124-007-9005-z
Toomela, A. (2008a). Variables in psychology: A critique of quantitative psychology. *Integrative Psychological and Behavioral Science, 42*(3), 245–265. https://doi.org/10.1007/s12124-008-9059-6
Toomela, A. (2008b). Vygotskian cultural-historical and sociocultural approaches represent two levels of analysis: Complementarity instead of opposition. *Culture and Psychology, 14*, 57–69. https://doi.org/10.1177/1354067X07085812
Toomela, A. (2009a). How methodology became a toolbox - and how it escapes from that box. In J. Valsiner, P. Molenaar, M. Lyra, & N. Chaudhary (Eds.), *Dynamic process methodology in the social and developmental sciences* (pp. 45–66). Springer.
Toomela, A. (2009b). Kurt Lewin's contribution to the methodology of psychology: From past to future skipping the present. In J. Clegg (Ed.), *The observation of human systems. Lessons from the history of anti-Reductionistic empirical psychology* (pp. 101–116). Transaction Publishers.

Toomela, A. (2010a). Methodology of idiographic science: Limits of single-case studies and the role of typology. In S. Salvatore, J. Valsiner, J. T. Simon, & A. Gennaro (Eds.), *Yearbook of idiographic science, volume 2/2009* (pp. 13–33). Firera & Liuzzo Publishing.

Toomela, A. (2010b). Modern mainstream psychology is the best? Noncumulative, historically blind, fragmented, atheoretical. In A. Toomela & J. Valsiner (Eds.), *Methodological thinking in psychology: 60 years gone astray?* (pp. 1–26). Information Age Publishing.

Toomela, A. (2010c). Poverty of modern mainstream psychology in autobiography. Reflections on *a history of psychology in autobiography*. *Culture and Psychology, 16*(1), 127–144. https://doi.org/10.1177/1354067X09344892

Toomela, A. (2010d). Quantitative methods in psychology: Inevitable and useless. *Frontiers in Psychology, 1*(29), 1–14. https://doi.org/10.3389/fpsyg.2010.00029

Toomela, A. (2011). Travel into a fairy land: A critique of modern qualitative and mixed methods psychologies. *Integrative Psychological and Behavioral Science, 45*(1), 21–47. https://doi.org/10.1007/s12124-010-9152-5

Toomela, A. (2012). Guesses on the future of cultural psychology: Past, present, and past. In J. Valsiner (Ed.), *The Oxford handbook of culture and psychology* (pp. 998–1033). Oxford University Press.

Toomela, A. (2014a). Methodology of cultural-historical psychology. In A. Yasnitsky, R. van der Veer, & M. Ferrari (Eds.), *The Cambridge handbook of cultural-historical psychology* (pp. 99–125). Cambridge University Press.

Toomela, A. (2014b). A structural systemic theory of causality and catalysis. In K. R. Cabell & J. Valsiner (Eds.), *The catalyzing mind. Beyond models of causality* (pp. 271–292). Springer.

Toomela, A. (2015). Towards understanding biotic, psychic and semiotically-mediated mechanisms of anticipation. In M. Nadin (Ed.), *Anticipation: Learning from the past* (pp. 431–455). Springer.

Toomela, A. (2016a). Six meanings of the history of science: The case of psychology. In S. H. Klempe & R. Smith (Eds.), *Centrality of history for theory construction in psychology* (pp. 47–73). Springer.

Toomela, A. (2016b). The ways of scientific anticipation: From guesses to probabilities and from there to certainty. In M. Nadin (Ed.), *Anticipation across disciplines* (pp. 255–273). Springer.

Toomela, A. (2017). *Minu Ise areng: Inimlapsest Inimeseks. (Development of my self: From the human child to the human.)*. Väike Vanker.

Toomela, A. (2019). *The psychology of scientific inquiry*. Springer.

Toomela, A. (2020a). *Culture, speech and my self*. Väike Vanker.

Toomela, A. (2020b). Psychology today: Still in denial, still outdated. *Integrative Psychological and Behavioral Science, 54*(3), 563–571. https://doi.org/10.1007/s12124-020-09534-3

Toomela, A., & Valsiner, J. (Eds.). (2010). *Methodological thinking in psychology: 60 years gone astray?* Information Age Publishing.

Tsvetkova, L. S. (1985). *Neiropsikhologicheskaja reabilitatsija bol'nykh. Rech i intellektual'naja dejatel'nost. (Neuropsychological rehabilitation of a sick person. Speech and intellectual activity. In Russian.)*. Izdatel'stvo Moskovskogo Universiteta.

Tulviste, P. (1988). *Kul'turno-istoricheskoje razvitije verbal'nogo myshlenija*. Valgus.

Valsiner, J. (2017). *From methodology to methods in human psychology*. Springer.

von Bertalanffy, L. (1968). *General system theory. Foundations, development, applications*. George Braziller.

von Uexküll, J. (1926). *Theoretical biology*. Harcourt, Brace & Company.

Vygotsky, L. S. (1926). *Pedagogicheskaja psikhologija. Kratkii kurs. (Educational psychology. A short course.)*. Rabotnik Prosveschenija.

Vygotsky, L. S. (1934). *Myshlenije i rech. Psikhologicheskije issledovanija. (Thinking and speech. Psychological investigations.)*. Gosudarstvennoje Social'no-ekonomicheskoje Izdatel'stvo.

Vygotsky, L. S. (1960). Problema razvitii is raspada vyshikh psikhicheskih funktsii. In A. N. Leontiev, A. R. Luria, & B. M. Teplova (Eds.), *L. S. Vygotsky. Razvitie vyshikh psikhicheskih funkcii. Iz neopublikovannykh trudov* (pp. 364–383). Moscow.

Vygotsky, L. S. (1982a). Istoricheski smysl psikhologicheskogo krizisa. Metodologicheskoje issledovanije. (Historical meaning of the crisis in psychology. A methodological study. Originally written in 1927; first published in 1982). In A. R. Luria & M. G. Jaroshevskii (Eds.), *L. S. Vygotsky. Sobranije sochinenii. Tom 1. Voprosy teorii i istorii psikhologii* (pp. 291–436). Moscow.

Vygotsky, L. S. (1982b). O psikhologicheskih sistemah. (Originally a lecture presented in 1930). In A. R. Luria & M. G. Jaroshevskii (Eds.), *L. S. Vygotsky. Sobranije sochinenii. Tom 1. Voprosy teorii i istorii psikhologii* (pp. 109–131). Moscow.

Vygotsky, L. S. (1982c). Psikhologija i uchenije o lokalizacii psikhicheskih funktcii. (Originally written in 1934). In A. R. Luria & M. G. Jaroshevskii (Eds.), *L. S. Vygotsky. Sobranije sochinenii. Tom 1. Voprosy teorii i istorii psikhologii* (pp. 168–174). Pedagogika.

Vygotsky, L. S. (1994a). The problem of the cultural development of the child. (Originally published in 1929). In R. v. d. Veer & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 57–72). Blackwell.

Vygotsky, L. S. (1994b). Thought in schizophrenia. (Originally published in 1934). In R. Van der Veer & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 313–326). Blackwell.

Vygotsky, L. S., & Luria, A. R. (1930). *Etjudy po istorii povedenija. Obezjana. Primitiv. Rebjonok*. Gosudarstvennoje Izdatel'stvo.

Watts, I. (1726). *Logick: Or the right use of reason in the enquiry after truth* (2nd ed.). Printed for John Clark and Richard Hett.

White, L. A. (1949). Science is sciencing. In L. A. White (Ed.), *The science of culture. A study of man and civilization* (pp. 3–21). Grove Press.

Yedronova, V. N., & Obcharov, A. O. (2013). Metody, metodologiya i logika nauchnykh issledovanii. *Ekonomicheskii Analiz: Teoria i Praktika, 9*(312), 14–23.

Chapter 9
Conclusion: From Experimental to Experiential Psychology

Jaan Valsiner and Davood Gozli

The experimental method is the cornerstone of psychology as a science. So we are told—over the past century in various disguises—by various experts and deep believers in the promise that psychology will one day become a "real" science. The label *method* is supposed to add credibility to what psychologists do, and the constant parallels made with the dependence of physics on experiments set the stage for playing the game of experimenter being in control of all the "variables" selected for inspection in a given study.

Yet there is a small feature of psychological experimentation that sets it drastically apart from the analogues with physics or chemistry—the phenomena studied in the latter *do not interpret what is going on with them* in an experiment. Human beings do. And even more fundamentally, their interpretation leads to change in their actions in the experimental context and their resistance to some "stimuli," the "effects" of which are supposedly being studied. The task that is initially given can shift in the process of the participation (as described in Chaps. 5 and 7 of this volume). Likewise, the motivation of the participant can change over the course of the experiment. The cherished notion of "control" by the experimenter is made indeterminate by the counter-active roles of the participants.

Thus, the experimental method is not a "standard conveyer belt" of testing cause-effect relations, but a theatrical encounter of different active persons—the experimenter (who pretends to "control" the situation) and the "research participant"[1]

[1] Note the historical changes in the labeling of these actors in the experimental situation (Bibace et al., 2009). First, they were called *observers*—as the experiments used introspective techniques. Then, they were called *Versuchsperson* in the German areas and *subjects* in English. Finally, by the twenty-first century, they are *research participants* who sign forms of giving up their rights of ownership of the data they produce for the anonymization of their person and the place. Note that the organizer of the study—the experimenter—is *not* considered to belong to the category of *participants*—even as her role in setting up an experiment is clear key participation. By that exclusion it becomes possible to remain uninformed of what actually happens in the experiment.

J. Valsiner
Communication and Psychology, Aalborg University, AALBORG, Denmark

D. Gozli (✉)
University of Macau, Taipa, Macau S.A.R., China

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9_9

(who is supposed to follow the instructions but whose generosity toward the experimenter actually lets the control illusion of the experimenter to thrive). The experimenter is the analogue of a theatre director who sets up the play, but does not play any part in it. As the director, she is completely dependent on the collaboration of the actors whose motivation to follow the given instruction may be lured by a small payment, a lottery with the chance to win some intermediate valued award, or getting points in the system of participants' pool.

This contrast—experiment as an administrative act which is fully under control of the director (experimenter), in contrast with the theatrical view where the role of the experimenter (theatre director) remains central, but her control is limited by the counter-actions of the participants, as well as of audience[2]—is worth further investigation as a culturally constructed and maintained social encounter. Its by-product can be new knowledge, yet its immediate significance of "research being conducted" or "experiment in progress" belongs to the category of societal rituals.

Experiment as an Administrative Act

Psychology over the twentieth century has managed to overlook the agentive roles of human beings in their lives and subsequently treated them as willing participants—once they have signed their "consent forms"—in various experiments set up under the models of basic sciences. Hence it has been relatively easy to present the experiment as a regular administrative act where the obedient citizens diligently and honestly carry out the instructions given by the administrator (Fig. 9.1) resulting in the production of the desired outcome (valid data).

Figure 9.1 provides a simple illustration—the experimenter (administrator) sets up the task and gives instructions, and the participant proceeds to perform the task, with the results duly collected and further analyzed. The act of performing the task is seen as that of benevolent obedience—it is assumed that the consenting participant is aligned with the administrator to achieve the results set up by expectations of the task. Even if there is a task shift in the middle of the performance, it is part of the commands coming from the administrator—the participant is tested as to her adaptation to that, rather than expected to introduce an uncontrollable personal take on the task (e.g., "I am tired of this boring task" or "I really do not like how the administrator treats me"). These acts of personal clandestine disobedience are not supposed to play any role in the experiment as the pure temple of science. They are either overlooked—easy to do as long as these personal constructs remain clandestine—or the given experimental session is trashed as a failure. The science of psychology over the past two centuries has been a science of obedient minds—while in the wider societies, these very same minds have been actively involved in divorces,

[2] The audience here is the readership of the published experimental results that judge these results through the culturally set prisms of societal relevance or through the sieve of moral responsibility.

Fig. 9.1 Traditional view of the experiment: controlling obedient participants

protests, revolutions, wars, and purposive efforts to resist structures, become rich, or live happily ever after.

Experiment as Theatre

We propose that the encounter of researchers with the researchees we call "experiment" belongs to the class of non-public theatre productions. The non-public nature of these productions is supposed to enhance their social prestige—as an event taking place behind the closed door of an ordinary room with the label "Laboratory of X" on the door. Thus, the experiment becomes a genre of theatrical productions in psychology, among others[3] (Davies & Harré, 1990; Harré & Secord, 1972). Figure 9.2 illustrates the theatrical structure of the experiment. The experimenter here is the theatre director who sets up the whole performance that entails not only creating the task (script) but also finding appropriate actors to carry out the roles in the script. Finding them is not an easy task—as it involves both the director's decision whether the given person fits the role and the person's willingness to participate. Rarely is the process of "*recruiting* subjects" reported in detail. When it is (see Günther de Araujo, 1998), the picture that emerges is not that of easy and casual invitation to participate, happily accepted by the researchees, but instead a complex set of ritualistic persuasion efforts intended to address various suspicions and counter-investigative strategies.

Once the set of actors is finalized, the theatrical act is ready to be enacted. The script here is set not only by its core instruction ("Do X!") but in contrast to outer conditions of what not to do or how not to perform. Thus, in giving rating scale tasks with focus on the first association of the object with the scale points, any

[3] In psychology, several other genres of comparable structures are used: "interview," "testing," "therapy," etc. These all have their own theatrical setup that differs in some details from that given in Fig. 9.2 but remains similar in the focus on scientific encounter as a form of performance art.

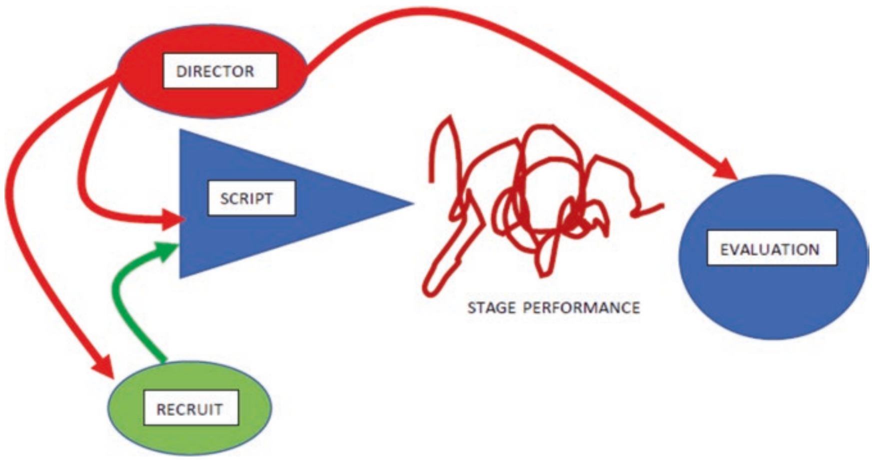

Fig. 9.2 Experiment as a theatrical stage performance

contemplation of the meanings of the scale end-points is de-emphasized (Rosenbaum & Valsiner, 2011). Likewise, there is no focus on elaborate and reflexive thinking through the meaning field in giving one's answer. The task is confined by borders, though the borders are not clearly set.

When participants switch between tasks, they are moving across normative situations and reconstructing the frame of their performance-participation. Chapter 6 (this volume) overviews the complexities of task switching in experiments, i.e., switching from one subset of instructions and rules to another set of instructions and rules. The authors point out that, despite the wealth of available data, the understanding of the underlying processes of task switching remains under-developed. Neglecting the theatrical structure of the situation of research, search continues for "inner" mechanisms and "underlying processes." A task is believed to be "loaded into working memory," rather than constructed in a collaborative process by the researchers and the researchees.

In a helpful illustration provided in Chap. 5, Ting (this volume) asks us to "imagine an experiment where participants are instructed to inhale pepper and then sneeze into the experimenter's face." She then asks, "Why is this instruction difficult to follow?" The norms that make such this strange—impolite, unhygienic, etc.—task difficult are, just as the instruction itself, in the social situation. The norms are not, at least not primarily, within the organism, inside the brain, or as contents of working memory. Instead, they are parts of larger, cooperative process in which a person can participate. The fact of conformity to a norm, or a request, is highlighted when there is a conflict between two norms (follow the instructions given by the researcher vs. maintain general good manners). The invisible "theatre director," by focusing all the attention on their research participants, downplays the presence of norms and the fact of social conformity.

Performance in an experiment, therefore, entails responsiveness to norms and the participants' co-construction of responses. While "the director" of the experiment might lose control over the actors' personal interpretations of their roles in the process of performing, she can certainly monitor the enactments and based on her evaluation of the process either trust or distrust the resulting data.

A good example here is the way in which a simple experimental procedure—the conservation of liquid quantity task of Jean Piaget—can be interpreted by actors of different background. The task is simple; the actor first evaluates the level of liquid in two similar-sized beakers (with the obvious result that the levels are judged to be the same). Then one of the beakers is poured into a third container with a wider base, and the actor is asked if the amount is the same or not. Obtaining the answer, "it is the same amount" is considered evidence of the cognitive achievement of conservation. This task works with children in occidental contexts as it follows the general principle that no liquid has been poured out from the system. The formally schooled actors in the West accept the general premise and do not question it.

However, when the task was carried out among the Qalandar in Pakistan (see account in Valsiner, 1984), the picture was very different. The Qalandar are peripatetic entertainers who earn their living from public presentation of tricks for their entertainment. The Piagetian experiment taken to the Qalandar was treated by them—the actors—as the task of finding out where is the "trick" in this show. And they easily found it—the amount of the liquid poured out is actually *not* the same because a very miniscule amount of it remains on the sides of the now empty beaker.

The two responses in the "conservation" experiment are based on two different ways of constructing the questions, only one of which aligns with the experimenter's assumptions. Studies reviewed in Chap. 7 (this volume) point to the fact that an explicit negotiation over interpreting the experimental situation is not necessary. Norms—which serve as the basis of co-construction—can be detected rapidly, effortlessly, and without direct instructions. Participants can move from "Are A and B contain the same amount of water?" to "Can you find how A and B are different despite appearances?" (Fig. 9.3). Based on their interpretations, research participants consider a type of action desirable or acceptable in a situation, without the researchers recognizing it. As a consequence, there are mismatches between how the situation interpreted by the participants and how it is interpreted by the researchers. Lacking a common interpretive basis—whereby the meaning of actions has already escaped the grasp of the researcher—there is no point in quantitative analysis of the "variables," just as there is no point in questioning the replicability of the results.

In Chap. 8 (this volume), Toomela emphasized the necessity of innovation with respect to methods. Scientific advancement requires advancement in methods, not the mindless application of the same methods. Methods, Toomela pointed out, cannot be separated from theories; methods are expressions of theoretical assumptions and commitments. Scientific activities are activities of knowing, which fundamentally differ from routinized (blind) production. The components that make up the whole of a scientific project are meaningful in light of that whole and because of their participation in the whole (see also Toomela, 2019; Valsiner, 2017). These

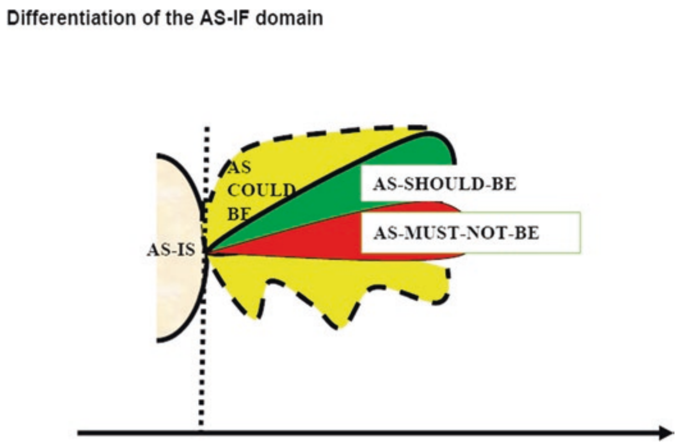

Fig. 9.3 How the experimental situation is interpreted shapes what is considered acceptable or desirable actions by the participants

components include the actions of research participants, which make sense only in light of the normative-descriptive framework surrounding the actions.

Seeking the Truth

> We are all engaged in truth seeking activities. When a politician suggests that we simply look at the facts—and that someone else's facts are fake—they are making a strong commitment to an objective reality. Similarly, when a psychologist reports that depressed people are more creative or that they have discovered the neural basis of semantic memory, they are making a commitment to a reality. An objective, reality is one that exists independently of human experience and can be known both formally through scientific method and informally by 'seeing it with your own eyes.' This stance toward reality is philosophical realism—the idea that things exist and have properties regardless of human minds (i.e., reality is mind-independent). (Matheson et al., this volume).

The idea of reality and commitments to that idea are expressed through controlled and controlling acts. When learning to perform such acts, we learn there are better and worse ways of representing the truth, i.e., better and worse ways of letting our acts be controlled by "matters of fact." However, control goes in both directions. While we might believe that it is only our acts that are controlled by reality, and our commitment to a faithful representation of reality, our acts of representation are themselves controlling acts. We cannot escape making pre-representational decisions, which enframe our representations of reality (Mammen, 2017), even though we might try our best to keep those decisions in the background.

Thus, it is worth asking whether and how a mere "commitment to reality" could be concealing another set of unexamined commitments that limit the kind of reality we find and describe. The unexamined commitments are akin to the theatre director's rigidity to see what unfolds on the stage in only one way, decided in advance. We understand Pfister's (this volume) proposal for bold claims, i.e., explicit assertions regarding how we decide, limit, and control the phenomena under investigation, to urge readers in this direction. Advancing dialogue between theoretical-critical and experimental psychologists would benefit from the willingness of both sides to pay attention to all important aspects of research, to what is given frontstage view in the "theater" of the laboratory and what is working behind the stage—operationalizations, interpretations, generalizations, etc.

It is possible for an experimental researcher to feel impatient about critical and theoretical inquiries, to evade any critique of the basis of their research. It is possible to feel that seeing the laboratory as a theatrical process will only slow down the happenings on the stage. Such evasions, however, are detrimental to the researcher's own work. We hope the chapters in this volume demonstrate that the disconnect between a theoretical and experimental psychology does not mean safety and freedom for experimental research. The disconnect is within the very heart of experimental research, expressed as an inattention to human experience.

We hope readers of the present volume are not only incentivized, but equipped with practical suggestions, regarding how to place experimental research within a broader view and how to think about the significance and relevance of experimental findings from a perspective that remains mindful of the totality of human-social life.

References

Bibace, R., Clegg, J., & Valsiner, J. (2009). What is in a name? Understanding the implications of participant terminology. *IPBS: Integrative Psychological & Behavioral Science, 43*(1), 67–77.

Davies, B., & Harré, R. (1990). Positioning: The discursive production of selves. *Journal for the Theory of Social Behaviour, 20*(1), 43–63.

Günther de Araujo, I. (1998). Contacting subjects: The untold story. *Culture and Psychology, 4*(1), 65–74.

Harré, R., & Secord, P. F. (1972). *The explanation of social behaviour*. Rowman & Littlefield.

Mammen, J. (2017). *A new logical foundation for psychology*. Springer.

Rosenbaum, P. J., & Valsiner, J. (2011). The un-making of a method: From rating scales to the study of psychological processes. *Theory & Psychology, 21*(1), 47–65.

Toomela, A. (2019). *The psychology of scientific inquiry*. Springer.

Valsiner, J. (1984). Cognitive socialization (book review: No five fingers are alike, by J. C. Berland, Harvard University Press, 1982). *Acta Pedologica, 1*(2), 175–178.

Valsiner, J. (2017). *From methodology to methods in human psychology*. Springer.

Index

D. Gozli, J. Valsiner (eds.), *Experimental Psychology*, Theory and History in the Human and Social Sciences, https://doi.org/10.1007/978-3-031-17053-9