Philosophy of science and the formalization of psychological theory

Eronen, Markus I.; Romeijn, Jan-Willem

Published in:
Theory & Psychology

DOI:
10.1177/0959354320969876

Publication date:
2020

Citation for published version (APA):
Eronen, M. I., & Romeijn, J-W. (2020). Philosophy of science and the formalization of psychological theory. *Theory & Psychology*, *30*(6), 786-799. https://doi.org/10.1177/0959354320969876

Theory & Psychology

Article

Theory & Psychology
2020, Vol. 30(6) 786–799

DOI: 10.1177/0959354320969876
journals.sagepub.com/home/tap
SAGE

Philosophy of science and the formalization of psychological theory

Markus I. Eronen
and Jan-Willem Romeijn
University of Groningen

Abstract

One of the original aims of this journal was to promote theory in psychology. Nowadays more and more psychological researchers are calling for more theory development, and articles on the "theory crisis" have also found their way into mainstream journals. In this article, we provide a further perspective to this theory debate. Over the past century, philosophy of science has staged extensive discussions on the mathematization of nature and on the role of mathematics in the development of theory and the connection of theory to empirical facts. We show that these discussions are highly relevant for the current debate in psychology. In particular, we emphasize the importance of conceptual work in the process of mathematization, and the role of mathematics in co-ordinating theory and observations. We then discuss the implications that these points have for statistically oriented psychology in general and for the recent theory debate in psychology.

Keywords

epistemic iteration, formalization, instrumentalism, statistics, theory

Formalizing psychology

One of the original aims of this journal was to promote theory in psychology (Gigerenzer, 2010; Stam, 2010). As we witness the 30th anniversary of *Theory & Psychology*, this goal is as important as ever: the development of theories still has a marginal role in psychology, and empirical methods, mostly of a quantitative and statistical nature, dominate the field (Borsboom, 2013; Muthukrishna & Henrich, 2019; Oberauer & Lewandowsky, 2019). However, what has changed dramatically in these 30 years is that discussions of the future of psychology now explicitly involve theory as a central theme. More and more influential researchers are calling for increased attention to theory in psychology,

---

Corresponding author:
Markus I. Eronen, Faculty of Philosophy, University of Groningen, Grote Kruisstraat 2/1, Groningen, 9712 TS, Netherlands.
Email: m.i.eronen@rug.nl

and papers advocating theory development have found their way into mainstream journals (e.g., Fiedler, 2017; Muthukrishna & Henrich, 2019; Smaldino, 2019). Also, *Theory & Psychology* continues to publish novel and important contributions to this debate (e.g., Hawkins-Elder & Ward, 2020; Klein, 2014; McGann & Speelman, 2020; Trafimow & Earp, 2016).

A central theme in these recent theory debates has been that psychological theories should be made more formal or mathematical. As has been often pointed out over the years (e.g., Meehl, 1978, 1990), psychological theories are typically verbal and vaguely formulated, and do not lead to precise predictions. This makes psychological theories difficult to falsify or test, hampering theoretical progress. Proponents of the formal turn in the development of psychological theory argue that theories should be made more precise through mathematical formulation, because in this way the assumptions and empirical implications of the theories become explicit and testable.

In this article, we provide a further perspective to this debate. Over the past century, the philosophy of science has staged extensive discussions on the mathematization of nature and on the role of mathematics in the development of theory and the connection of theory to empirical facts. These discussions are highly relevant for the current debate in psychology, but have not been adequately connected to it yet. This is what we set out to do in this paper.

First, we point out that there is a long history of arguments about the mathematization of nature, starting in astronomy and continuing on into psychological science. Next, we connect these arguments to classical philosophy of science, in particular logical empiricism, as well as more recent work by Nancy Cartwright and Hasok Chang. We emphasize the importance of conceptual work in the process of mathematization, and the role of mathematics in co-ordinating theory and observations. In the last sections, we discuss the implications that these points have for statistically oriented psychology and for the recent theory debate in psychology.

From astronomy to psychology: Uses of mathematics

It is of course impossible to fully summarize the history of the mathematization of nature here. Instead, we would like to highlight three points of specific interest for our discussion on the mathematization of psychology. The first concerns the role of mathematics as a means to capture the structure of reality, and the second and third are about the role of mathematics as mediator between theory and empirical facts, pertaining to the empirical and the conceptual adequacy of theory.

The use of mathematics in the sciences is as old as science itself. In fact its starting point lies even further back in time than modern science. Already in antiquity, natural philosophers used mathematical models to describe the movements of fixed and "wandering" stars, or as we call them now, stars and planets (cf. Hoskin, 1997a, 1997b). The idea that mathematics would be applicable to the heavens fitted nicely with a world view that positioned the Divine among the stars, far away from the messy earthly environment. However, in the premodern era the mathematical models were directed and constrained by metaphysical or theological assumptions, most importantly that the Earth is the immovable center of the universe (Hoskin, 1997a).

A major shift in scientific thinking took place when early-modern astronomers began taking mathematical structure itself more seriously, up to the point where arguments of mathematical elegance and accuracy were allowed to compete with, and eventually offset, arguments of a metaphysical or theological nature. In the work of Copernicus and later on Galileo we see that systems that place the sun in the middle were preferred because they captured the phenomena more elegantly, despite the fact that these systems violated the Biblical and Aristotelian dogma that the Earth rests and the Sun moves. The position of the Church, as formulated by Cardinal Bellarmine in his polemic with Galileo, was that there will no doubt be many mathematical descriptions of the solar system that fit the empirical facts, for example, Ptolemy's system that placed the Earth at the center and Aristarchus' heliocentric alternative, but that the choice among those equivalent systems would be up to theologians, not mathematical scientists (Hoskin, 1997b). The scientists, by contrast, considered mathematical structure to be much more than a mere instrument. They adopted a new role for mathematical structure as a means to capture not only how the world appears, but how the world is. This tension between instrumental and realistic interpretations of mathematical models is the first point we want to emphasize.

The second point concerns the role of mathematics in connecting theories to empirical observation. When mathematics took up a more central position in the new science, empirical accuracy became a point of focus. The meticulous observations of Brahe forced Kepler and other early astronomers to accept that planets could not be tracing circular orbits, but rather elliptical ones, necessitating both the mathematical models of Ptolemy and, at that point, Copernicus to be adapted (Hoskin, 1997b). The achievements of empirical precision were remarkable and revolutionary. Brahe constructed and collected extensive tables with positions of planets and stars, resulting in what in philosophy of science is called a *data model*: a corrected, organized, and to some degree idealized representation of the data (Frigg & Hartmann, 2020). Through the application of mathematics to these data models, it was then possible to describe the precise orbits of the planets. Based on these mathematical descriptions, scientific theories could be put to the test, and corrected where needed. This illustrates how, in astronomy, mathematics attained a mediating role, serving to connect data models to theory and thereby supplying the latter with precise empirical content.

For our purposes we would like to lift out a further aspect of this emergence of the sciences: the mathematization of nature was eventually extended to other domains besides astronomy. In other words, mathematical structure, with its successful use in the movements of the stars, was brought to bear on the world below the moon as well. Using the famous Newtonian illustration, the same mathematically formulated theory of gravitation was assumed to describe both the astronomical process of the Earth falling in a circular motion around the Sun, as well as the mundane event of the apple falling to the ground. Apparently, then, the mathematical structure that is inherent to the astronomical domain was also inherent to the objects that make up our immediate surroundings.

This expansion in the domain of applicability for mathematics was, certainly at the time, a daring move. How might disorganized earthly surroundings be governed by mathematical law, analogous to heavenly objects? It may seem obvious to us that mesoscopic objects like billiard balls, pendulums, and magnets can be described with mathematical laws. However, it took substantial conceptual effort to identify and characterize the elements and relations that could fruitfully be included in a mathematical description of these mesoscopic objects. For example, the development of Newton's theory of gravitation required replacing the Aristotelian concepts of natural place and natural motion with concepts such as mass and force. Moreover, Newtonian theory relied on concepts such as duration, velocity, and acceleration, which had been carefully defined and empirically grounded by earlier scientists.

It may be tempting to think that these basic concepts, used routinely in the natural sciences, have always been manifest to us, and only needed to be snapped up by theorists who would place them in the right mathematical relations to each other. But the history and philosophy of science suggest a very different picture, one in which the construction of these concepts was more than half of the scientific work. The same goes for later developments in physics, for instance with concepts like electric charge, polarity, current, and resistance. This is our third point: successful mathematical descriptions of nature rely on clearly defined and empirically grounded concepts.

In what follows our main interest is in another daring move, that of mathematizing not only the behavior of mesoscopic objects but the workings of our own minds, an environment that is, in its social and cognitive nature and in its stochasticity and variation, potentially even more messy. In this further discussion we will return to the three observations listed above, to wit, (a) the realist tendencies inherent in efforts to mathematize, (b) the role of mathematics as mediator between data and theory, and (c) the conceptual work as precondition for mathematization.

Mathematization of nature in logical empiricism

Before we engage specifically with the current debate on theory in psychology, and on the status of mathematical psychological theory in particular, it will be insightful to provide further nuance to the three points above, based on the philosophical reception, over the first half of the 20th century, of the use of mathematics in the sciences.

We consider in particular how the logical empiricists viewed the use of mathematics in the sciences, focusing on Reichenbach as a primary representative of the movement. First of all, the focus on empirical observations as opposed to theoretical structures is immediately apparent among all logical empiricists (see Creath, 2017, for an overview). The first of our three points of interest discussed in the previous section manifests in their rejection of metaphysics in general, and more specifically, in the rejection of the idea that mathematical structure represents or captures reality. In contrast to the early-modern scientists, the logical empiricists adopted a radical instrumentalist point of view, opposing any suggestion that the mathematical modeling tools were anything other than that: tools in the service of achieving accurate empirical predictions. This is where we also find the second of our three points. Mathematics was employed to achieve empirical adequacy and to mediate between theories and observations. Finally, as we will elaborate below, the logical empiricists pointed explicitly to conceptual frameworks as prerequisites for any scientific endeavor, which is where we can recognize the third of our three points of interest.

These three points are particularly salient in debates on spacetime theories, that is, theories in physics that are concerned with the structure of space and time. Before the invention of nonstandard geometry in the middle of the 19th century, there was no discussion over the structure of space: it was assumed to be captured by the definitions and postulates of Euclid. Theorems derivable in the Euclidean geometric system were *ipso facto* applicable to the physical space that we inhabit. It was almost automatic that the mathematical structure of Euclidian geometry served as a faithful description of the structure of the physical world. However, all of that changed when alternative geometries were devised, halfway through the 19th century. Even before Einstein published his revolutionary ideas on alternative geometrical structures for physical space and time, Poincaré (1902) was discussing how exactly we manage to make the mathematical structures of geometry applicable to the experiential reality of space from a more philosophical angle.

The availability of alternatives to Euclidean geometry made a whole new question about the structure of space and time possible: How are geometrical concepts, structures, and theorems relevant to the physical space around us? The answer, according to Poincaré, and later on, Reichenbach (1938, 1928/1957) as well as other logical empiricists, was that we have to *actively co-ordinate* the mathematical structure onto empirical reality. That is, we have to lay down definitions that connect key theoretical concepts to experimental procedures and measurements, so that claims cast in terms of these concepts are provided with empirical content. As an example, consider the theoretical claim that light travels along geodesics, that is, straight lines in space. This claim connects the geometric concept of a geodesic to a physical phenomenon. However, unless we define what a geodesic is by reference to a measurement procedure (i.e., give a co-ordinative definition), we have not given empirical content to this claim about light's behavior. Thus, theoretical claims cast in a mathematical format have empirical import only in virtue of co-ordinative definitions.

The above excursion on spacetime theories teaches us some important lessons on how, according to the logical empiricists, formal or mathematical theory relates to empirical reality in general. First, the picture of co-ordination between mathematical structure and empirical facts implies a degree of flexibility in how mathematical structure applies to the domain that is under investigation. To take an example from contemporary science, the same mathematical structure of network theory can be applied to railroads, brain networks, or networks of psychological symptoms (Barabási, 2012; Borsboom & Cramer, 2013). In line with this, it does not seem natural to take the mathematical structure as a representation of the target domain simpliciter. If it does so at all, it only represents under a chosen conceptual framework.

Besides casting doubt on the role of mathematics in representing, the Reichenbachian analysis highlights the constructive role that mathematics plays in identifying empirical patterns and salient concepts. If we want to apply the mathematical structures that make up our theory to the target domain, we have to specify co-ordinative definitions. This requires us to isolate the specific empirical concepts that can feature in the definitions, and make them fully precise in mathematical terms. The co-ordinative process thus involves a continuous mutual tweaking of theoretical and empirical frameworks, a process that is facilitated by the precision and transparency of mathematical structure. This supports our second and third points: mathematization offers precision in linking theoretical and empirical structures, but it requires conceptual work because the structures connect to empirical facts only via preconceived conceptual frameworks.

As an aside, we note that, at around the same time, Husserl developed very similar views on mathematical science. In the first part of his *Crisis of European Sciences* (1954), Husserl discusses mathematics as a means to describe the natural world, and points out that we must not conclude from the prominence of mathematics in the natural sciences that the natural world itself is in essence mathematical, and that our own experiences of it are merely derivative. Instead, Husserl argues, we need to give priority, ontologically and epistemically, to the world that is directly accessible to us in experience. And this holds all the more for psychological science, or for any other science concerned with the social and cultural reality that we inhabit. Of course, Husserl is usually not discussed in the context of analytic philosophy of science, mostly because he took his analysis in a wholly different direction. We too are skeptical about the phenomenological methods with which Husserl hoped to renew psychological science. But in our view, he saw correctly that mathematics is, first and foremost, a tool for describing the world, rather than being constitutive of the world itself. And he rightly directed us instead to the world that we experience, that is, the world of facts at the mesoscopic scale, and of facts about our social and cognitive lives. Moreover, he emphasized that mathematization relies on a conceptual framework that needs to be constructed in advance and then carefully matched with the world to which it can be applied.

The logical empiricist picture that we have presented in this section still needs to be refined and nuanced, especially on the third point concerning conceptual frameworks. What is not sufficiently worked out yet is the dynamic nature of the co-ordinative process, its continuous to-and-fro between the conceptual and the observational. This is what we will focus on in the next section.

Conceptual co-ordination in contemporary philosophy of science

One key point that emerged from our discussion above is that mathematization of nature requires conceptual work, and a careful co-ordination between the empirical world on the one hand, and theoretical or mathematical structures on the other. These co-ordination efforts have been an important topic in more recent philosophy of science.

Let us start with Nancy Cartwright, who throughout her career has studied the multiple steps required in linking scientific concepts to empirical facts, both in physics and the social sciences (e.g., Cartwright, 1983, 2007). In both fields, a crucial step in achieving (or improving) the validity of measurements is ensuring that the concept or construct that we aim to measure is sufficiently well defined (see also Alexandrova, 2017). And importantly, this is not achieved in one step, but is an ongoing process: "Usually, we need to start with some rough, defeasible characteristics of the concept and through a gradual back-and-forth process refine the characterization simultaneously while refining our procedures for measuring it and our claims about its relations to other concepts" (Bradburn et al., 2017, p. 76). In psychology, operational definitions of concepts have been popular, such as "IQ is what intelligence tests measure." However, such operational definition does not leave much room for gradual refinement: the concept is already set in stone, and is not responsive to theoretical or empirical developments (Bradburn et al., 2017).

In Cartwright's work it becomes very clear that the path from theoretical concepts to practical action and empirical fact is lengthy and intricate. To see exactly how long and intricate this connection is, it will be helpful to turn to another contemporary philosopher of science, Hasok Chang, who has studied the back-and-forth of conceptual and empirical co-ordination in detail. (The following is based on Chang, 2004; see also Bringmann & Eronen, 2016; van Fraassen, 2008.) He calls this back-and-forth process *epistemic iteration*, and characterizes it as the business of "getting on," where each successive stage of knowledge is building on the previous one, in the absence of any universally accepted or infallible conceptual foundations (Chang, 2016). Chang's main case study is the history of temperature measurement. Let us therefore briefly return to 16th-century science, but this time from a different perspective. Galileo and other early scientists were not only interested in astronomy, but also in phenomena of heat and cold, and made attempts to build instruments for measuring temperature. It was known since antiquity that liquids (and air) tend to expand when they are heated, and based on this simple empirical principle, it was possible to build "thermoscopes" by enclosing a liquid (or air) in a glass container or other closed vessel.

These simple instruments, in turn, made it possible to discover further robust phenomena: for example, that the temperatures at which water boils and freezes are remarkably constant across measurements. These constants could then serve as fixed points and the interval between them could be divided into units, resulting in a numerical scale. This was the beginning of the mathematization of temperature. The simple numerical scale allowed for more precise measurements, which again allowed for establishing new robust phenomena (e.g., boiling or freezing points of other liquids), which then led to better measuring instruments and a better understanding of the concept of temperature. In this way, the concept was refined in cycles of co-ordinating the concept and its mathematical structure with empirical observations. The process was also not purely empirically driven. Key advances in defining and quantifying temperature were the results of theoretical developments: for example, the absolute zero was only successfully calculated after the concept of temperature was connected to statistical physics in the 19th century. Similar kinds of epistemic iteration can be witnessed throughout the sciences, for example, in the development of classification systems of chemical kinds (Chang, 2016).

In general, the upshot of our discussion in this and the previous section is that concepts only make contact with empirical reality through a long chain of assumptions, concretizations, and heuristics, and that the mathematization of nature needs to be seen as an ongoing process of co-ordinating conceptualizations with the empirical world. In the next section, we consider what this implies for the mathematization of psychology, by looking into the statistical nature of this science, and then discussing the three points of the foregoing, that is, how mathematical structure does not itself represent but rather facilitates the connection between theory and empirical fact.

Statistification of psychology

So far, we have mostly discussed philosophical issues related to mathematization of nature in general. But in this section we turn to the specifics of mathematization in psychology. As readers of this journal know, this mathematization is not a hypothetical scenario: just like Galileo and other scientists aimed at giving the movement of stars or the phenomena of heat and cold a mathematical structure, psychologists have attempted to quantify attributes such as intelligence or memory capacity. However, the way this mathematization took place in psychology was very different from the natural sciences. Instead of formal or mathematical theories that were co-ordinated with empirical observations, it occurred via the statistification of psychology in the early 20th century.

In fact, the history of statistics and the history of psychology are closely intertwined (Hacking, 1990): many key statistical concepts were developed with psychology in mind, such as the correlation coefficient (by Francis Galton) or factor analysis (by Charles Spearman). A guiding idea of the pioneers of statistical psychology was that although human behavior at the individual level is complex, fleeting, and largely intractable, it is possible to study human populations with statistical methods and thereby get reliable results concerning human behavior (Danziger, 1990; Hacking, 1990). Early successes of this approach included the discovery (or construction) of the *g*-factor of general intelligence and discoveries of phenomena such as regression to the mean.

What made the use of statistics especially attractive for psychology was that it provided a strong argument for the scientific status of psychology, which was still very much under debate in the early 20th century (Danziger, 1990). Even though the subject matter (the human mind) remained elusive and its quantification a matter of debate, the (often successful) use of clearly defined statistical techniques that were also used in other fields gave psychology a strong claim for being a respectable science (Danziger, 1990; Michell, 1999). In the run of the 20th century, statistics came to completely dominate psychological science, which we still witness today.

This mathematization of psychology through statistics is not, however, just a success story. The (mis)use of statistics in psychology has been the subject of much criticism over the decades and up to this day (Danziger, 1990; Gigerenzer, 2004; Meehl, 1967, 1978, 1990). Most importantly, the criticism has focused on the utter dominance of one problematic method: null hypothesis significance testing (Gigerenzer, 2004; Meehl, 1978, 1990). Gigerenzer (2004) refers to this as the "null ritual" and argues that psychology is dominated by "mindless statistics." Meehl (1978, 1990) also points out (drawing from philosophy of science) that null hypothesis significance testing is exceptionally ill-suited for theory testing, but psychologists nevertheless use it for that purpose.

We agree with this criticism, but would like to emphasize here a different issue, building on our discussion in the previous sections. The null ritual and the cookbook-style application of statistical methods in general made it possible to sidestep all the difficult questions about the exact nature of the concepts that were applied (see also Danziger, 1990). As we pointed out in the previous section, concepts in psychology are often given operational definitions, which do not allow for much epistemic iteration. However, even more common in contemporary psychology is that concepts are not explicitly defined at all (Flake & Fried, in press; Flake et al., 2017). To give one recent example, the concept of "self-control" plays a key role in many areas and theories of social psychology, including famous ego-depletion experiments, but recently it has been emphasized that the concept has actually never been clearly defined, and often refers to different things in different contexts (Friese et al., 2019; Inzlicht & Friese, 2019; Lurquin & Miyake, 2017).

In other words, the sort of co-ordination or epistemic iteration we have emphasized in our discussion of the mathematization of science has hardly taken place in psychology. One important reason for this is that when statistical methods such as null hypothesis significance testing are used instead of mathematical theories, it is possible to skip the difficult steps of conceptual work and co-ordination. The statistical methods will easily produce "significant" results even when the conceptual basis is shoddy: for example, following Meehl's well-known argument (1967, 1978), a significant difference between the means of two groups can practically always be found, on any variable, because even tiny (random) differences in the means will reach statistical significance if the sample size is large enough.

There have been extensive discussions throughout the decades on construct validation and psychological measurement, but they have not resulted in sustained efforts to co-ordinate theoretical concepts with the empirical world through an iterative process (Eronen & Bringmann, in press; Flake et al., 2017; Fried & Flake, 2018). Taken together, the above considerations suggest that the mathematization of psychology via formal theories may be premature as long as the conceptual basis is not solid enough (see also Eronen & Bringmann, in press).

Statistical psychology and epistemic iteration

With this rough sketch of the use of mathematics, in particular statistics, in psychology, let us return to the three points raised and further developed earlier. On the first point we can be relatively brief. Looking at the introduction and deployment of statistical methods and representations into psychology, we can safely say that they did not take up a role as representations of reality. The Reichenbachian analysis of how hypotheses apply to the empirical domain of psychology indeed seems more apt: a lot of "co-ordinative labor" has to be carried out to match the hypotheses to empirical patterns. The theoretical structures, in this case the statistical ones, only make contact with empirical reality through a long chain of concretizations, delimitations, and rules of thumb. And the mathematical structures themselves evolve through cycles of epistemic iteration. This makes it difficult to treat them as corresponding to reality in any straightforward sense (cf. Chang, 2004; van Fraassen, 2008). Thus, as we have already seen in our discussion of the logical empiricists, the overall picture fits naturally with a broadly instrumentalist view of the role of mathematical structure in science (in contrast to, e.g., the realism of Michell, 1999, 2000).

The other two points can also be identified in the context of statistified psychology, but this requires a little more attention. How exactly does the co-ordinative work, alluded to in the foregoing, get done in statistical modeling? And do the statistical structures indeed facilitate this labor, by introducing stringency and precision? We run into an interesting paradox here. We argued that statistics helped to establish psychology as a legitimate empirical science. But precisely because psychologists were so focused on establishing empirical phenomena, they were impaired in a more speculative mode of thinking. This arguably hampered the development of new concepts that could have assisted them in identifying new and salient empirical patterns. Conceptual development is of course still possible in statistical psychology and mathematical structures are still instrumental in it. But the strict focus on the empirical has slowed it down, and we believe that this explains, for a large part, the renewed interest in theory among psychologists.

Unfortunately, we can only provide some preliminary suggestions on how statistics can perform the function of connecting psychological theory to empirical facts. We might once again take inspiration from the analysis of geometry applying to physical space, which is in a sense exemplary for how the sciences connect mathematical structures to empirical reality in general, also beyond the confines of natural science. In their discussion of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), van Loo and Romeijn (2015) rely on the Reichenbachian idea of co-ordinative definitions to clarify how disease concepts from the DSM-5 function, deploying the basic set-theoretical structure of the DSM-5. We believe that such clarifications can also be given for theories in other subdisciplines within psychology. In short, the Reichenbachian analysis of how mathematical geometry relates to physical space offers something of a blueprint for understanding how formal or mathematical theory relates to empirical reality in general.

There are other examples of how co-ordinative labor is carried out as part of statistical modeling. Descriptive statistics (e.g., principal component analysis, support vector machines) can provide “data models” that summarize and represent the data in useful ways (see second section). Statistical tools (e.g., machine learning techniques) are also efficient in finding patterns in data and thereby can help to find robust phenomena, which are essential for testing and improving theories (Eronen & Bringmann, in press; Haig, 2013). In short, if the aim is to mathematize psychology, statistics should be seen as an instrument that can help in the development of theories.

Concluding remarks on theory and psychology

Based on what we have discussed, we can draw several implications regarding the current theory debate in psychology, where we witness a call for the development of formal theories (Borsboom et al., 2020; Fried, in press; Muthukrishna & Henrich, 2019; Oberauer & Lewandowsky, 2019; Robinaugh et al., 2020; Smaldino, 2019; van Rooij & Baggio, 2020).

First, as we have pointed out, mathematization often goes together with the tendency to realistically interpret theories and mathematical structures, which we also see in the recent debates in psychology (e.g., Borsboom et al., 2020; Fried, in press; Robinaugh et al., 2020). However, the history of science suggests that one should be very careful with such interpretations. This is especially true for psychology, where there is no agreement even on the basic concepts that should be used as building blocks for formal theories. Moreover, throughout the history of the statistification of psychology, we see that mathematical structures have not taken up the role of representing reality. Instead, it is clear that a lot of co-ordinative work needs to be carried out to connect mathematical structures and psychological concepts to empirical patterns. In general, statistical and mathematical methods are best seen as tools or instruments, and not as providing representations of reality. This fits well with a broadly instrumentalist view on psychological theories.

Second, our analysis highlights the constructive role that mathematics plays in mediating between theory and the empirical domain. If we want to apply the mathematical structures of a theory to the target domain, we have to specify co-ordinative definitions. And this requires us to isolate the specific empirical concepts that can feature in the definitions, and make them fully precise in mathematical terms. In psychology, this also implies facing the challenges in quantifying psychological attributes, such as intelligence or personality traits, which have been extensively debated in the pages of *Theory & Psychology* (e.g., Michell, 2000; Trendler, 2009, 2019; see also Borsboom & Mellenbergh, 2004; Bringmann & Eronen, 2016). Advocates of formal and mathematical theories in psychology should keep these issues in mind, as such theories are likely to be successful only if the conceptual basis is sufficiently well-defined.

There are also further obstacles to developing good psychological theories, as pointed out by Meehl (1978, 1990), and as one of us has argued in another article (Eronen & Bringmann, in press). Most importantly, in psychology we do not find a broad range of widely agreed upon patterns and phenomena that can constrain theory building, or at least we have not found such an uncontroversial conceptual and empirical basis yet (see also McGann & Speelman, 2020). Moreover, finding psychological causes and mechanisms is extremely hard due to difficulties in measuring and manipulating psychological variables, and this makes the development of a formalism that can serve to express psychological theory even harder (Eronen, 2020). Therefore, although we find this renewed attention to theories important and laudable, we also believe that the current call for the development of formal and mathematical theories would merit more critical attention.

At a more general level, with this article we hope to have shown that there are interesting connections between the philosophy of science literature on mathematization of nature and the recent theory debate in psychology. We encourage others to continue the debate and explore these connections further, also in the pages of *Theory & Psychology*, which provides a much-needed venue for such interdisciplinary work.

Acknowledgements

The authors would like to thank Laura Bringmann and Freek Oude Maatman for their very helpful comments on an earlier draft of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

Alexandrova, A. (2017). *A philosophy for the science of well-being*. Oxford University Press.

Barabási, A. L. (2012). The network takeover. *Nature Physics*, *8*(1), 14–16. https://doi.org/10.1038/nphys2188

Borsboom, D. (2013, November 20). Theoretical amnesia. *Open Science Collaboration Blog*. http://osc.centerforopenscience.org/2013/11/20/theoretical-amnesia/

Borsboom, D., & Cramer, A. O. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Borsboom, D., & Mellenbergh, G. J. (2004). Why psychometrics is not pathological: A comment on Michell. *Theory & Psychology*, *14*(1), 105–120. https://doi.org/10.1177/0959354304040200

Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A., & Haig, B. D. (2020, September 22). Theory construction methodology: A practical framework for theory formation in psychology. *PsyArXiv Preprints*. https://doi.org/10.31234/osf.io/w5tp8

Bradburn, N. M., Cartwright, N. L., & Fuller, J. (2017). A theory of measurement. In L. McClimans (Ed.), *Measurement in medicine: Philosophical essays on assessment and evaluation* (pp. 73–87). Rowman & Littlefield.

Bringmann, L. F., & Eronen, M. I. (2016). Heating up the measurement debate: What psychologists can learn from the history of physics. *Theory & Psychology*, *26*(1), 27–43. https://doi.org/10.1177/0959354315617253

Cartwright, N. (1983). *How the laws of physics lie*. Oxford University Press.

Cartwright, N. (2007). *Hunting causes and using them: Approaches in philosophy and economics*. Cambridge University Press.

Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford University Press.

Chang, H. (2016). The rising of chemical natural kinds through epistemic iteration. In C. Kendig (Ed.), *Natural kinds and classification in scientific practice* (pp. 53–66). Routledge.

Creath, R. (2017, April 5). Logical empiricism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2020 ed.). Stanford University. https://plato.stanford.edu/archives/sum2020/entries/logical-empiricism/

Danziger, K. (1990). *Constructing the subject*. Cambridge University Press.

Eronen, M. I. (2020). Causal discovery and the problem of psychological interventions. *New Ideas in Psychology, 59*, Article 100785. https://doi.org/10.1016/j.newideapsych.2020.100785

Eronen, M. I., & Bringmann, L. F. (in press). The theory crisis in psychology: How to move forward. *Perspectives on Psychological Science*.

Fiedler, K. (2017). What constitutes strong psychological science? The (neglected) role of diagnosticity and a priori theorizing. *Perspectives on Psychological Science*, *12*(1), 46–61. https://doi.org/10.1177/1745691616654458

Flake, J. K., & Fried, E. I. (in press). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*.

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, *8*(4), 370–378. https://doi.org/10.1177/1948550617693063

Fried, E. I. (in press). Lack of theory building and testing impedes progress in the factor and network literature. *Psychological Inquiry*.

Fried, E. I., & Flake, J. K. (2018). Measurement matters. *APS Observer*, *31*(3), 29–31.

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2019). Is ego depletion real? An analysis of arguments. *Personality and Social Psychology Review*, *23*(2), 107–131. https://doi.org/10.1177/1088868318762183

Frigg, R., & Hartmann, S. (2020, Februrary 4). Models in science. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2020 ed.). Stanford University. https://plato.stanford.edu/archives/spr2020/entries/models-science/

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, *33*(5), 587–606. https://doi.org/10.1016/j.socec.2004.09.033

Gigerenzer, G. (2010). Personal reflections on theory and psychology. *Theory & Psychology*, *20*(6), 733–743. https://doi.org/10.1177/0959354310378184

Hacking, I. (1990). *The taming of chance*. Cambridge University Press.

Haig, B. D. (2013). Detecting psychological phenomena: Taking bottom-up research seriously. *The American Journal of Psychology*, *126*(2), 135–153. https://doi.org/10.5406/amerjpsyc.126.2.0135

Hawkins-Elder, H., & Ward, T. (2020). Theory construction in the psychopathology domain: A multiphase approach. *Theory & Psychology*, *30*(1), 77–98. https://doi.org/10.1177/0959354319893026

Hoskin, M. (1997a). Astronomy in antiquity. In M. Hoskin (Ed.), *The Cambridge illustrated history of astronomy* (pp. 22–47). Cambridge University Press.

Hoskin, M. (1997b). From geometry to physics: Astronomy transformed. In M. Hoskin (Ed.), *The Cambridge illustrated history of astronomy* (pp. 98–141). Cambridge University Press.

Husserl, E. (1954). *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie* [The crisis of European sciences and transcendental phenomenology: An introduction to phenomenological philosophy]. Martinus Nijhoff.

Inzlicht, M., & Friese, M. (2019). The past, present, and future of ego depletion. *Social Psychology*, *50*(5–6), 370–378. https://doi.org/10.1027/1864-9335/a000398

Klein, S. B. (2014). What can recent replication failures tell us about the theoretical commitments of psychology? *Theory & Psychology*, *24*(3), 326–338. https://doi.org/10.1177/0959354314529616

Lurquin, J. H., & Miyake, A. (2017). Challenges to ego-depletion research go beyond the replication crisis: A need for tackling the conceptual crisis. *Frontiers in Psychology*, *8*, Article 568. https://doi.org/10.3389/fpsyg.2017.00568

McGann, M., & Speelman, C. P. (2020). Two kinds of theory: What psychology can learn from Einstein. *Theory & Psychology*, *30*(5), 674–689. https://doi.org/10.1177/095935432037804

Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, *34*(2), 103–115. https://doi.org/10.1086/288135

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, *46*(4), 806–834. https://doi.org/10.1037/0022-006X.46.4.806

Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports*, *66*(1), 195–244. https://doi.org/10.2466/pr0.1990.66.1.195

Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept*. Cambridge University Press.

Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory & Psychology*, *10*(5), 639–667. https://doi.org/10.1177/0959354300105004

Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour*, *3*, 221–229. https://doi.org/10.1038/s41562-018-0522-1

Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, *26*(5), 1596–1618. https://doi.org/10.3758/s13423-019-01645-2

Poincaré, H. (1902). *La science et l'hypothèse* [Science and hypothesis]. Flammarion.

Reichenbach, H. (1938). *Experience and prediction: An analysis of the foundations and the structure of knowledge*. University of Chicago Press.

Reichenbach, H. (1957). *The philosophy of space and time* (Reichenbach, M., & Freund, J., Trans.). Dover Publications. (Original work published 1928)

Robinaugh, D., Haslbeck, J. M. B., Waldorp, L., Kossakowski, J. J., Fried, E. I., Millner, A., McNally, R., van Nes, E. H., Scheffer, M., Kendler, K. S., & Borsboom, D. (2020, July 18). Advancing the network theory of mental disorders: A computational model of panic disorder. *PsyArXiv Preprints*. https://doi.org/10.31234/osf.io/km37w

Smaldino, P. (2019, November 6). Better methods can't make up for mediocre theory. *Nature*, *575*(7781), Article 9. https://doi.org/10.1038/d41586-019-03350-5

Stam, H. J. (2010). Theoretical communities and *Theory & Psychology*: A decade review. *Theory & Psychology*, *20*(6), 723–731. https://doi.org/10.1177/0959354310391871

Trafimow, D., & Earp, B. D. (2016). Badly specified theories are not responsible for the replication crisis in social psychology: Comment on Klein. *Theory & Psychology*, *26*(4), 540–548. https://doi.org/10.1177/0959354316637136

Trendler, G. (2009). Measurement theory, psychology and the revolution that cannot happen. *Theory & Psychology*, *19*(5), 579–599. https://doi.org/10.1177/0959354309341926

Trendler, G. (2019). Conjoint measurement undone. *Theory & Psychology*, *29*(1), 100–128. https://doi.org/10.1177/0959354318788729

van Fraassen, B. (2008). *Scientific representation: Paradoxes of perspective*. Oxford University Press.

van Loo, H. M., & Romeijn, J.-W. (2015). Psychiatric comorbidity: Fact or artifact? *Theoretical Medicine and Bioethics*, *36*(1), 41–60. https://doi.org/10.1007/s11017-015-9321-0

van Rooij, I., & Baggio, G. (2020, February 28). Theory before the test: How to build high-verisimilitude explanatory theories in psychological science. *PsyArXiv Preprints*. https://doi.org/10.31234/osf.io/7qbpr

Author biographies

Markus I. Eronen is an assistant professor at the Faculty of Philosophy, University of Groningen, the Netherlands, and an associate editor of *Theory & Psychology*. His current research is focused on the nature of theories and levels of organization in psychology, and recent publications include (with L. Bringmann) “The Theory Crisis in Psychology: How to Move Forward,” in *Perspectives on Psychological Science* (in press).

Jan-Willem Romeijn is a professor of philosophy of science at the Faculty of Philosophy, University of Groningen, the Netherlands. His research focuses on general philosophy of science, statistical methods, and social epistemology. Among his recent publications is (with H. M. van Loo & K. S. Kendler) “Changing the Definition of the Kilogram: Insights for Psychiatric Disease Classification,” in *Philosophy, Psychiatry and Psychology* (2019).