

“Are psychological attributes quantitative?” is not an empirical question: Conceptual confusions in the measurement debate

Theory & Psychology
2022, Vol. 32(1) 131–150
© The Author(s) 2021



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09593543211045340
journals.sagepub.com/home/tap



David J. Franz 
University of Würzburg

Abstract

Critics of psychological measurement have accused quantitative psychologists of ignoring the empirical hypothesis that psychological phenomena are quantitative (Michell), or have claimed that it is impossible in principle to find out whether psychological phenomena are actually quantitative (Trendler). By drawing on Bennett and Hacker (2003), I argue that both criticisms do not go far enough because they sidestep the fundamental conceptual problem of the measurement debate: It is impossible to give concrete formulations of the question “Are psychological attributes quantitative?” without transgressing the boundaries of meaningful language. Conceptual confusions and questionable philosophical assumptions have contributed to the misguided idea that the quantity of psychological phenomena must or can be demonstrated empirically. First, the measurement debate is characterized by misleading examples and ambiguous terminology. Second, the idea of psychological measurement is inherently Cartesian. In summary, psychological measurement is even more problematic than Michell and Trendler have argued.

Keywords

Cartesianism in psychology, conceptual analysis in psychology, philosophical foundations of quantitative psychology, psychological measurement, quantification

Quantification is the foundation of current empirical psychology. Psychological research objects (e.g., emotions, attitudes, personality, skills) are translated into numbers and these numbers are statistically analyzed. However, since the introduction of quantitative

Corresponding author:

David J. Franz, Institut für Psychologie, Universität Würzburg, Lehrstuhl IV, Röntgenring 10, Würzburg, 97070, Germany.

Email: david.franz@uni-wuerzburg.de

methods to psychology, the practice of quantification has always been the target of criticism (Michell, 1997, 2000). Beyond that, despite the widespread acceptance and usage of quantitative methods in psychology today, the debate about potential downsides of quantification continues and has grown considerably in the last 25 years (e.g., Barrett, 2003, 2008; Borsboom & Mellenbergh, 2004; Guyon et al., 2018; Humphry, 2017; Lamiell, 2010, 2013; Maraun, 1998; Martin, 2003; Maul et al., 2016; Michell, 1997, 2000; Sherry, 2011; Slife & Williams, 1997; Tafreshi et al., 2016; Trendler, 2009; Uher, 2018, 2019; van Lill, 2012; Westerman, 2006a, 2006b, 2014; Yanchar, 2006).

The central question in the debate about quantification is whether psychological phenomena, such as intelligence, personality, emotions, and attitudes, are measurable at all. Since concepts of psychological measurement discussed in this debate are inspired by measurement in natural sciences such as physics, answers to this question boil down to exploring whether psychological phenomena are quantitative just as physical attributes are. Two influential positions in this debate that are highly critical of quantitative psychology¹ are those put forward by Joel Michell (1997, 2000, 2003, 2010, 2011, 2012, 2020) and Günter Trendler (2009, 2013, 2019a, 2019b). Michell criticizes quantitative psychologists for assuming that psychological attributes are quantitative without attempting to test this hypothesis. Trendler goes a step further by arguing that testing this hypothesis is impossible and that there can be no psychological measurement.

The purpose of this article is to show that—contrary to widespread assumptions in the measurement debate—the question “Are psychological attributes quantitative?” is not an empirical question. The prevalent interpretation of this question and attempts to answer it both rest on conceptual confusions and questionable philosophical assumptions. Searching for an empirical equivalent of mathematical relations in psychological phenomena is a misguided endeavor. My arguments are highly inspired by the work of Bennett and Hacker (2003).² While Bennett and Hacker’s conceptual analyses are primarily concerned with the meaninglessness of ascribing psychological attributes to the brain, I will show that their conceptual approach can provide valuable insights for the psychological measurement debate.

The article is structured as follows. First, I provide a brief sketch of Michell’s and Trendler’s criticism of quantitative psychology. Second, I show that central to their criticism and the measurement debate is the question of whether psychological phenomena satisfy Hölder’s axioms. This question, however, is conceptually confused, as I will argue in the subsequent section. After that, I will analyze the core conceptual confusions and philosophical assumptions that lie beneath the measurement debate: misleading examples, ambiguous terminology, and a hidden Cartesianism. I conclude with an expansion of Michell’s and Trendler’s challenges for psychometricians.

Criticizing quantitative psychology’s foundations

Joel Michell has characterized quantitative psychology as “pathological science” (2000, p. 639) and as suffering from “methodological thought disorder” (1997, p. 374). According to Michell, quantitative psychology rests on a double failure. The first failure is psychologists assuming that psychological attributes are quantitative without putting this assumption to empirical test. The second failure, which makes psychological

research pathological, is that psychologists do not realize that most of their research is based on this untested empirical hypothesis. Michell has identified three fundamental methodological and philosophical beliefs that maintain the pathology of psychological research. First, quantitative psychology is pervaded by “Pythagoreanism” (2011, p. 244): the deep-seated belief that everything that exists is quantitatively structured. Second, when psychologists bring forward arguments for a Pythagorean view on psychological attributes, they oftentimes commit the fallacy of inferring quantity from order (Michell, 2012). As Michell points out, the fact that psychological phenomena are ordered (e.g., that pain can be more or less intense) does not logically entail the conclusion that these phenomena are quantitative (e.g., that pain is numerical). Third, Michell (2003) identifies the “quantitative imperative” (p. 5) as psychology’s fundamental methodological commitment. This imperative dictates that only research that is based on quantitative measurement can be considered real science. However, not all research programs in psychology disguise the fundamental hypothesis of quantitative psychological attributes. According to Michell (1997, 2000), the theory of conjoint measurement, which quantitative psychologists hardly ever rely on, can provide the tools for constructing empirical tests of the quantity assumption.

Günther Trendler is influenced by Michell’s arguments, but he is even more critical of quantitative psychology (Trendler, 2009, 2013, 2019a, 2019b). Trendler (2009) confronts quantitative psychologists with the “Milleian quantity objection” (p. 589). This objection states that psychological phenomena “are neither manipulable nor are they controllable to the extent necessary for an empirically meaningful application of measurement theory” (p. 592). Trendler’s argument for this objection is complex, but the basic idea is that psychological phenomena cannot be manipulated with such precision that one can be sure that equal amounts of a psychological phenomenon are related to equal measurement values. Trendler (2009) discusses the example of motivation: in order to show that motivation can be measured quantitatively, one must show that equal amounts of motivation lead to equal measurements of motivation within the limits of measurement error. To establish this relation, one has to implement equal levels of motivation in one person at various points in time or in different persons simultaneously. However, this is not possible because one would have to shield the participant’s motivation from all systematic influences and one would have to manipulate solely the motivation by the exact same amount. Therefore, psychological attributes are not measurable. It is important to note that Trendler does not argue that psychological attributes are not quantitative. He rather argues that testing this hypothesis is impossible and that we, therefore, will never find out if psychological phenomena are quantitative or not (Trendler, 2013).

Both Michell and Trendler contend that it is an empirical question whether psychological attributes are quantitative. Michell (2000) characterizes “the hypothesis that psychological attributes are quantitative” as a “basic, empirical hypothesis” (p. 650). Trendler (2009) explicitly agrees with Michell in writing that “quantitative structure can be ascribed to an attribute only if it empirically satisfies the conditions of quantity” (p. 582). Moreover, at least one other critic of quantitative psychology shares this assumption (Barrett, 2003, 2008), and several scholars that defend quantitative psychology claim that empirical research must and can prove psychological phenomena to be quantitative (Borsboom & Mellenbergh, 2004; Kyngdon, 2008, 2013; Saint-Mont, 2012).³

What does it mean to ask whether psychological attributes are quantitative?

The question whether psychological attributes are quantitative is synonymous with the question whether psychological attributes are (at least) interval scaled since real measurement and many statistical operations require (at least) an interval scale. If one conceives of measurement as a homomorphic mapping of empirical relations onto numerical relations, then a numerical interval scale can only be justified if this scale is the numerical equivalent of an interval scaled empirical relation. In other words, if numerical relations ought to capture empirical relations that exist between different psychological phenomena, then it must be possible to describe these empirical relations numerically. Since numerical relations have the minimum requirement of being interval scaled, the question is whether psychological phenomena can be conceived of as being interval scaled.

To justify an interval scale, not only the relations of equality (= the central feature of a nominal scale) and order (= the central feature of an ordinal scale) must be present, but also the relation of additivity. The relation of additivity has been described comprehensively by Otto Hölder's axioms, which I cite in the words of Joel Michell (1997):

A range of instances of an attribute, Q , constitutes a continuous quantity if and only if the following five conditions obtain (in each case an attempt has been made to state first a more accessible explanation of what the condition means, free of mathematical symbols and technical terms).

1. Any two magnitudes of the same quantity are either identical or different and, if the latter, there must exist a third magnitude, the difference between them, i.e. for any a and b in Q , one and only one of the following is true: (i) $a = b$, (ii) there exists c in Q such that $a = b + c$, and (iii) there exists c in Q such that $b = a + c$;
2. A magnitude entirely composed of two discrete parts is the same regardless of the order of composition, i.e. for any a and b in Q , $a + b = b + a$;
3. A magnitude which is a part of a part of another magnitude is also a part of that same magnitude, the latter relation being unaffected in any way by the former, i.e. for any a , b and c in Q , $a + (b + c) = (a + b) + c$;
4. For each pair of different magnitudes of the same quantity there exists another between them, i.e. for any a and b in Q such that $a > b$, there exists c in Q , such that $a > c > b$; and
5. Given any two sets of magnitudes, an 'upper' set and a 'lower' set, such that each magnitude belongs to either set but none to both and each magnitude in the upper set is greater than any in the lower, there must exist a magnitude no greater than any in the upper set and no less than any in the lower, i.e. every non-empty subset of Q that has an upper bound has a least upper bound. (p. 357)

Several scholars in the measurement debate refer to these axioms as central indicators of quantity (Barrett, 2003, pp. 422–423; Kyngdon, 2013, p. 233; Michell, 2000, p. 650, 2010, pp. 48–49; Trendler, 2009, pp. 581–582).

At this point, it is possible to summarize the central topic of the measurement debate. The empirical question "Are psychological attributes measurable?" is equated by scholars

with the question “Are psychological attributes quantitative?” This question, in turn, can be put into concrete terms by asking whether psychological attributes satisfy Hölder’s axioms.

Conceptual confusions and misleading questions

Many scholars debate whether, or to what extent, different measurement models, such as classical test theory, Rasch modeling, or conjoint measurement, are suited for testing empirically whether psychological phenomena satisfy Hölder’s axioms (Barrett, 2003; Borsboom & Mellenbergh, 2004; Humphry, 2017; Kyngdon, 2008, 2013; Michell, 2000; Saint-Mont, 2012; Trendler, 2009). In contrast, I do not think that more complex statistics and well thought out experiments are needed, but more conceptual clarifications. The main reason for this is that, as Michael Maraun (1998) has pointed out, psychology’s conceptual foundation largely consists of “common-or-garden concepts” (p. 454).⁴ The central concepts of psychology, such as fear, attitude, character trait, or motivation, are taken from everyday social practices and, thereby, their meaning depends on these practices. This is a key difference between the common-or-garden concepts that lie at the heart of psychological research and the technical concepts of physics, which have a precise meaning that is established by an expert community through explicit definition.

In marked contrast to technical concepts, common-or-garden concepts are not developed, laid down or modified at the outset of empirical investigation. This is because these concepts already have meanings, as manifest in their everyday use, use being governed by grammar. Hence, there exist grammatical restrictions on what one may legitimately do with them. . . . it is not the case that common-or-garden concepts *must* provide the conceptual foundation for empirical work in psychology, but merely that if the phenomena they denote are to be the focus of investigation, coherent empirical work necessitates that they be employed correctly. For when the meaning of a concept is subverted, the link between the phenomena and the concept that was supposed to denote them is severed: The denotational link is not established. (Maraun, 1998, p. 454)

While physicists determine the meaning of their discipline’s central concepts, psychologists cannot do that because the meaning of most psychological concepts is dictated by the commonplace use of these concepts. If one wants to understand what is meant by important psychological concepts, such as fear, attitude, character trait, or motivation, one has to reflect on the way that “fear,” “attitude,” “character trait,” and “motivation” are used in our daily lives. Because common-or-garden concepts are central to psychology, the interpretation of psychological research depends fundamentally on the meaning of these concepts. From this follows that conceptual clarification is of vital importance for a coherent understanding of psychological research in general and for debates about psychological measurement in particular. Neglecting conceptual questions can result in misdirected research efforts and conceptually confused interpretations of empirical research results. As Bennett and Hacker (2003) aptly point out:

Distinguishing conceptual questions from empirical ones is of first importance. When a conceptual question is confused with a scientific one, it is bound to appear singularly refractory. It seems in such cases as if science should be able to discover the truth of the matter under investigation by theory and experiment—yet it persistently fails to do so. . . . Furthermore, when empirical problems are addressed without adequate conceptual clarity, misconceived questions are bound to be raised, and misdirected research is likely to ensue. (p. 2)

I think that this warning points to a central problem of the measurement debate. The persistent difficulties of psychological measurement are not the result of insufficient statistics, imperfect measurement models, or limits of empirical investigations. They are the result of a lack of conceptual clarity (Maraun, 1998). The main conceptual problem of the measurement debate is that it is impossible to give a detailed formulation of Hölder's axioms in regard to psychological phenomena without transgressing the boundaries of meaningful language. Therefore, the question whether psychological attributes satisfy Hölder's axioms has no meaning. Consider three examples of concretizations of Hölder's axioms:

1. If I add the anxiety I had this morning to the anxiety I had this afternoon, I get exactly my current anxiety (cf. axiom 1, ii).
2. Adding my opinion about the deterrent effect of the death penalty to my opinion about the moral permissibility of the death penalty leads to the same result as a reverse addition (cf. axiom 2).
3. On one occasion, I first added my motivation for regular physical exercise to my motivation to quit smoking and then I added my motivation for a healthy diet. On another occasion, I first added my motivation for a healthy diet to my motivation to quit smoking and then I added my motivation for regular physical exercise. On both occasions, I got the exact same result (cf. axiom 3).

From Michell's and Trendler's statements that it is an empirical question whether Hölder's axioms hold true for psychological phenomena follows that all three examples should be regarded as empirical claims. It also follows that it is up to empirical research to explore whether the three examples are true (i.e., whether the "calculations" in these examples are correct). However, it is impossible to investigate these claims empirically because they have no meaning. Speaking of the addition of emotions, motivations, or opinions is misconceived. It is entirely unclear what it means to sum mental phenomena and it is, therefore, equally unclear what it means to arrive at the same computational result in different "additions of mental phenomena."

In general, it is easy to demonstrate what it means to perform an addition. Just write the numbers down and perform each step of the calculation, for example, by using columnar addition. For this purpose, you first write down the single digits, tens digits, and hundreds digits in one column each, then add the numbers in one column from right to left and, if this sum turns out to be a two digit number, write down only the single digit while adding the tens digit to the next column on the left. If you further want to explain how exactly you add the single numbers in one column, you can illustrate this by simply counting objects that serve as examples. None of this can be done with psychological

phenomena. There is no meaning to the notion of adding single instances of opinions by counting them and illustrating that they add up to a sum. I cannot write down my anxieties from different points in time and make sure that the single digits and tens digits of different anxieties stand in one column. Operations of addition are impossible with psychological phenomena because there is no way of verbalizing these operations meaningfully. The language community has not assigned any meaning to phrases like “counting up opinions to a sum,” “writing anxieties in one column,” or “the single digits of my anxiety.”

Similar considerations apply if one does not interpret Hölder’s axioms in a strict mathematical sense. For example, Michell (1997) interprets the “+” in the axioms not as denoting mathematical addition but as denoting “a relation between the magnitudes *a*, *b* and *c*. . . . The relation I have in mind is this: magnitude *c* is entirely composed of discrete parts, magnitudes *a* and *b*” (p. 357). This alternative reading also leads to sentences without meaning: “My current anxiety is entirely composed of the anxiety I had this afternoon and the anxiety I had this morning” or “I can divide my motivation to write this paper into discrete parts of motivation that can be recombined to my overall motivation.” The point is not that we lack the empirical methods to validate these claims. The point is that there is no way of explaining the meaning of phrases like “discrete parts of motivation” or “the composition of my anxiety.”

It might be objected that we are simply faced with the limits of our everyday psychological vocabulary and that sophisticated empirical research is needed to reveal what really lies behind the mere words we use to talk about psychological phenomena. While everyday language might not be able to explore the additivity of psychological phenomena, modern measurement models (e.g., Rasch modeling or conjoint measurement) and sophisticated experiments are suited for this task. However, it is misleading to think that there are fundamental shortcomings of our everyday psychological language that can be compensated for by empirical methods (for a comprehensive analysis of arguments that draw on alleged shortcomings of everyday language see Bennett & Hacker, 2003, pp. 74–81, 378–381). The main reason for this is that

whether a putative hypothesis makes sense depends upon the meanings—that is, the correct uses—of the words that formulate it. The meanings of words are determined by their rule-governed use, and they are given by what are accepted as correct explanations of community of speakers. For explanations of meaning function as rules or standards for the correct use of the expression concerned. (Bennett & Hacker, 2003, p. 382)

Questions of meaning are prior to empirical questions. A clarification of the meaning of the term “quantitative psychological phenomena” has to be given before we can empirically investigate the measurability of a psychological phenomenon. One can only interpret the results of empirical attempts to prove that psychological phenomena are quantitative if one knows what is meant by “discrete parts of motivation,” “the result of adding up my different anxieties,” or “the composition of my joy.”

The vocabulary we use to describe psychological phenomena is only meaningful insofar as it is embedded in everyday social practice (Maraun, 1998). Common language use constitutes the meaning of terms like “anxiety,” “attitude,” or “personality.” If one

wants to know what, for example, the meaning of the term “pain” is, one has to pay close attention to how competent speakers use it, to the contexts in which the term is used, and to the various other terms that are related to it (e.g., suffering, screaming, etc.). It is not possible to deprive the term “pain” (or other psychological terms) of the social practice of language use without depriving the term of its meaning.

Psychologists cannot avoid relying on our everyday psychological vocabulary because results of empirical research in psychology are completely uninformative without an interpretation that makes use of this vocabulary (Westerman, 2006a, 2006b, 2011; Yanchar, 2006). Knowing that an empirical study on the effectiveness of anxiety therapy yielded a Cohen's $d = .45$ in favor of the experimental group does not tell one anything of interest as long as one is not fundamentally familiar with the social practice of using the term “anxiety.” However, as a competent language user, one could interpret the result as suggesting that the therapy effectively leads to a reduction of anxiety. Interpreting the result in this way means relying fundamentally on everyday anxiety talk. Consequently, if researchers reacted to the points I raised above by claiming that their goal was to get rid of our deficient everyday psychological vocabulary through the use of statistics and empirical research methods, they would actually deprive themselves of the possibility of making sense of any psychological research.

In summing up my arguments so far, I first want to emphasize the distinct value of Michell's and Trendler's works.⁵ Both have put forward considerable challenges for psychological researchers that logically follow from the assumption that psychological measurement can be modeled after physical measurement. Michell and Trendler are right in pointing out that if the measurability of psychological phenomena actually was an empirical question, then psychologists would have to show empirically that the Hölder axioms hold true for psychological phenomena. They are also right in asserting that psychologists have largely ignored this task, despite many psychologists believing that psychological measurement can emulate measurement in the natural sciences. Furthermore, Michell's and Trendler's skepticism about the possibility of bringing forward empirical proof for the measurability of psychological phenomena is a consistent elaboration of the assumption that the additive relationships of numerical scales are mappings of empirical relationships between psychological phenomena. If it made sense to conceive of psychological measurement as an empirical issue, it probably would be very hard or even impossible to prove empirically that certain psychological phenomena actually are quantitative. Michell and Trendler have thought through the idea of psychological measurement in a much more consistent manner than most psychologists have.

Nevertheless, both Michell and Trendler do not question psychometricians' problematic assumption that investigating the measurability of psychological phenomena is an empirical issue. As my analysis above has shown, however, the question “Are psychological attributes quantitative?” is not an untested or untestable empirical hypothesis; it is a linguistic deception. From this follows that the reason why there is and cannot be a psychological answer to the question whether “attributes are quantitative” is not that it is very hard or impossible to bring forward empirical evidence that might answer that question. The reason is that this question loses its meaning entirely when it is transferred from the physical sciences to psychology.

The roots of the conceptual confusions

To advance the measurement debate, it is vital to understand the philosophical roots that lie beneath it. In this section, I argue that misleading examples, ambiguous terminology, and a hidden Cartesianism all contributed to the dubious belief that the quantity of psychological phenomena must be demonstrated empirically.

Misleading examples

It is striking that many arguments in the measurement debate revolve around examples of measurement of physical properties, such as hardness (Michell, 2010, pp. 50, 53; 2012, p. 256), density (Barrett, 2003, p. 428), temperature (Sherry, 2011), length (Kyngdon, 2013, pp. 233–234), and tension (Trendler, 2009, pp. 581–587). In addition, most authors oftentimes only speak in abstract terms of “psychological attributes” and avoid discussing the semantics of psychological phenomena. Scholars do not try to give meaningful interpretations of the Hölder axioms for psychological phenomena.

I do not want to suggest that trying to learn from physics about successful measurement is a bad idea. However, I think it is misleading to focus heavily on the measurement of physical properties because, thereby, the logic of language of psychological phenomena is disregarded. Concentrating on examples of physical measurement without paying close attention to the conceptual issues surrounding psychological phenomena has contributed to the misleading question that lies at the heart of the measurement debate.

Ambiguous terminology

Closely related to the point about misleading examples is the fact that the widespread usage of the broad term “psychological attributes” is itself problematic. Of course, it is difficult to come up with an unambiguous term denoting all phenomena that can be the object of psychological research. However, the vagueness of the term “attribute” renders further conceptual errors likely.

On the one hand, “attributes” can be understood as attributions or ascriptions that we make in the course of psychological talk. In this interpretation, attributes denote the mental predicates that we use to describe humans or animals (e.g., “anxious,” “motivated,” “convinced”) or the speech acts of ascribing such predicates (e.g., the act of saying “She is very anxious right now”). Certainly, it would be beneficial to interpret “attributes” in this way because such an understanding would point to the importance of conceptual analyses. However, the lack of such analyses in the measurement debate renders it unlikely that the disputants understand the term in this way.

On the other hand, “attribute” can be read as a synonym for “property,” “characteristic,” “feature,” or “trait,” and mean the property of a certain object. Of course, this interpretation is legitimate. Nevertheless, it can be very deceptive since it misleadingly suggests that psychological phenomena should be understood as analogous to properties of physical systems. It might be tempting, at this point, to insist that the central question of the measurement debate is exactly whether it is empirically justified to treat

psychological phenomena like physical properties. This, however, is conceptually dubious for several reasons.

At first, many psychological phenomena, like emotions, opinions, wishes, motivation, dreams, and worries, are not properties of a person. To say that Peter has the property of being anxious right now or that he has the property of being convinced of the moral permissibility of the death penalty is no more than an unnecessarily complicated way of saying that Peter is anxious or that he thinks the death penalty can be justified. In many cases, the addition of the term “property” to ordinary psychological talk does not come with more meaning or with a more nuanced description. It is superfluous at best and misleading at worst because speaking of psychological phenomena as properties bears the risk of conceiving sensations, emotions, opinions, wishes, motivation, and so forth as mysterious unobservable entities that exist hidden inside animals and humans. Such a Cartesian conception of the mind is a philosophical position—not an empirical hypothesis—and it is highly questionable (see next section).

Secondly, even in cases of psychological phenomena in which talking of properties can be reasonable, there are important conceptual differences between physical and psychological properties (Ryle, 1949/2000). Our ordinary talk of character “traits” certainly points to a close conceptual relationship with the term “properties.” Nevertheless, it is vital to pay close attention to the semantics of character talk. If we ascribe a certain personality to another person, we speak of constant tendencies of this person to act, think, and feel in a certain way under various circumstances. When we ascribe Peter an aggressive personality, we mean that Peter very likely reacts with aggressive statements or actions to different events that can trigger aggression. Personality traits are to be understood as dispositions, behavioral tendencies, or inclinations. In contrast, the length of a rod is not the rod’s behavioral tendency and the temperature of a gas is not a disposition to act, feel, or think in a certain way. Similarly, while intelligence can be seen as a person’s property (e.g., “A special feature of Peter is that he is very intelligent”), the conceptual differences from physical properties are considerable. Speaking of a highly intelligent person means to ascribe various intellectual capabilities to that person, for example, the capability of solving complex problems, understanding complicated interrelations, or learning difficult contents in a short period of time. The length of a rod, however, is not a capability of the rod and the temperature of a gas is not a special skill the gas possesses.

One has to pay close attention to all the conceptual distinctions just pointed out when one talks about “psychological attributes.” The fact that many disputants in the measurement debate are focusing on the measurement of physical properties is clear evidence that decisive conceptual distinctions between physical properties and psychological attributes are overlooked.

Finally, in an attempt to justify the uncritical equation of psychological and physical properties, one might argue for some kind of reductionism. According to this objection, psychological attributes can be reduced completely to physical properties (e.g., certain states of the brain or nervous system). Consequently, there is no fundamental difference between psychological and physical properties. It is not necessary to spell out the details of this reductionist point of view (e.g., whether the reduction is semantic or ontological). The main point is that reductionism is a philosophical position—and not an empirical

hypothesis—that has been met with illuminating philosophical critique by scholars from different schools of thought (e.g., Bennett & Hacker, 2003; Bergner, 2016; Fuchs, 2011; Nagel, 1979, 1989). Consequently, a reductionist analysis of psychological phenomena has to be defended by philosophical arguments and not by empirical investigations into the quantity of psychological phenomena. Beyond that, depending on the specific version of reductionism one prefers, psychological research and, thus, a debate about psychological measurement seems to be superfluous. At least, the reductionist has to give extensive arguments why there is still need for psychological research if psychological phenomena are no more than brain cells firing. Consequently, a reductionist analysis of the mind does not seem to be a promising position for scholars in the measurement debate.

Hidden Cartesianism

Cartesian thinking is ubiquitous in different branches of quantitative psychology and neuroscience (Bennett & Hacker, 2003; Westerman & Steen, 2007), and it can also be found in the works of some theoretical psychologists (Westerman, 2014). Westerman and Steen (2007) give a helpful characterization of Cartesian thinking in psychology:

The [Cartesian] framework is based on a split between the subject (identified with mind), on the one hand, and everything else, on the other, including body, material objects, and other people. The person is a thinker, or spectator . . . , who reflects on the world from a distance. Material things are meaningless contents, or facts, essentially unrelated to one another except insofar as Mind finds abstract meanings behind them. This framework provides psychology with its views on basic substantive matters by suggesting that perceptions, cognitions, affects, and goals are the “inner” processes of a removed subject, that these processes are fundamentally different in kind and isolable from “outer” events and behavior, and that psychological phenomena can be explained by putting together accounts made up of terms from the two sides of the polarity. (p. 326)

The idea of psychological measurement is Cartesian in nature because it implies that “outer” measurement results are mappings of “inner” psychological phenomena. According to widespread terminology, psychological measurement instruments gauge the “internal,” “hidden,” or “unobservable” psychological phenomena that “underlie,” “influence,” or “cause” overt behavior. For example, according to one textbook on personality psychology, “*constructs* are invisible internal attributes, measurable by personality tests, whose existence can be used to help explain and predict behavior” (Carducci, 2009, p. 46). The author of a textbook on statistics writes that “non-verbal measurements are made to quantify the hidden behaviour of the subjects such as motivation, frustration and anxiety” (Verma, 2019, p. 24). Conceiving the relationships between questionnaire answers as the numerical equivalents of relationships that are inherent in “hidden” psychological phenomena presupposes a fundamental divide between an enigmatic “inner” mind and a given “outer” world that provides hints about “the inner.”

Cartesian approaches to psychological measurement can also be found in more comprehensive analyses of quantitative research methods. For example, Jana Uher (2018) states that psychological phenomena “can be perceived *only from within the individual*

itself and by nobody else in principle under all possible conditions” and that they “can be explored by others *only indirectly* through individuals’ externalizations (e.g., behaviors, language)” (p. 8). In another article, Uher (2019) elaborates on the specific methods that, according to her, are necessary to investigate psychological phenomena:

Introquestive methods are needed to help individuals become aware of and conceive the psychical phenomena under study, such as through inner self-observation. The introquesting individual must then externalise the outcomes of its introquestion to make them accessible to others, such as through self-report. These externalisations can only be made by the individual under study. (p. 235)

Uher’s remarks are conceptually questionable. When a person declares their love to a loved one, the person does not “relocate” or “transfer” their love from an “inner” privacy to the “outer” world. Rather, the person simply performs a gesture of love. Supporting another person unconditionally or trying to be physically close to that person is as much a part of being in love as are feelings of deep affection and happiness. When a person reports a pain in their left shoulder, the person does not “become aware of” or “conceive” the pain by “inner self-observation.” The person simply feels pain in their shoulder and tells this to other people. If it could literally (and not metaphorically) be said that one observes a stabbing pain in one’s shoulder, then one could also meaningfully say that the person waited several hours to finally observe the pain, or that the observation was only an illusion, or that the pain was covered and difficult to see (cf. Bennett & Hacker, 2003, pp. 90–92). However, all these claims have no meaning. Talking about the “externalization” of hidden psychological phenomena that “individuals become aware of . . . through inner self-observation” (Uher, 2019, p. 235) fuels a misleading Cartesian understanding of psychological measurement according to which empirical investigations into the quantity of enigmatic “inner” entities are needed. There is neither an obscure act of “looking inside” and “finding” or “observing” pain nor one of “externalizing” this pain to the “outer.”

The Cartesian terminology that Uher (2018) uses is especially baffling in light of her claim that her interdisciplinary metatheoretical framework “takes a metaphysically neutral stance without making assumptions of either ontological dualism or monism while emphasizing the necessity for *methodical dualism* to account for observations of two categorically different realities that require different frames of reference, approaches and methods” (p. 4). Since the methodological remarks by Uher that I cited above are built on the Cartesian split between “inner” psychological phenomena and “outer” “externalizations,” her claim of being not committed to ontological dualism seems to be in contradiction with her own analyses of measurement methods. Regardless of these inconsistencies, it is important to note that Uher’s theoretical analyses of psychological research methods reinforce the Cartesian idea that “outer” instruments are needed to measure “inner” psychological phenomena.

Apart from Uher’s work, it is insightful to see that psychological measurement and Cartesian thinking are so closely interwoven that even the highly critical analyses of psychological measurement brought forward by Michell and Trendler rest at some points on elements of Cartesian thinking. By showing that even these comprehensive theoretical

treatments of psychological measurement are not free from Cartesian assumptions, I will complete the reasons why psychological measurement is not an empirical issue.

Psychological phenomena are not inner entities we draw inferences about

As already suggested by Michael Westerman (2014), there is evidence that Michell has incorporated elements of Cartesianism into his critique of quantitative psychology. Although Michell (2000) is highly critical of psychometricians, he does not question their assumption that we can only make claims about psychological phenomena by “first observing something else and making inferences” (p. 648). This short remark by Michell implies the Cartesian thesis that psychological concepts denote unobservable entities whose existence and attributes have to be inferred from overt behavior. There are further passages that buttress this reading of Michell’s work. For example, Michell explains different ways of gaining scientific knowledge and summarizes the most important points of his analysis as follows:

Regarding things that we cannot directly observe, we must come to know them via reason. We have nothing else to reason from other than what we already consider true. That is, given observations and general views about the logic of things (i.e., the logic of causality, of quantity, etc.), scientists must reason their way to conclusions about things not directly observed. (Michell, 2013, p. 19)

For Michell (2012), the task of a scientist is “to find hidden structures” (p. 261) or “to uncover nature’s hidden ways of working” (p. 267). In cases in which it is not possible to observe these structures or ways of working directly, scientists must base their conclusions on observable phenomena. Michell (2013) conceives psychological phenomena as hidden or unobservable entities that scientists have to draw inferences about when he asks: “Is it reasonable to infer from the phenomena of testing that the theoretical concept (or attribute) assessed by a psychological test possesses continuous quantitative structure?” (p. 17). Michell (2013) answers this question in the negative. Central to his critique of quantitative psychology is the notion that an inference from observable phenomena to the thesis that psychological attributes are quantitative is unjustified: “Therefore, my conclusion is that there seems little basis within the phenomena of testing from which to infer that the theoretical concepts assessed by tests are continuous quantities” (p. 13). However, both the question Michell asks and the answer he gives are philosophically questionable because they presuppose a Cartesian theory of meaning according to which our psychological vocabulary denotes “inner” psychological phenomena.

This Cartesian framework is problematic because it is a misrepresentation of our ordinary use of psychological predicates to assume that they refer to unobservable entities whose existence and attributes are inferred. We do not infer that a person is in pain; we see that they painfully scream or that their face is distorted with pain. We do not infer that another person has seen an obstacle; we observe that they have avoided it. We do not infer that Mary loves Jane; we are touched by her caring, sacrificial, and deeply committed behavior towards Jane over many years. The grounds for ascribing psychological

predicates are not conclusions based on inferential reasoning but specific behaviors that are logically sufficient conditions for such ascriptions.

It is not an empirical discovery that when people are in pain, they groan, cry out and assuage their injury. Nor is it intelligible possibility that pain might systematically be correlated with smiling and laughing, as opposed to being correlated with crying and groaning – that is with pain behavior. Similarly, it is not an empirical discovery that when a creature sees, it responds to *visible* objects, uses *its eyes* to *follow* them, cannot see when its eyes are closed, or when it is pitch dark. Rather, the primary warrant for the ascription of psychological predicates to another person or to an animal is *conceptually* bound up with the meaning of the prevalent predicate. Pain-behavior is a criterion – that is, *logically* good evidence for being in pain – and perceptual behavior . . . is a criterion for the animal's perceiving. (Bennett & Hacker, 2003, p. 82)

Screaming is not a cue based on which we make inductive inferences about “inner” pain. It is part of the pain and, consequently, a logical criterion for pain ascriptions. Intensely caring for a person and repeatedly seeking physical proximity means loving that person and is not a hint to some obscure “inner” entity we call love.

Since observable behavior is a part of psychological phenomena just like subjective feelings, sensations, and impressions, it is a conceptual error to extrapolate from the fact that some psychological phenomena can be kept secret to the philosophical thesis that the mind is an inaccessible “inner” entity.

We are prone to confuse the fact that we often do not show our feelings, and indeed sometimes make an effort to conceal them, with the misguided idea that the emotions are in some deep sense “private” and “hidden.” But this is confused. We can often see delight and rage in a person's face, joy, anguish or horror in their eyes, contempt or amusement in their smile. We can hear the love and tenderness, the grief and sorrow, the anger and contempt, in a person's voice. We can observe the tears of joy or grief, the cries or terror, joy or amazement, and the blushes of embarrassment or shame. (Bennett & Hacker, 2003, pp. 221–222)

The possibility of disguising one's feelings or thoughts and the fundamental privacy of psychological phenomena are two distinct claims and the second does not follow from the first because, in many situations, psychological phenomena are overt.

Despite Michell's critical stance toward psychological measurement, he has not called into question the Cartesian framework surrounding psychometrics. The fact that even Michell's comprehensive critique of psychological measurement incorporates Cartesian thoughts further illustrates that psychological measurement is not an issue open to empirical investigation but rather a misbelief that rests on problematic philosophical assumptions.

Psychological talk is not measurement executed by people

According to Trendler (2019b), in attempts to test empirically the hypothesis that psychological phenomena are quantitative, “what inevitably enters as an auxiliary hypothesis is the question of whether humans have the capabilities of measuring devices” (p. 146). It is insightful to look in more detail at what Trendler (2019b) means by the hypothesis that humans have the capabilities of measurement devices:

More precisely, what is tacitly assumed is that, first, humans have “internally” the capability to determine magnitudes of psychological attributes, compare them for more or less, or determine ratios between them and, second, that they are able to communicate, partly or completely, the result of the “internal” measurement operations “outwardly” to the experimenter. (p. 146)

It is important to note that Trendler does not share the view that humans actually can be measurement devices. Subsequent to the cited passage, Trendler (2019b) argues that there can be no measurement of psychological phenomena because in attempts to prove that psychological phenomena are quantitative “one will have to make sure before repeating an experiment that the test participants are valid and undisturbed devices for measurement” (p. 146). This however, is impossible, according to Trendler:

In the case of artificial, man-made instruments it is clear how this can be done. But how are we to proceed with human beings? We cannot simply call the craftsman or the mechanic to check and, if necessary, fix them. The only alternative consists in the assumption that humans are by nature perfect, i.e., undamageable measuring devices. In my view this hypothesis is problematic because in the real world where disturbances abound there are no such things as perfect instruments; i.e., they can always break down, in which case they must be repaired or replaced. (p. 146)

In summary, Trendler (2019a) claims that humans do not have the capabilities of serving as reliable and valid measuring devices, or more precisely, he calls the assumption that humans have solid measurement abilities “an unrealistic hypothesis” (p. 116).

Although Trendler is rightfully critical of measurement in psychology, his argument sidesteps fundamental conceptual issues. Prior to answering the question of whether humans can serve as measurement devices, one must consider whether this is a sensible question to ask in the first place. Trendler (2019b) overlooks this fundamental conceptual point when he states “that the hypothesis that humans have the capabilities of measuring devices . . . is logically coherent and though, when considered superficially, it has the appearance of a testable empirical hypothesis” (p. 147). Trendler (2019a) also states: “Note that the logical possibility that humans may have the capabilities of measurement instruments is not disputed; though it is in my view—for reasons stated elsewhere; Trendler, 2009—an unrealistic hypothesis” (p. 116). While Trendler is skeptical whether humans are able to serve as measuring devices and while he is even more skeptical that it can be empirically proven that humans have the capabilities to be measuring devices, he contends that it is logically sound to ask whether humans can be measuring devices. This, however, is a conceptually confused question. To ask if humans can be measurement devices makes as much sense as asking if humans can be thermometers or Geiger counters. Humans are not measurement instruments; they use such devices for measurement or make decisions based on results provided by them. Humans can be doctors, enemies, slaves, or role models but there is no meaningful way in which humans can be measurement devices. This is not an empirical but a conceptual issue.

If it actually was at least logically possible that humans could measure psychological phenomena, and if humans could always fail in their measurement attempts, as Trendler argues, one could say sentences like the following: “I have a sharp pain in my shoulder, however, I could be wrong and actually feel no pain”; “His anxiety overwhelmed him;

however, he was not really afraid”; or “I have been in love with her for 10 years, however, I might be wrong about that and actually have felt deep-seated hatred for her all this time.” However, all of these sentences are nonsensical and this further illustrates that “humans being measurement devices” is a term without meaning.

In addition, it is important to note that attempts to give an answer to the question whether humans can be measurement devices presuppose a Cartesian conception of the mind. Since it is conceptually impossible that a measurement device is identical with what is measured (e.g., heat cannot be measured by heat and a thermometer cannot measure a thermometer), humans cannot be identical with what they supposedly are trying to measure. Consequently, there has to be an “inner” entity that is separate from psychological phenomena and tries to “measure” them. When Trendler critically asks whether humans really have the abilities to be measuring devices, his questions implies that there must be a ghostly “I” (or something else) that resides “inside” people and is responsible for attempts to “measure” emotions, motivation, and opinions, which likewise are “inside” but separate from the “I.” After the attempts of “measurement” have taken place, this “I” tries to communicate the measurement results to the “outer” world. As I read Trendler, his basic point is that there is no possibility of ruling out that many errors take place in every step of this complex process. However, the outlined process is not an empirical theory but a rather confused Cartesian philosophy. Claiming that humans cannot be measurement devices because they are too error prone in their attempts to register psychological phenomena presupposes an enigmatic “inner” psychological reality, in which errors of measurement can take place. It presupposes the philosophical thesis that our psychological common-or-garden concepts denote unobservable “inner” entities. In consequence, the question whether humans can be measurement devices can only be a “logical possibility” (Trendler, 2019a, p. 116) in a questionable Cartesian conception of the mind.

Moreover, if this analysis is correct, then Trendler’s (2019b) characterization of the assumption that “humans have ‘internally’ the capability to determine magnitudes of psychological attributes” (p. 146) is imprecise. Actually, what must be assumed is that there is some “inner” entity (The mind? The “I”? The “Self”? The soul?) that might possess the capability of “measuring” psychological phenomena and communicating the results to the “outside.” This, however, is an instance of what Bennett and Hacker (2003) have termed the “mereological fallacy”: “ascribing to a part of a creature attributes which logically can be ascribed only to the creature as a whole” (p. 29). Just as it can only be said meaningfully that a car can drive fast and not that the gear shifter can drive fast, it does only make sense to say that humans (or animals) have a capability and not that some inner entity possesses a capacity. This further illustrates that it is not possible to talk about humans being measurement devices without running into conceptual confusions.

In summary, Trendler does not question psychometricians’ fundamental assumption that psychological measurement is an empirical issue. This uncritical adoption prevents Trendler from getting past the Cartesian framework that surrounds psychological measurement. The fact that even Trendler’s highly critical analysis of psychological measurement presupposes elements of Cartesian thinking further illustrates that psychological measurement is not an issue open to empirical investigation.

Conclusion

Despite the differences between Michell's and Trendler's stances toward psychological measurement, their challenges for quantitative psychology can be summarized in a simplified manner as follows. If psychologists want to model psychological measurement after physical measurement, they have to accomplish three tasks: First, psychologists have to realize that empirical evidence is needed to prove that psychological phenomena actually are quantitative (i.e., that they satisfy Hölder's axioms). Second, they have to come up with thoughtful research designs and adequate measurement models that might be able to deliver this empirical evidence. Third, psychologists have to refute all the skeptical arguments why the second task cannot be accomplished with most of the empirical methods that are currently widespread in psychology (Michell) or why this task cannot be accomplished at all (Trendler).

However, the challenge for advocates of psychological measurement is even bigger. Before psychologists can even try to meet Michell's and Trendler's challenges, they first have to accomplish three more fundamental tasks: First, they have to give an interpretation of the question "Are psychological attributes quantitative?" that is actually meaningful. Second, they have to show how all the conceptual confusions that pervade the measurement debate can be avoided. Third, they have to defend the Cartesian assumptions that lie beneath the idea of psychological measurement. The arguments I outlined are sufficient to remain highly skeptical as to whether even one task can be accomplished, let alone all three of them. Consequently, it is reasonable to conclude that psychological measurement is no more than a meaningless pseudotechnical term.

Acknowledgements

I would like to thank Michela Summa, Karl Mertens, Diego D'Angelo, and Manuel Klein for helpful comments on an earlier draft of this paper. The paper also benefitted from comments provided by three anonymous reviewers.

Funding

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

ORCID iD

David J. Franz  <https://orcid.org/0000-0002-3639-2056>

Notes

1. It is frequent practice in the debate about quantification to speak about the usage of quantitative methods in "mainstream psychology" or by "mainstream psychologists." However, not only is the term "mainstream psychology" quite vague, it is also often used as a rather pejorative term that creates an unhelpful juxtaposition of "us" (the critics?) and "them" (the targets of the critique?; Fowers, 2015; Westerman, 2006a). For this reason, the more neutral terms "quantitative psychology" and "quantitative psychologists" are used in this article. These terms denote psychological research that relies on the statistical analysis of numerical data to answer research questions.

2. The reader should note that many of the arguments that were put forward by Bennett and Hacker (2003) are based in large parts on the philosophical works of Gilbert Ryle (1949/2000) and Ludwig Wittgenstein (1974).
3. An exception is the argumentation by Markus and Borsboom (2012). In their objection to Trendler, the authors several times raise doubts as to whether Trendler's arguments rest on empirical assumptions.
4. The value of Maraun's (1998) paper for my arguments was pointed out to me by an anonymous reviewer.
5. I thank an anonymous reviewer for reminding me of the points of this paragraph.

References

- Barrett, P. (2003). Beyond psychometrics: Measurement, non-quantitative structure, and applied numerics. *Journal of Managerial Psychology*, 18(5), 421–439. <https://doi.org/10.1108/02683940310484026>
- Barrett, P. (2008). The consequence of sustaining a pathology: Scientific stagnation—a commentary on the target article “Is psychometrics a pathological science?” by Joel Michell. *Measurement: Interdisciplinary Research & Perspective*, 6(1–2), 78–83. <https://doi.org/10.1080/15366360802035521>
- Bennett, M. R., & Hacker, P. (2003). *Philosophical foundations of neuroscience*. Blackwell.
- 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. <https://doi.org/10.1037/teo0000026>
- Borsboom, D., & Mellenbergh, G. J. (2004). Why psychometrics is not pathological. *Theory & Psychology*, 14(1), 105–120. <https://doi.org/10.1177/0959354304040200>
- Carducci, B. J. (2009). *The psychology of personality: Viewpoints, research, and applications* (2nd ed.). Wiley-Blackwell.
- Fowers, B. J. (2015). The promise of a flourishing theoretical psychology. *Journal of Theoretical and Philosophical Psychology*, 35(3), 145–159. <https://doi.org/10.1037/a0038646>
- Fuchs, T. (2011). Hirnwelt oder Lebenswelt? Zur Kritik des Neurokonstruktivismus [Brainworld or lifeworld? On the critique of neuroconstructivism]. *Deutsche Zeitschrift Für Philosophie*, 59(3), 347–358. <https://doi.org/10.1524/dzph.2011.0028>
- Guyon, H., Kop, J.-L., Juhel, J., & Falissard, B. (2018). Measurement, ontology, and epistemology: Psychology needs pragmatism-realism. *Theory & Psychology*, 28(2), 149–171. <https://doi.org/10.1177/0959354318761606>
- Humphry, S. M. (2017). Psychological measurement: Theory, paradoxes, and prototypes. *Theory & Psychology*, 27(3), 407–418. <https://doi.org/10.1177/0959354317699099>
- Kyngdon, A. (2008). Treating the pathology of psychometrics: An example from the comprehension of continuous prose text. *Measurement: Interdisciplinary Research & Perspective*, 6(1–2), 108–113. <https://doi.org/10.1080/15366360802035570>
- Kyngdon, A. (2013). Descriptive theories of behaviour may allow for the scientific measurement of psychological attributes. *Theory & Psychology*, 23(2), 227–250. <https://doi.org/10.1177/0959354312468221>
- Lamiell, J. T. (2010). Why was there no place for personalistic thinking in 20th century psychology? *New Ideas in Psychology*, 28(2), 135–142. <https://doi.org/10.1016/j.newideapsych.2009.02.002>
- Lamiell, J. T. (2013). Statisticism in personality psychologists' use of trait constructs: What is it? How was it contracted? Is there a cure? *New Ideas in Psychology*, 31(1), 65–71. <https://doi.org/10.1016/j.newideapsych.2011.02.009>

- Maraun, M. D. (1998). Measurement as a normative practice. *Theory & Psychology*, 8(4), 435–461. <https://doi.org/10.1177/0959354398084001>
- Markus, K. A., & Borsboom, D. (2012). The cat came back: Evaluating arguments against psychological measurement. *Theory & Psychology*, 22(4), 452–466. <https://doi.org/10.1177/0959354310381155>
- Martin, J. (2003). Positivism, quantification and the phenomena of psychology. *Theory & Psychology*, 13(1), 33–38. <https://doi.org/10.1177/0959354303013001760>
- Maul, A., Torres Iribarra, D., & Wilson, M. (2016). On the philosophical foundations of psychological measurement. *Measurement*, 79, 311–320. <https://doi.org/10.1016/j.measurement.2015.11.001>
- Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology*, 88(3), 355–383. <https://doi.org/10.1111/j.2044-8295.1997.tb02641.x>
- Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory & Psychology*, 10(5), 639–667. <https://doi.org/10.1177/0959354300105004>
- Michell, J. (2003). The quantitative imperative: Positivism, naïve realism and the place of qualitative methods in psychology. *Theory & Psychology*, 13(1), 5–31. <https://doi.org/10.1177/0959354303013001758>
- Michell, J. (2010). The quantity/quality interchange: A blind spot on the highway of science. In A. Toomela & J. Valsiner (Eds.), *Methodological thinking in psychology: 60 years gone astray* (pp. 45–68). Information Age.
- Michell, J. (2011). Qualitative research meets the ghost of Pythagoras. *Theory & Psychology*, 21(2), 241–259. <https://doi.org/10.1177/0959354310391351>
- Michell, J. (2012). “The constantly recurring argument”: Inferring quantity from order. *Theory & Psychology*, 22(3), 255–271. <https://doi.org/10.1177/0959354311434656>
- Michell, J. (2013). Constructs, inferences, and mental measurement. *New Ideas in Psychology*, 31(1), 13–21. <https://doi.org/10.1016/j.newideapsych.2011.02.004>
- Michell, J. (2020). Thorndike’s credo: Metaphysics in psychometrics. *Theory & Psychology*, 30(3), 309–328. <https://doi.org/10.1177/0959354320916251>
- Nagel, T. (1979). What is it like to be a bat? In T. Nagel (Ed.), *Mortal questions* (pp. 165–180). Cambridge University Press.
- Nagel, T. (1989). *The view from nowhere*. Oxford University Press.
- Ryle, G. (2000). *The concept of mind*. Penguin Books. (Original work published 1949)
- Saint-Mont, U. (2012). What measurement is all about. *Theory & Psychology*, 22(4), 467–485. <https://doi.org/10.1177/0959354311429997>
- Sherry, D. (2011). Thermoscopes, thermometers, and the foundations of measurement. *Studies in History and Philosophy of Science Part A*, 42(4), 509–524. <https://doi.org/10.1016/j.shpsa.2011.07.001>
- Slife, B. D., & Williams, R. N. (1997). Toward a theoretical psychology: Should a subdiscipline be formally recognized? *American Psychologist*, 52(2), 117–129. <https://doi.org/10.1037/0003-066X.52.2.117>
- Tafreshi, D., Slaney, K. L., & Neufeld, S. D. (2016). Quantification in psychology: Critical analysis of an unreflective practice. *Journal of Theoretical and Philosophical Psychology*, 36(4), 233–249. <https://dx.doi.org/10.1037/teo0000048>
- 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. (2013). Measurement in psychology: A case of *ignoramus et ignorabimus*? A rejoinder. *Theory & Psychology*, 23(5), 591–615. <https://doi.org/10.1177/0959354313490451>
- Trendler, G. (2019a). Conjoint measurement undone. *Theory & Psychology*, 29(1), 100–128. <https://doi.org/10.1177/0959354318788729>

- Trendler, G. (2019b). Measurability, systematic error, and the replication crisis: A reply to Michell (2019) and Krantz and Wallsten (2019). *Theory & Psychology*, 29(1), 144–151. <https://doi.org/10.1177/0959354318824414>
- Uher, J. (2018). Quantitative data from rating scales: An epistemological and methodological enquiry. *Frontiers in Psychology*, 9, Article 2599. <https://doi.org/10.3389/fpsyg.2018.02599>
- Uher, J. (2019). Data generation methods across the empirical sciences: Differences in the study phenomena's accessibility and the processes of data encoding. *Quality & Quantity: International Journal of Methodology*, 53(1), 221–246. <https://doi.org/10.1007/s11135-018-0744-3>
- van Lill, J. B. (2012). Psychometrics: Signs of pathology, anxiety or a misdiagnosis? *Journal of Social Sciences*, 30(3), 205–210. <https://doi.org/10.1080/09718923.2012.11892997>
- Verma, J. P. (2019). *Statistics and research methods in psychology with excel*. Springer. <https://doi.org/10.1007/978-981-13-3429-0>
- Westerman, M. A. (2006a). Quantitative research as an interpretive enterprise: The mostly unacknowledged role of interpretation in research efforts and suggestions for explicitly interpretive quantitative investigations. *New Ideas in Psychology*, 24(3), 189–211. <https://doi.org/10.1016/j.newideapsych.2006.09.004>
- Westerman, M. A. (2006b). What counts as “good” quantitative research and what can we say about when to use quantitative and/or qualitative methods? *New Ideas in Psychology*, 24(3), 263–274. <https://doi.org/10.1016/j.newideapsych.2006.10.004>
- Westerman, M. A. (2011). Conversation analysis and interpretive quantitative research on psychotherapy process and problematic interpersonal behavior. *Theory & Psychology*, 21(2), 155–178. <https://doi.org/10.1177/0959354310394719>
- Westerman, M. A. (2014). Examining arguments against quantitative research: “Case studies” illustrating the challenge of finding a sound philosophical basis for a human sciences approach to psychology. *New Ideas in Psychology*, 32, 42–58. <https://doi.org/10.1016/j.newideapsych.2013.08.002>
- Westerman, M. A., & Steen, E. M. (2007). Going beyond the internal—external dichotomy in clinical psychology. *Theory & Psychology*, 17(2), 323–351. <https://doi.org/10.1177/0959354307075048>
- Wittgenstein, L. (1974). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Basil Blackwell.
- Yanchar, S. C. (2006). On the possibility of contextual—quantitative inquiry. *New Ideas in Psychology*, 24(3), 212–228. <https://doi.org/10.1016/j.newideapsych.2006.09.005>

Author biography

David J. Franz is both a trained psychologist and philosopher working as a research associate at the University of Würzburg. He specializes in various topics at the intersection of psychology and philosophy. The primary objective of his scientific work is to bring psychology and philosophy closer together in order to improve research in both disciplines.