The Journal of Mind and Behavior
Winter 2019, Volume 40, Number 1
Pages 29–52
ISSN 0271–0137

What is Scientific Definition?

Fiona J. Hibberd

University of Sydney

Expertise in *what* is being measured, predicted, explained, replicated, etc. requires the development of scientific (conceptual) definitions. Yet an account of what that involves is absent from the psychology literature. This article explains what the standard (classical) account of scientific definition is, what it is not, the steps involved in developing a scientific definition, the distinction between definition and classification, and other key points. The article then turns to recent research in the philosophy of science which proposes an alternative to the standard account. This alternative is said to provide scientific definitions that better accommodate the inherent variation of certain kinds of phenomena. Its implications are examined and they point to it *not* being an alternative to the standard account at all, but an approach that captures the early-to-mid stages of research towards genuine scientific definition. Attention then turns to psychology's conception of the operational definition, how it relates to scientific definition, and what else psychologists sometimes mean by "operationism." Operational definitions are not scientific definitions and it is time that psychology give consideration to the latter.

Keywords: definition, essentialism, non-essentialism, operationism

In a recent article, Gigerenzer (2017) argued that psychology's most pressing challenge was the development of its theoretical base through the testing, construction, and unification of theoretical frameworks. He also made the point that this required the development of rigorous definitions. Without these, no discipline advances from the vagueness and ambiguities of a less than technical language and the follies that result. We are probably all familiar with research articles which conclude with the observation that the concept in question has either suffered under the weight of a plethora of definitions, has not been defined consistently, has not been defined without ambiguity, has not been defined at all, or that there is a mismatch between the conceptual definition and its operationalization. For example:

Correspondence concerning this article should be addressed to Fiona J. Hibberd, Ph.D., School of Psychology, University of Sydney, Sydney, NSW 2006, Australia. Email: fiona.hibberd@sydney.edu.au

> A definition of emotion common to the affective sciences is an urgent desideratum. Lack of such a definition is a constant source of numerous misunderstandings and a series of mostly fruitless debates. (Mulligan and Scherer, 2012, p. 345)

> … in many applications of psychological measurement, the definition of the attribute of interest is vague at best and incoherent or entirely absent at worst … affairs might be improved via greater attention to definitional clarity and the a priori articulation of testable theories. (Maul, 2017, pp. 61, 65)

> … psychology cannot become a science at all until, like physics and other true sciences, we have our own set of optimal standard measures. There is, however, a necessary prior step, which is to agree on single *definitions* of our major constructs, given that the content of the measure depends critically on the content of the construct definition. (Rossiter, 2017, p. 490)

In fact, *scientific* definition (I do not mean "operational definition") is not a subject that psychology's method texts and journals address, despite it being foundational to the research process. This absence has its roots in psychology's history where, during its formative years, psychologists were urged to focus on the *operational* definition of concepts (e.g., Stevens, 1939).[1] Operationism has, for roughly 80 years, been hegemonic and few method topics have received as much attention. The 1930s, 1940s, 1950s, 1980s, 1990s and 2000s, in particular, saw formulations, elaborations, re-formulations, critiques, and symposia on operationism in psychology (Feest, 2005) and, to this day, the importance of operationally defining variables continues to be cemented into the thinking of undergraduate students. However, operational definition is not scientific definition and psychology's ongoing attachment to the former has not helped to address the challenges that Gigerenzer and others have noted.

Without an account of what scientific definition involves and how far removed it is from operationism, the prospects of any change in thinking and practice are, understandably, negligible. My aim, then, is to (i) bring to psychology's attention the standard (or classical) account of scientific definition; (ii) comment on some implications and point to the steps involved in achieving a scientific definition; (iii) evaluate recent arguments (from other disciplines) for an alternative to this standard account; and (iv) make clear why operationism is not scientific definition, explain the inherent difficulties of any operational definition and address what else psychologists sometimes mean when they use the expression "operational definition."

The actual doing of science is invariably messy, and there are vast differences between psychology's many domains with respect to the scientific challenges faced and the methods employed to meet those challenges. However, the formal

[1] Perhaps its absence has been exacerbated by the recent zeitgeist, postmodernism, where the supposed impossibility of definition has, arguably, affected societal and intellectual expectations?

account of scientific definition provided here concerns *all* of psychology. Psychologists from across the discipline can draw from this article (and the examples provided) to make connections with their specific research area. My examples of practices in psychology which are at odds with scientific definition are intended as instructional. I begin with with some preliminary, but necessary, comments on what scientific definitions define and their role in science generally.

What Do Definitions Define?

I assume that research psychologists operate against the backdrop of realism, i.e., that they have beliefs and intentions consistent with a realist philosophical framework. This just means that they think behaviour and cognition can be investigated scientifically, practical difficulties and human fallibility notwithstanding, because the complex bio–psycho-social systems of interest to psychologists exist or occur independently of their observing, thinking, talking or writing about them. This independence entails treating questions about the nature or features of psychological phenomena as ontological questions, quite distinct from epistemic issues about the extent to which, and manner in which, we can come to know them. Realism also involves the assumption that complex systems, and reality generally, are made up of situations or states of affairs, each situation being a particular-instance-of-a-certain-general-sort or *kind*; any situation occurrence just is the occurrence or existence of a certain sort or kind of thing (e.g., Armstrong, 1997). The liquid in my glass is a particular sample of a general sort, the kind *water*. And yesterday, for a short period of time I was the bearer of something general — a type or kind of *emotion*, viz., frustration.

Understood realistically, then, the call for scientific definitions from Gigerenzer and others, is not a call to define words. It is not a lexical enterprise. At its most elemental, it is a call to answer the ontological question "what is it?" or "what kind of thing is it?" Although we use words to propose or state what kind of thing we think it is, a definition references the *what-it-is-to-be* that kind of thing — its principal features or structure — in order to delimit it from other kinds and to make possible a systematic study of it and its connections. The kind or type or sort of thing defined may be a chemical substance, a fundamental particle, a virus, a relation — such as intentionality or attachment, a process — such as extinction learning, a behaviour — such as vigilance test performance, cognition — such as remembering, an emotion — such as fear, an effect or phenomenon — such as confirmation bias, and so on.

Some may object that, for a non-operational definition, what is being defined is a concept, hence the distinction between conceptual and operational definitions (e.g., American Psychological Association, 2012, p. xxix). This objection may be made for various reasons. In particular, some may have in mind a distinction between "theoretical" and "observational" terms. For example, "trait"

could be regarded as an unobservable theoretical term distinct from the observable behaviour that traits supposedly cause, and so "trait" should be treated as a concept, i.e., to define "trait" is to define a concept, not a kind. I do not think there is much at stake here. The plausibility of any strict theory–observation distinction has long been dismissed (e.g., Putnam, 1962). Where the line is drawn between the two depends on a range of contextual parameters, such as notions of "direct" and "indirect" observation, theoretical presuppositions, and background knowledge. Moreover, concepts need not reference only that which is hypothetical or theoretical — think "H_2O" or "behaviour," for example. Still, if some wish to insist that it is concepts which are defined, this is a minor point which should not distract from the substantive issues involving definition. My assumption will be that psychologists use the term "concept" in its logical sense — a sense attuned to realism — where, in speaking of the concept of *X*, what is being "pointed to" is kind *X*, viz., a kind of relation, process, behaviour, state, etc., *thought to be real regardless of whether it is observable*.

The Standard Account of Scientific Definition

With its roots in Socratic thought, the standard account of scientific definition was developed by Aristotle and has been upheld most recently by various realist philosophers (e.g., Devitt, 2008; Groarke, 2009; Kripke, 1980; Putnam, 1975). To provide this type of definition is to describe the kind's *essential* or *defining* features — to describe what it is in virtue of that makes kind *X*, *X*. Essential or defining features are the characteristics or features or conditions without which that kind could not be the kind it is, i.e., the kind has them necessarily. It is, then, the mark of any real genuine kind, such as a particular kind of behaviour or cognitive process, that something identical — some set of common features — necessarily runs through *all* of its instantiations.

To elaborate on how the classical view is understood, let's take water as a kind or type. The genus of water is *substance*. The species is what water *is* — its unique set of features present in every sample of water which distinguishes it from other substances — its differentia.[2]

With respect to its differentia, see Table 1. Water has the essential characteristics of an oxygen atom binding with two hydrogen atoms. This is its molecular structure and water has these properties necessarily. We may think that calcium is also a necessary property because it also occurs in water naturally, but calcium can be removed from water (using water purification technology) and water is still water. Calcium, then, is not necessary. With respect to water, it is an accompanying or contingent property (although, it is a necessary property of chalk). So, even

[2] The terms "genus" and "species" are relative (like "parent" and "offspring"), as are the terms "kinds" and "sub-kinds."

if these accompanying features *always* accompany the necessary characteristics, they are not themselves necessary if the kind could still be the kind it is without them. The point is that we should not assume that what happens all the time is necessary — frequency is neither necessary nor sufficient for necessity.

Table 1
Essential and Accompanying Properties of Water

	Location 1		Location 2		Location 3		Location 4	Location 5	Location 6
WATER	H_2	Ca^{2+}	H_2	Ca^{2+}	H_2	Ca^{2+}	H_2	H_2	H_2
	O		O		O		O	O	O

H_2	Two hydrogen atoms
O	Oxygen atom
Ca^{2+}	Calcium

Water's essential features, oxygen and hydrogen, are themselves kinds which will, for scientific purposes, require definition, but it is in virtue of their conjunction (of an oxygen atom binding with two hydrogen atoms) that we categorize water as water and not some other substance. It is also in virtue of water having this *unique* set of essential/defining features that (i) it has the causal effects it has; (ii) it can be referred to as a *mono*thetic kind; and (iii) the concept "water" is a monothetic concept (see Sokal and Sneath, 1963, pp. 13–14). As a genuine kind, water is the ontological springboard for induction, especially for causal explanation, prediction, and theoretical integration. The fact that a large quantity of water will have certain effects on a particular area of arid land, for example, can only be explained by reference to the real kind — water — and the other relevant kinds (dry land, fertile land) involved in that causal process. That process cannot be explained scientifically if what water *is*, is not understood.

On this standard view of definition, then, identifying real, genuine kinds means identifying the kind's genus and its differentia, the latter being those conditions or features without which it would not be the kind of substance it is. These essential/defining features constitute the kind's ontological structure, they are not "principles," "essences," or "meanings." In describing water's essential features, there may be any number of linguistic alternatives, but that requires them to be synonymous with the original kind term. We may use the term "water" or the term "aqua" to refer to the essential features of water but the two are synonymous

because the word "aqua" has the same intension as the word "water" — both refer to water's essential features and so mean the same.[3]

We can now see why recommendations to write brief formal descriptions of a target concept (e.g., Clark and Watson, 1995, p. 310) is not definition. It may help the early stages of our thinking about that concept, but if we have not identified the genus nor any features of the target concept as essential (rather than accompanying), we are far from having circumscribed the concept and made possible a systematic study of it and its connections. Hence the dictum: all definition is description, but not all description is definition.

The fact that any psychological condition, process, state, relation, or attribute is not a substance is not an inpediment to the scientific definition of psychology's concepts. Defining features can be (i) intrinsic, e.g., those of the amygdala; (ii) partly intrinsic, partly extrinsic and relational, e.g., human cognition; or (iii) relational and extrinsic, e.g., an individual's connection to some aspect of government. For example, where there are different *kinds* of relations, identifying their defining characteristics is, in principle, no different from identifying the defining characteristics of water. Take, for instance, dissociative identity disorder as a real kind. The DSM–5 (p. 291) has defined it scientifically, identifying its genus as *dissociative disorders* and its differentia as:

(a) the presence of two or more distinct personality states or an experience of possession which is not part of a broadly accepted cultural or religious practice, and
(b) recurrent episodes of amnesia.

Let's agree that both conditions are partly intrinsic, partly extrinsic and relational.[4] Each condition or feature is taken as necessary, together they are sufficient.[5] If we assume that this definition is right, what it is to have a dissociated identity is the conjunction of these essential conditions. There may be any number of accompanying features, but these two essential conditions run through all of its instantiations (the occasions where both conditions are simultaneously borne by certain individuals) necessarily, i.e., those with dissociated identities are homogenous with respect to these two conditions. And, as with the water example, each essential condition is complex with parts that are also complex and require definition — amnesia, for example. There are kinds nested within kinds.

These essential conditions of a dissociated identity are not to be confused with any *diagnostic* criteria for the disorder, just as *signs* of water should not be

[3] A term's intension is the defining features or characteristics to which the term refers (Grayling, 1997).

[4] I have in mind biochemical states/processes/systems linked to a perceptual system within a body as necessary for the individual to stand in primarily cognitive relations to his or her environment.

[5] The possibility that condition (b) is an accompanying feature (and, therefore, of the same status as calcium in water) would require justification from researchers in the DID area.

confused with water. The DSM reports on how the essential features of a dissociated identity will manifest behaviourally. Generally, one can expect to observe discontinuities in speech, affect and behaviour, conversion symptoms, and so on. The scientific concern (as opposed to clinical concerns, e.g., arriving at an accurate diagnosis, treatment) is that these generalizations about behaviour require an explanation — an answer as to *why* those with dissociative identities behave in this way. No credible explanation can be given without reference to the essential/defining conditions; the explanation must appeal to the very nature of a dissociated identity. It is the essential conditions of a dissociated identity as a real kind that are explanatory and do the causal work, just as in the water example.

We can see this standard account of definition at work in Gigerenzer's (2017) article. He discusses the features of, and relations between, distinct memory phenomena so as to illustrate theory integration. (Unfortunately, in doing so, he misdescribes the type of definition as "operational," a matter I return to later.) In effect, Gigerenzer identifies the genus as *memory* and then goes on to single out two kinds of memory phenomena — the *reiteration effect* is identified by two essential conditions and is distinguished from *hindsight bias* which has two different essential conditions. Either definition could be refuted because each feature/condition proposed has the status of an empirical hypothesis (in being either true or false) but given that both definitions are widely accepted, Gigerenzer has good reasons for assuming their truth.

Obviously, the possibility of true definition does not mean that it is easily achieved. The seemingly successful identification of a kind's essential or necessary features may signal the end of a demanding and lengthy program of research, albeit one that paves the way for new lines of investigation. But that definition always remains open to revision. Helmholtz's (1887/1977) contribution to the definition of a quantitative attribute, for example, was important but incomplete. Burali–Forti (1898) and Hölder (1901/1996, 1901/1997) soon followed with what is now regarded as a complete definition but one that still may be revised — if, for instance, infinitesimals were readmitted into measurement (cf. Ehrlich, 2006). On the other hand, a concept may repeatedly elude scientific definition suggesting that although the term functions as a linguistic marker in everyday language, it fails to reference any genuine scientific kind (e.g., Wrigley, 2015 on the concept of vulnerability).

What Scientific Definition is Not

From the standard account of scientific definition above, we can now see what scientific definition is not and, therefore, should not be confused with.

(i) It is not conventional — criteria for the term's use have not been decided by agreement and so established by convention *without referencing the kind's essential features*.

(ii) It is not stipulative — no-one is arbitrarily assigning a meaning to the expression "dissociative identity" or the "reiteration effect," for example.
(iii) It is not nominal — we are not assuming that the kinds exist in name only and that what we are defining is nothing more than the name/word or the idea/imagined features attached to the name/word (e.g., the word "unicorn") or defining a word by using other words. Our concern is scientific, not linguistic, because our working hypothesis is that there are psychological relations, processes, or states with the features or conditions indicated.
(iv) It does not confuse *is* with *does* — we are not conflating what something *is* with a description or definition of its goal, purpose, or function. Conventional definitions are inclined to define something by its function or causal role, e.g., a heart is a pump, an eye sees, a shoe protects the foot, but this is not to identify the essential features of the heart, eye, or shoe. "What is *X*?" is a question different from "what does *X* do?" or "what function/role does *X* have?" or "what are the effects of *X*?" There is nothing wrong with a functional definition as long as what the kind *does* is not muddled with what the kind *is*.
(v) It is not operational — we are not defining or giving meaning to a kind "… by spelling out what the investigator must do to measure it and evaluate that measurement." Nor is it "… a sort of manual of instructions to the investigator" (Kerlinger and Lee, 2000, p. 42).
(vi) It is not ostensive — we are not pointing to an example of hindsight bias and saying "hindsight bias is *this*" or "the term 'hindsight bias' means *this*." To refer to an example of *X* is not to describe the kind *X*'s essential features.
(vii) It is not classification — we are not collecting examples of an empirical phenomenon and making the case that each example is a member of some particular class. We classify according to our knowledge, needs, and interests. In the scientific context, an outcome will be a collection of things (a class) classified according to the kind of thing they are, that is according to the kind's essential features. So, classification *depends on* definition and, therefore, cannot be definition. For instance, we could point to a particular act of tweeting a threatening message through social media and make the case that this act exemplifies aggression and, therefore, belongs to the class of aggressive behaviours.[6] But "making the case" is not definition for it depends on the definition of aggression, on what kind of behaviour aggression is. It depends on the conditions required for membership to that class, those conditions being the essential features that run through every act of aggression necessarily. This definition–classification distinction may not always be well understood. For instance, the claim that DSM–5 allows more individuals to be diagnosed with a mental disorder than previous editions does not amount to an "expansion of the concept of mental disorder" (the title of Boysen

[6] This class would be the term's (aggression) extension — the range of acts to which the term applies (Grayling, 1997).

and Ebersole, 2014), because the definition of mental disorder hasn't altered. It is rather that the number of cases of mental disorder (the cases in this class) is likely to increase because the diagnostic criteria have been made more inclusive.

Further Considerations

The diminution of categorical thinking in place of dimensional thinking in some areas of psychology is of little consequence to this standard account of scientific definition. Some do regard most concepts in, for example, personality and psychopathology as dimensional (e.g., Haslam, Holland, and Kuppens, 2012) and assume that these dimensions are continuous (e.g., American Psychiatric Association, 2013, p. 770; Markon, Chmielewski, and Miller, 2011). Nevertheless, the two poles of a dimension are usually empirical opposites, each requiring definition (e.g., extroversion–introversion). And it should not be assumed that matters of degree or differences between degrees are quantitative (Michell, 2012). Where they are not, the distinctions between degrees or levels are categorical or heterogeneous. The researcher is then dealing with different kinds (or "sub-kinds," which are kinds nevertheless) each requiring its differences to be identified for a finely-tuned discrimination between kinds.

But it can reasonably be asked whether specifying a concept's essential/defining features is really necessary? Is it not enough to just recognize a particular case as a case of kind *K*? Consider the following scenarios:

1. You take concept *A* to be a genuine kind but you are wrong, it isn't.
2. You take concept *A* to be a genuine kind and you happen to be right, it is. You manage to identify some of its features but are unable to distinguish between its essential/defining and accompanying features.
3. You take concept *A* to be a genuine kind and you are right, it is. And you are also able to specify its essential/defining features.

Scenario 1 is sub-optimal, scenario 3 optimal, but what of scenario 2? Does it matter if, in your field of research, the kind's essential features have not been identified? Perhaps this depends on your research question. Conceivably, you may not need to know water's essential properties to conduct research on the effects of water, you may simply need to recognize your samples of water as water. But if asked to justify that your samples were really water and not some other clear colourless liquid, could you do this without specifying water's essential features? You could examine your samples to determine how they respond under certain conditions and conclude that they are water because only water responds in this way. But if then asked to explain why water responds in this way, would you not struggle to do so without citing its essential features or do you really think the explanation need not appeal to them? In another context, it could be that your research question requires a controlled setting where different properties of the treatment variable

(water) are manipulated instead of simply varying the amount of water. Adequate control may involve some conditions where you do not use samples that consist of accompanying properties. This requires the expertise to distinguish between water's essential and non-essential properties, if the experimental design and any causal account is not to be compromised. Where a degree of expertise is required and where scientific inference is concerned, scenario 2 will not suffice, not even with water, let alone with the more slippery concepts that permeate psychology.

Recent Non-Essentialist Programs of Scientific Definition

With an understanding of the standard, classical account in place, we should now consider recent non-essentialist programs of scientific definition as an apparent alternative. First, some history. Prior to the attention it received from realist philosophers during the second half of the twentieth century, the standard account had been largely set aside during the first half. The reasons were many. There was Russell's and Whitehead's persistent denunciation of most Aristotelian doctrines as they advocated for their modern mathematical logic (e.g., B. Russell, 1946/1984, p. 212). There was Bridgman's (1927) striking response to the conceptual upheaval brought about by Einstein's theory of special relativity and the then new quantum mechanics. It included his now infamous operationist solution, that "In general we mean by any concept nothing more than a set of operations; *the concept is synonymous with the corresponding set of operations*" (p. 5).[7] There was Wittgenstein's early (1930/1975, p. 59) identification of the meaning of a word with its use (instead of features or conditions referred to in using a particular word) and his 1929–1930 selling of this thesis to Vienna's logical positivists. They recognized its conceptual connection to Bridgman's operationism and employed both to defend their verifiability principle (Schlick, 1936/1979). There was, from 1935 onwards, the promotion of operationism (through a series of papers) to psychologists aligned with psychology's methodological behaviourism and its assumed unobservability of the mental (e.g., Stevens, 1935). And there was Wittgenstein's (1953/1967) later claim that there are no essential, defining features in what we call "language" and "games" — the language-games that constitute language share nothing more than family resemblances (§§65–67).

By the late 1950s the move away from the standard view was well underway and two seemingly distinct research programs began to develop (Sutcliffe, 1999). One proliferated in the areas of mathematics, statistics, engineering, and computational research. Here, concerns about the classification of entities or individuals were explored through technique-driven practices. For example, methods (various algorithms) of clustering were devised to best group entities or individuals

[7] This "solution" was quite at odds with Einstein's epistemology and explicitly dismissed by him (Einstein, 1949).

with respect to a criterion *imposed on them* (Sokal, 1974, p. 1120). The second program was conceptually driven and concerned to argue that (i) there are no essential defining conditions for *biological species*, but (ii) it is absurd to suggest that biological species are not natural kinds (e.g., Mayr, 1963).

This move away from the standard view is now referred to as "the death of essentialism." The central supposition (certainly explicit in the second research program if not the first) is that many real kinds in the living sciences have no unique set of essential or defining features — they are not monothetic. Justification for this is that across living agents, research continues to reveal variation and change through time — a variability or heterogeneity with respect to genetics, morphology, abilities, diseases, diagnoses, and so on. Your response may be that these variations could be the expected changes in agents' *accompanying* features as their environments change while the essential features remain unchanged. However, the non-essentialists' position is that the variability within some kinds is such that the kinds have no unique features or sets of necessary and sufficient conditions common to all of their particular instantiations.

It was the philosopher Morton Beckner (1959) who reflected Wittgenstein's move away from the Socratic–Aristotelian tradition and set out the defining conditions of any polythetic (non-monothetic) concept for biology. Notwithstanding the recent return to Aristotelianism (cf. Berti, 2011; Groarke, 2009) and realist arguments for the standard account of definition (cited earlier), there is now a widely sanctioned account concerning the definition of non-essential natural kinds. Originally proposed by the scientific realist, Richard Boyd (e.g., Boyd, 1999, 2010; Keller, Boyd, and Wheeler, 2003) and known as the Homeostatic Property Cluster (HPC) account, its core thesis is that some kinds are defined by a cluster of features that regularly *but not exceptionlessly* co-occur; *and* a set of factors (causal homeostatic mechanisms) that maintain their systematic co-instantiation or clustering, factors that provide some necessary cohesiveness or stability to the cluster. Again, you may think that perhaps the causal homeostatic mechanisms are the essential/defining features or conditions — after all, the binding of hydrogen and oxygen atoms is a complex causal process — but this need not be the case because, according to Boyd (2010), different causal profiles may obtain over time. Suffice it to say, a virtue of the HPC account is thought to be that it better accommodates reality because:

> Lots of phenomena — [biological] species, islands, kinds of economic systems, biological populations, and higher taxa — have just the sort of "vague" boundaries recognized by the HPC conception. (Boyd, 2010, p. 698)

Although Boyd (1991) recognizes that, for induction or explanation, kinds "… must always cut the world at its joints … [so] that we accommodate our categories to the causal structure of the world" (p. 139), at times that "cutting" cannot be clean.

Difficulties with the Non-Essentialist Account

If some kinds have no essential or defining features, identifying their structure in terms of those features is a fool's errand; the standard account of scientific definition fails to apply, and any concept referencing a genuine HPC kind would be polythetic, not monothetic. Let's consider this possibility — that *some* kinds just do not have unique sets of essential features or conditions — and its implications. Imagine some psychological state or condition which we reference as concept *Y*. Let's assume that, consistent with the HPC account, the best available research tells us that *Y* is a cluster of single homeostatic features *A, B, C, and D*, where *Y* has no essential features because no one feature appears in every instantiation (see Table 2).

Table 2

Y as a Polythetic Concept or "Kind"

State or condition *Y*	Occurrence 1	Occurrence 2	Occurrence 3	Occurrence 4	Occurrence 5
	A		*A*	*A*	*A*
	B	*B*	*B*		*B*
	C	*C*		*C*	*C*
			D	*D*	

Y appears, then, as a non-monothetic concept which may, nevertheless, reference a real kind. Given that it cannot be defined essentially, its definition will, at best, "… reflect the imperfect clustering of relevant properties …" (Boyd, 1999, p. 153). This means that there will always be some indeterminacy or vagueness as to what the kind actually is. The vagueness, is ontological, not epistemological — it's "out there" — and so no "clean cuts" could ever be made between *Y* and neighbouring concepts. This "death of essentialism" means that the basis for definition has shifted. Our task as researchers is not to look for a unique set of essential conditions that run through every case of *Y*, but to look for a cluster of features across the many *Y*'s, so that the many cases of *Y* at best bear some "family resemblance" or are similar to one another.

Still, the scientific question remains: What *is Y* or what does concept *Y* reference?[8] If *Y* is a real kind (of process/relation/phenomenon/state) but we cannot

[8] In the analysis that follows, I have drawn from Sutcliffe (1993, 1999) and I am grateful to Joel Michell for many helpful discussions.

say what it is, this will hamper genuine scientific advance. We still want to offer a realist definition, not any of those identified above, so given the HPC account the question is: In what does *Y* consist (or, if you prefer, what does concept Y refer to) if it isn't always the same cluster of four features? To begin to answer that question, we must specify the different sub-types that constitute the kind but this requires us to stipulate what is to count as a sub-type. Say we decide on the following constraint: that at least any two features from the cluster are sufficient to identify the occurrence as an occurrence of that kind, kind *Y* (see Table 3).

Table 3
Y as a Disjunction of Sub-Types

State or condition *Y*	Occurrence 1	Occurrence 2	Occurrence 3	Occurrence 4	Occurrence 5
	A		*A*	*A*	*A*
	B	*B*	*B*		*B*
	C	*C*		*C*	*C*
			D	*D*	
Sub-types of *Y*	Y_1	Y_2	Y_3	Y_4	Y_1

It follows that *Y* is anything up to 11 possibilities. *Y* can, in principle, unfold into 11 different sub-types each with their own unique set of necessary and sufficient conditions. In Table 3, we can see two instantiations where only *A*, *B*, and *C* co-occur. This means that features *A, B,* and *C* are necessary and sufficient for sub-type Y_1, *B* and *C* constitute the sub-type Y_2 and so on. Therefore, our best answers to the question "What is condition *Y*?" is that it is a disjunction of sub-types: it is Y_1 or Y_2 or Y_3 or Y_4 … or Y_{11}; that *Y* unfolds into 11 monothetic essential sub-kinds; that *Y* is anything up to 11 possibilities, each with their own unique set of necessary and sufficient conditions. We could set the constraint that any three features from the cluster will suffice, in which case *Y* is anything

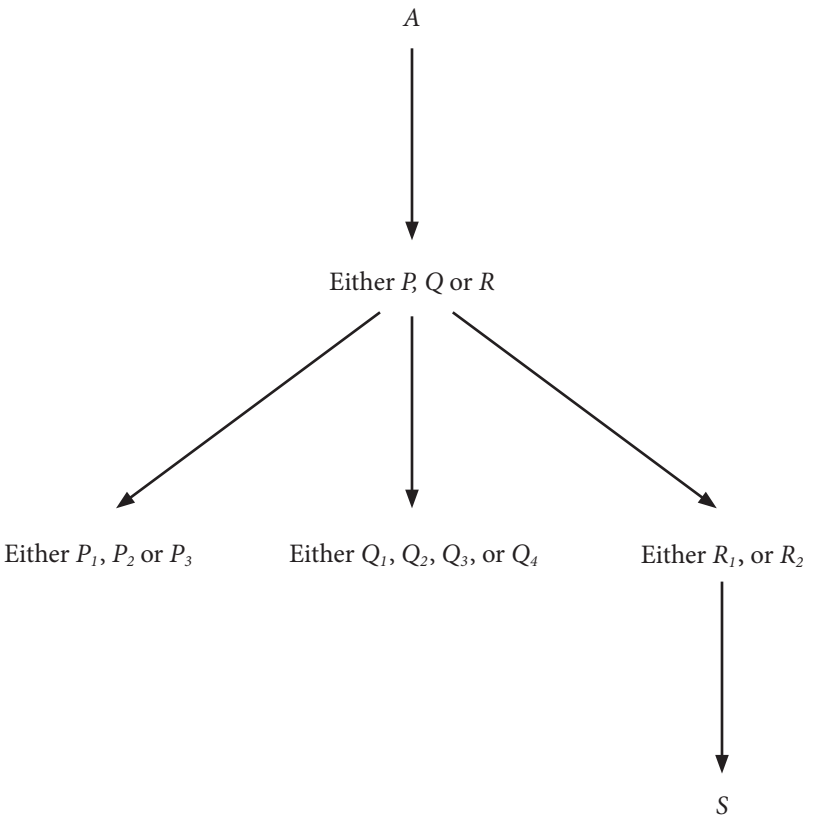

Figure 1: A regress of disjunctions which at some point must stop.

up to five possibilities. Either way, not only has an element of subjectivity been introduced into the method of definition in the form of constraint setting, the relative lack of specification means that we cannot currently make sense of *Y*; we cannot give an account of *Y*. If we are not barking up the wrong tree and *Y* is a genuine kind, what it *is* hasn't been identified and delimited from neighbouring kinds. The science is undeveloped because disjunctions are inadequate in providing a foothold for scientific inference. A disjunctive predicate doesn't reference a real kind.

It is possible that a supposed feature, such as *A*, may itself be disjunctive. What then is the implication? If feature *A* is either *P* or *Q* or *R*, and *P* itself then unfolds into P_1 or P_2 or $P_3 \ldots$ or P_n, as in Figure 1, we have a regress.

It is entirely possible for the regress to venture further down the series of steps but, at some point, it has to stop. Without the regress of disjunctions stopping, we could not say of any random feature (say *E, F, G* …) that it is *not* a feature of the domain we are inquiring into. No feature has been ruled out as being necessary;

no feature gets excluded. This is a *reductio ad absurdum* because if there were no unique feature or set of features common to a type or sub-type, there would be nothing in virtue of which it was one kind of thing and not some other kind of thing. If there were no kinds of things, there would be no different kinds of things. If there were no different kinds of things, no real categorical distinctions could ever be made.

A related implication of non-essentialist definition is that it yields contradictions. Any instantiation of *Y* may consist of *A* (or *B*, or *C*, or *D*) or it may not. If *Y* is either *A* or not *A*, any proposal can validly be deduced (Copi, 1954). Again, there's no foothold for scientific inference because no specific prediction, explanation, or replication can be made from a contradiction.

Moreover, recall that the "death of essentialism" involves a conceptual shift from identity to similarity as the basis for definition; that non-essentialist accounts depend on the concepts of similarity and/or resemblance. In effect, such accounts maintain that although the class of *Y* does not have the same feature or features running through each case, the members are nevertheless all similar, i.e., they resemble one another, they are alike, if not identical. The question then is: If the terms "similar," "resemblance," and "likeness," cannot mean "of the same kind," what do they mean? Let's say: "having some features in common while being different in significant respects"? But this returns us to the standard, classical account of definition and, if those common features remain unspecified, our response is vacuous (Ghiselin, 1966).

Perhaps, then, the notion of similarity can be improved through the application of quantitative concepts? (This is often psychologists' go-to strategy in an attempt to sharpen up that which is subjective or lax.) A coefficient of similarity could be selected, but on what basis if you have no theory of similarity from which to justify your choice? And it still requires the specific features to be identified, the number of features to be used in the analysis, and so on. Without some theoretical basis, you can only make arbitrary choices from a wide range of ad hoc techniques (Sutcliffe, 1999). Even if the degree of similarity is expressed as a correlation or as "distance," the situation fares no better. Both characterize features or relations between features quantitatively when "similarity" is not a quantitative concept — one feature isn't twice as similar as some other feature (Ghiselin, 1966, pp. 213–214) and, again, both require knowledge of the features to begin with. In short, attempts to quantify similarity are a non-starter.

How, then, did the much esteemed Wittgenstein (1953/1967) justify his assertion that the language-games which constitute language share nothing more than family resemblances? The answer is "inadequately." His claim was without adequate warrant because all of the problems identified above were in play. It may also be that, in looking for a feature of language common to specific language-games and failing to find one, Wittgenstein was not looking in the best place. His task may have been helped by "moving up a level" to that of the genus and considering

the differentia between human language and other systems of communication (Sutcliffe, 1993).

Summary

The arguments above point to the impossibility of "non-essential kinds." Any polythetic concept, such as *Y*, is parasitic on monothetic kinds and cannot, therefore, licence a genuine alternative to the standard account of definition. The arguments also point to stages of discovery in the process of definition and identifying clusters of features is not the end point. If *Y* is a genuine kind, we are still unable to say what *Y* is and we are unable to demarcate *Y* from *not-Y*. At best, identifying a cluster of features may serve as a precursor to identifying the kind's essential features. It is worth noting that Beckner's (1959) work on polythetic concepts recognizes this. He treats polythetic concepts as important steps en route to identifying monothetic kinds — a first approximation to the scientifically valued monothetic kinds. At no point does he state, or imply, that polythetic concepts can replace concepts which reference the essential features of monothetic kinds.

How, then, might those defending the standard, realist account of definition accommodate the fact that the world is constantly changing? If "non-essential kinds" are impossible, what of Boyd's (2010) claim that kinds just do evolve and change over time? The classical, essentialist response is that a kind's accompanying features may change but a *genuine kind's essential or defining features or conditions do not — the very notion of a changing kind is incoherent*. Consider the following from Groarke (2009):

> Suppose that every nitrogen molecule in the universe was changed into oxygen. It would not follow that nitrogen or oxygen lacked essences. The species "nitrogen" would become extinct and the species "oxygen" would become more prevalent than it was before. But nitrogen cannot and could not become oxygen. (p. 420)

So, the very idea that a kind is *X* while becoming *Y* (and, therefore, not *X*), or that a kind is sometimes *X* and sometimes not *X*, is contradictory. What something is, is its own unique form or structure — its essential features. Lose that structure and the kind is no longer real, it ceases to be. Some kinds come and go, others are more permanent. Conceptual change reflects that.

Operational Definitions

Now we can turn to the operational definition (operationism). To operationally define one's variables has been a (perhaps *the*) methodological imperative in psychology since the 1930s. Despite criticism from outside the discipline and within (e.g., Benjamin, 1955; Bickhard, 2001; Leahey, 1980; Michell, 1990; L. J. Russell, 1928; Wallach, 1971), operationism continues to be upheld as an

essential feature of rigorous psychological research (e.g., American Psychological Association, 2012). We should first consider the standard treatment of what an operational definition is and then turn to why this treatment is illogical before considering what else psychologists may sometimes mean when they invoke operationism.

An early statement of what operationism is can be found in Stevens (1935): "Operationism consists simply in referring any concept for its definition to the concrete operations by which knowledge of the thing in question is had" (p. 323). If we move to the year 2000, there is little change: "[An operational definition] defines or gives meaning to a variable by spelling out what the investigator must do to measure it and evaluate that measurement" (Kerlinger and Lee, 2000, p. 42).[9] Similarly, from the glossary of Rosenthal and Rosnow's (2008) text: "**operational definition:** An empirically based definition, that is, the meaning of a variable in terms of the operations used to measure it or the experimental method involved in its determination" (p. 753). The standard treatment, then, is that a variable or concept is defined in terms of the operations that facilitate knowledge of it.

It should not be difficult to discern just how far removed this is from scientific definition. Working hypotheses about a relation, process, behaviour, or state's independently existing essential and non-essential (accompanying) features or conditions are suspended or avoided because what a variable is or what a concept means just is the procedures used. It is this which marks operationism as a false doctrine because it is illogical — illogical because it confuses the activity of manipulating and/or measuring a variable with what that variable is (or with what the term referring to the variable stands for). This conflation is akin to confusing the operations or activities involved in baking with the cake itself. It is a confusion between the putative kind or concept and the means (the operations) by which that kind or concept is identified (Byerly and Lazara, 1973; Michell, 1990).

A corollary of this logical error is that the crucial distinction between cause and effect is compromised. For instance, if the intensity of food deprivation *is defined in terms of* the hours spent withholding food (a feature of the method), then logically the hours spent withholding food cannot be a cause of that intensity, not even partially. And yet the time spent withholding food has surely been at least a partial cause of the intensity of the food deprivation. Similarly, if fluid reasoning ability is defined in terms of performance on a fluid reasoning test, then logically that ability cannot also be a cause of such performance.

This conflation of cause and effect can be found in method texts for undergraduate and early graduate students in psychology. For example:

[9] At the time of writing, Google Scholar reports that the various editions of the Kerlinger and Lee text have received over 21,000 citations. It is not, then, a text of little import.

> … we use operational definitions, which consist of manifest variables, to represent hypothetical constructs in research. In doing so, we assume that the hypothetical construct is causing the presence and strength of the manifest variable used as its operational definition …. By measuring the manifest variable, we measure the hypothetical construct …. (Whitley and Kite, 2013, p. 148)

Here the hypothetical construct is operationally defined in terms of behaviour (or anything else that we can observe directly) *and* it is also causing the behaviour. But this is not logical — if the construct is *X*, it cannot cause *X*. The confusion lies in the authors conflating two senses of "represent." It cannot mean "to define one thing in terms of another" while also meaning "be the cause of." Given that cause and effect are logically independent, it follows that no hypothetical construct can be, even partially, constructed from, or defined in terms of, its alleged effects.

Proponents of operationism thought they could deal with operationism's earlier critics by promoting multiple independent operations across different contexts or Campbell and Fiske's (1959) strategy of "convergent operationalism," so as to address any "mono-operation bias" (e.g., Reis and Judd, 2014, p. 18). This was considered appropriate because "… operational definitions generally 'underdetermine' (i.e., only partly define) perfectly valid concepts" (Rosenthal and Rosnow, 2008, p. 46). Yet these adjustments do not address the *illogic* of operational definitions, viz., the confusion between what the putative kind *is* — its essential or defining features — and the means by which those features are identified. Green (2001) expresses the problem this way: you have to recognise the concept *apart from* the operations employed to manipulate or measure it, but if there is no independent concept, there is "… nothing for the operations to converge upon" (p. 47). "Recognising the concept" involves scientific definition and for that, as Green makes clear, the kind's features must have an existence logically independent of method. "Convergent operationalism" would, at best, provide the class of operations involved in bringing about the putative kind but this is not to identify the kind's essential features.

In more recent text-book treatments of operationism, as Slife, Wright, and Yanchar (2016) have noted, we often see a yoking of definition with measurement. The texts I have quoted from are not unique in this respect — others also run the two together. For example: "… how the [latent] construct is measured can shape how we understand or define it" (Barker, Pistrang, and Elliott, 2016, p. 53); operational definition translates the concept into "… terms of the operations (procedures, actions, or processes) by which it could be observed and measured" (Zedeck, 2014, p. 245). The difficulty here is that such yoking yields an invalid a priori answer to the scientific question about what the concept or kind *is*. The invalid a priori answer is "Regardless of what it is, it can be measured — *it* is a continuous quantity." Yet we cannot say what *it* is — the concept currently has no intensional meat. And whether *it* has quantitative structure and is, therefore, a continuous quantity is an empirical hypothesis, not an a priori assumption

(Michell, 1997, 2012). It is contingent on having first identified the concept's defining features or conditions and only then working to determine whether each is classificatory, ordinal, or quantitative. Scientific definition is a logically prior scientific task to assessment.

As psychology's critics of operationism have counselled, operationism is not only illogical, it is inimical to sound scientific practice. Unsurprisingly, then, it has shown itself to be a relatively unproductive scientific "tool" (e.g., Borsboom, 2006; Maul, 2017). How limiting is it to "define" short-term memory as "… the longest list of digits on which the participant has perfect recall in more than 80% of the trials" or intelligence as "… that which is measured by this particular intelligence test" (Coolican, 2014, p. 38)? Tying a variable, concept, or putative kind to a set of operations, procedures or activities, leaves a profusion of splinter concepts (Byerly and Lazara, 1973; Michell, 1990) and the absence of unification noted at the beginning of this article.

What can Legitimately be Meant by "Operationism"?

Psychologists' use of the term "operationism" may not always signal anything illogical. When they mean only "the specifying of operations or procedures employed to test for, bring about, assess," etc. a variable, effect or phenomenon, this is unobjectionable. *But it is not definition* and it is at odds with the afore mentioned standard treatment of what an operational definition is. Let's call it "procedurism." Sometimes, psychologists' use of the expression "operational definition" just refers to procedurism when their research is that of scientific definition. Gigerenzer (2017), for example, identifies what the "reiteration effect" is and then sets out what he says will be a precise operational definition of the phenomenon (p. 136). But it is not — he in fact sets out the experimental procedure or operations designed to test for or bring about that effect accompanied by a simple formula to estimate effect size. This is not definition. The same occurs with "hindsight bias" where *what it is* is identified by Gigerenzer separately from the experimental design (p. 136). Koch's (1992) recollections of mid twentieth century psychology capture Gigerenzer's twenty-first century oversight:

> Within a year or two after the word [operationism] was first heard (probably via one of Stevens' 1935 articles), virtually everyone in psychology — not to mention their relatives and their dogs — was some kind of operationist. *It was as if the adjective "operational" had become cemented to the noun "definition."* (p. 269, italics added)

This conflating of two distinct scientific practices — reporting design or method details and defining the putative kind or concept — is also evident in the Kerlinger and Lee (2000) text. In the very same paragraph that I have already quoted from, the authors also say:

> An operational definition assigns meaning to a construct or a variable by specifying the activities or "operations" necessary to measure it and evaluate the measurement. *Alternatively, an operational definition is a specification of the activities of the researcher in measuring a variable or manipulating it. An operational definition is a sort of manual of instructions to the investigator.* (p. 42, italics added)

In consecutive sentences, Kerlinger and Lee (2000) have identified an operational definition as both the operations constituting the meaning of a variable *and* the operations merely being made explicit in the study of that variable. The former is illogical, the latter is the necessary procedurism but, again, procedurism is not definition.

Apropos such nonsensical blending, what scientific definition actually involves, the putative kind, *K*, should not be conflated with one or more of the following:

1. the conditions, processes or procedures that brought *K* about, i.e., what has caused *K*.
2. the signs of *K*.
3. the effects or consequences of *K*.
4. *K*'s function.
5. the so-called "measurement" of *K*.

In short, the relations or connections that *K* is involved in tell us nothing about what it is to be *K*. In particular, the pervasive, singular presence of the operational definition in psychology since the 1930s has many treating the words "operationalize" and "define" as synonymous when, in fact, to operationalize is *not* to define at all, nor can it be. Operationism is not coherent and procedurism is not definition. Regardless of whether psychologists operationally define their variables or merely engage in the necessary procedurism, neither scenario provides anything approaching scientific definition.

Conclusion

Prompted by psychology's neglect of scientific definition, this paper has provided a formal account of how scientific definition has been traditionally understood, that is through the classical/standard realist account. The paper has also outlined a presumed alternative to this account — the homeostatic property cluster (HPC) account. In contemporary philosophy of science, the HPC account prevails primarily because it accommodates an assumption that some scientific kinds are without a fixed set of defining features or conditions as they change over time. However, a number of arguments cast doubt on this. The HPC account appears not to be a *genuine* alternative to the standard account because of its logical dependence on the latter. At best, it accommodates the early to mid stages of scientific definition.

Given that scientific definition identifies a kind's essential features or conditions, it is quite different from psychology's standard treatment of the operational definition. Here, the putative kind or concept is confused with the means by which it is

to be identified. Obviously genuine expertise in *what* is being assessed, measured, predicted, explained, etc. cannot rest on this confusion. It depends on identifying the kind's essential features, not on identifying that kind with the operations that brought it about, nor with any other connection that the kind is involved in.

I assume that scientific definition matters in psychology because without it, the discipline is left with the notorious vagueness and ambiguity of everyday language and the errors (logical and empirical) that so readily emanate; because certain areas of psychology are particularly prone to employing daily concepts on the false assumption that "we all know what they mean;" because the more a discipline's concepts rest uncritically on that assumption the less the discipline advances; because if the supposed definition of a concept is variously based on, for example, operations involved in task performance, or working memory, or inhibitory control, the pointless disagreements that follow divert attention from real theoretical research; because, given the standard classical realist account, where the essential features or structure of a concept have not been identified, it follows that we are mistaken in assuming that we have expertise in *what* that concept or kind is; and because the distinction between essential and accompanying features or conditions is scientifically significant. We may have evidence of the presence of various features, but until we can identify which are defining/essential, we cannot justify claims about *what* exactly is being assessed, measured, predicted, explained, manipulated, controlled, etc. Where conceptual replication is concerned, for example, this means that testing the *same* concept from an earlier study requires that concept's *defining* features to have been made explicit. It seems odd that, despite the epistemic virtues scientific definition brings, the topic has been overlooked by the very sources aiming to facilitate psychology's scientific literacy.

If the arguments here are sound, the direction for psychology is clear. Research should leave operationism behind and work towards the standards that comprise the standard realist account of scientific definition, treating the HPC account as a way station. Explaining those standards should feature in outlets which aim to facilitate psychology's scientific literacy and in the education of psychology students. Of course, when the mantra to "operationally define" has been ingrained across generations of psychologists, persuading the discipline's gatekeepers to withdraw their endorsement of operationism will not be easy. And psychologists often appear unconcerned about such matters. Nevertheless, operationism is a practice at odds with their realism because it is at odds with the ontology of psychological phenomena and so completely undermines what the discipline is concerned to achieve.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (fifth edition). Arlington, Virginia: American Psychiatric Association.

American Psychological Association. (2012). *APA handbook of research methods in psychology*. Washington, D.C.: American Psychological Association.

Armstrong, D. M. (1997). *A world of states of affairs*. Cambridge: Cambridge University Press.
Barker, C., Pistrang, N., and Elliott, R. (2016). *Research methods in clinical psychology: An introduction for students and practitioners* (third edition). Malden, Massachusetts: John Wiley & Sons Ltd. doi:10.1002/9781119154082
Beckner, M. (1959). *The biological way of thought*. Berkeley: University of California Press.
Benjamin, A. C. (1955). *Operationism*. Springfield, Ilinois: Charles C. Thomas.
Berti, E. (2011). The contemporary relevance of Aristotle's thought. *Iris, 3*(6), 26–35.
Bickhard, M. H. (2001). The tragedy of operationalism. *Theory and Psychology, 11*(1), 35–44.
Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika, 71*(3), 425–440.
Boyd, R. N. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies, 61*, 127–148.
Boyd, R. N. (1999). Homeostasis, species, and higher taxa. In R. A. Wilson (Ed.), *Species: New interdisciplinary essays* (pp. 141–185). Cambridge, Massachusetts: MIT Press.
Boyd, R. N. (2010). Homeostasis, higher taxa, and monophyly. *Philosophy of Science, 77*, 686–701.
Boysen, G. A., and Ebersole, A. (2014). Expansion of the concept of mental disorder in the DSM-5. *Journal of Mind and Behavior, 35*(4), 225–244.
Bridgman, P. W. (1927). *The logic of modern physics*. New York: Macmillan.
Burali-Forti, C. (1898). Les propriétés formales des opérations algébriques. *Rivista di Matematica della Università di Parma, 6*, 141–177.
Byerly, H. C., and Lazara, V. A. (1973). Realist foundations of measurement. *Philosophy of Science, 40*(1), 10–28.
Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81–105.
Clark, L. A., and Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309–319.
Coolican, H. (2014). *Research methods and statistics in psychology*. East Sussex: Routledge.
Copi, I. M. (1954). *Symbolic logic*. New York: Macmillan.
Devitt, M. (2008). Resurrecting biological essentialism. *Philosophy of Science, 75*, 344–382.
Ehrlich, P. (2006). The rise of non-Archimedean mathematics and the roots of a misconception I: The emergence of non-Archimedean systems of magnitudes. *Archive for History of Exact Sciences, 60*(1), 1–121. doi:10.1007/s00407-005-0102-4
Einstein, A. (1949). Remarks concerning the essays brought together in this co-operative volume. In P. A. Schilpp (Ed.), *Albert Einstein: Philosopher–Scientist* (Vol. II, pp. 665–688). New York: Harper and Brothers.
Feest, U. (2005). Operationism in psychology: What the debate is about, what the debate should be about. *Journal of the History of the Behavioral Sciences, 41*(2), 131–149.
Ghiselin, M. T. (1966). On psychologism in the logic of taxonomic controversies. *Systematic Zoology, 15*(3), 207–215.
Gigerenzer, G. (2017). A theory integration program. *Decision, 4*(3), 133–145.
Grayling, A. C. (1997). *An introduction to philosophical logic*. Oxford: Blackwell.
Green, C. D. (2001). Operationism again. What did Bridgman say? What did Bridgman need? *Theory and Psychology, 11*(1), 45–51.
Groarke, L. (2009). *An Aristotelian account of induction*. Montreal: McGill–Queen's University Press.
Haslam, N., Holland, E., and Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: A quantitative review of taxometric research. *Psychological Medicine, 42*, 903–920.
Helmholtz, v. H. (1977). Numbering and measuring from an epistemological viewpoint [M. F. Lowe, Trans.]. In R. S. Cohen and Y. Elkana (Eds.), *Epistemological writings: The Paul Hertz/Moritz Schlick centenary edition of 1921 with notes and commentary by the editors* (pp. 72–114). Dordrecht: D. Reidel. (Originally published 1887)
Hölder, O. (1996). The axioms of quantity and the theory of measurement. Translated from Part I of Otto Hölder's German text "Die Axiome der Quantität und die Lehre vom Mass" by Joel Michell and Catherine Ernst. *Journal of Mathematical Psychology, 40*, 235–252. (Originally published 1901)
Hölder, O. (1997). The axioms of quantity and the theory of measurement. Translated from Part II of Otto Hölder's German text "Die Axiome der Quantität und die Lehre vom Mass" by Joel Michell and Catherine Ernst. *Journal of Mathematical Psychology, 41*, 345–356. (Originally published 1901)

Kagan, J. (2012). *Psychology's ghosts: The crisis in the profession and the way back*. New Haven: Yale University Press.
Keller, R. A., Boyd, R. N., and Wheeler, Q. D. (2003). The illogical basis of phylogenetic nomenclature. *The Botantical Review, 69*(1), 93–110.
Kerlinger, F. N., and Lee, H. B. (2000). *Foundations of behavioral research* (fourth edition). South Melbourne, Australia: Thomson Learning.
Koch, S. (1992). Psychology's Bridgman vs Bridgman's Bridgman: An essay in reconstruction. *Theory and Psychology, 2*, 261–290.
Kripke, S. (1980). *Naming and necessity*. Cambridge, Massachusetts: Harvard University Press.
Leahey, T. H. (1980). The myth of operationism. *Journal of Mind and Behavior, 1*(2), 127–143.
Markon, K. E., Chmielewski, M., and Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin, 137*(5), 856–879.
Maul, A. (2017). Rethinking traditional methods of survey validation. *Measurement: Interdisciplinary Research and Perspectives, 15*(2), 51–69. doi:10.1080/15366367.2017.1348108
Mayr, E. (1963). *Animal species and evolution*. Cambridge, Massachusetts: Harvard University Press.
Michell, J. (1990). *An introduction to the logic of psychological measurement*. New Jersey: Lawrence Erlbaum.
Michell, J. (1997). Quantitative science and the definition of *measurement* in psychology. *British Journal of Psychology, 88*, 355–383.
Michell, J. (2012). Alfred Binet and the concept of heterogeneous orders. *Frontiers in Psychology, 3*, 1–8. doi:10.3389/fpsyg.2012.00261
Mulligan, K., and Scherer, K. R. (2012). Toward a working definition of emotion. *Emotion Review, 4*(4), 345–357. doi:10.1177/1754073912445818
Orr, D. (2002). *The nature of design: Ecology, culture, and human intention*. Oxford: Oxford University Press.
Putnam, H. (1962). It ain't necessarily so. *Journal of Philosophy, 59*(22), 658–671.
Putnam, H. (1975). *Mind, language and reality*. Cambridge: Cambridge University Press.
Reis, H. T., and Judd, C. M. (Eds.). (2014). *Handbook of research methods in social and personality psychology* (second edition): Cambridge University Press.
Rosenthal, R., and Rosnow, R. L. (2008). *Essentials of behavioral research: Methods and data analysis* (third edition). Boston: McGraw–Hill.
Rossiter, J. R. (2017). Optimal standard measures: Comment on Matthews et al. (2016). *American Psychologist, 72*(5), 489–490.
Russell, B. (1984). *A history of western philosophy*. London: Unwin Hyman Ltd. (Originally published 1946)
Russell, L. J. (1928). Review of Bridgman's 'The logic of Modern Physics'. *Mind, 37*, 355–361.
Schlick, M. (1979). Meaning and verification. In H. L. Mulder and B. F. B. van de Velde–Schlick (Eds.), *Moritz Schlick: Philosophical papers* (Vol. II, pp. 456–481). Dordrecht: D. Reidel Publishing. (Originally published 1936)
Slife, B. D., Wright, C. D., and Yanchar, S. C. (2016). Using operational definitions is research: A best-practices approach. *Journal of Mind and Behavior, 37*(2), 119–140.
Sokal, R. R. (1974, Sept 27). Classification: Purposes, principles, progress, prospects. *Science, 185*(4157), 1115–1123.
Sokal, R. R., and Sneath, P. H. A. (1963). *Principles of numerical taxonomy*. San Fransisco: W. H. Freeman and Co.
Stevens, S. S. (1935). The operational basis of psychology. *American Journal of Psychology, 47*, 323–330.
Stevens, S. S. (1939). Psychology and the science of science. *Psychological Bulletin, 36*, 221–263.
Sutcliffe, J. P. (1993). Concept, class, and category in the tradition of Aristotle. In I. Van Mechelen, J. Hampton, R. S. Michalski, and P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 35–65). London: Academic Press Ltd.
Sutcliffe, J. P. (1999). *A critical enquiry into the classification movement of the last half-century*. A compilation of published research submitted for the Doctor of Science degree, School of Psychology, University of Sydney.
Wallach, L. (1971). Implications of recent work in philosophy of science for the role of operational definition in psychology. *Psychological Reports, 28*, 583–608.

Whitley, B. E., and Kite, M. E. (2013). *Principles of research in behavioral science* (third edition). New York: Routledge.
Wittgenstein, L. (1967). *Philosophical investigations* [G. E. N. Anscombe, Trans., third edition]. Oxford: Basil Blackwell. (Originally published 1953)
Wittgenstein, L. (1975). *Philosophical remarks*. Oxford: Basil Blackwell. (Originally published 1930)
Wrigley, A. (2015). An eliminativist approach to vulnerability. *Bioethics, 29*(7), 478–487.
Zedeck, S. (Ed.) (2014). *APA dictionary of statistics and research methods*. Washington: American Psychological Association.