A Defense and Definition of Construct Validity in Psychology

Caroline Stone*

Psychologists say a measure has construct validity when it, in fact, measures the construct it is intended to measure. Construct validity is both an important notion in psychological research methods and the source of much confusion and debate among psychologists. I argue that this confusion arises, in part, because of a failure to distinguish between construct validity, a feature of measures relative to a construct, and construct legitimacy, a feature of the construct itself. I propose a prescriptive account of construct validity on the basis of this distinction, then provide evidence for my account through two examples from research psychology.

1. Introduction. Controversies in psychology often turn on the construct validity of key measures definitive of a research program.[1] Psychologists say that a measure has *construct validity* when it measures what it is intended to measure, a psychological construct like intelligence.[2] Establishing the construct validity of a measure is difficult. Even measures that are well established and widely used are subject to criticism and sometimes drastic revision. For example, despite being the most commonly used—and one of the oldest—measures of intelligence, some psychologists still object to the Wechsler Adult Intelligence Scale-III on the basis of evidence that it is culturally biased (Gottfredson and Saklofske 2009). A more recently developed measure, the implicit association test (IAT), has been widely adopted by researchers studying implicit attitudes despite arguments that the IAT fails to both predict behavior

*To contact the author, please write to: Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO 63130; e-mail: carolinestone@wustl.edu.

1. In psychology, "validity" refers to the property of being sufficiently supported by evidence, rather than a property of formalized arguments as it does in logic. For the purposes of this article, "validity" will always be used in the psychological sense.

2. "Construct" can be replaced with "theoretical term," which is used more commonly in philosophical contexts. I use the term "construct" to respect the common nomenclature in research psychology.

Philosophy of Science, 86 (December 2019) pp. 1250–1261. 0031-8248/2019/8605-0037$10.00
Copyright 2019 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/705567 Published online by Cambridge University Press

and have sufficient construct validity (De Houwer et al. 2009). In comparative cognition, measures of cognitive capacities developed and validated in human and primate populations are extended to novel populations. A study by Ari and D'Agostino (2016) claimed to demonstrate prerequisite behavior for self-recognition in manta rays. Yet it remains an open question whether their measure, Gallup's mirror self-recognition test, is construct valid for nonprimate species (Thünken et al. 2009; Brown 2014). Within psychopathy, practitioners and researchers are concerned with the construct validity of the behavioral criteria used to diagnose mental illness, referred to as diagnostic validity (Feighner and Herbstein 1987). Since the publication of Cronbach and Meehl's (1955) seminal paper "Construct Validity in Psychological Tests," construct validity has been both an important concept in psychological measurement and the source of much confusion and debate. Different theories of construct validity disagree about what construct validity is and what the requirements are for establishing the construct validity of a measure. In this article, I propose a prescriptive account of construct validity that distinguishes construct *validity*, a feature of a measure relative to a construct, from construct *legitimacy*, a feature of the construct itself. This account can resolve the ambiguity about what construct validity is and, further, permits a better understanding of current debates about psychological measures.

In section 2 of this article, I further motivate the need for an account of construct validity by detailing sources of both confusions and disagreement. In section 3, I outline my account of construct validity and distinguish it from a related, and often conflated, concept: *construct legitimacy*. In section 4, I provide evidence for this distinction by way of examples in psychological research in which the two concepts come apart. In section 5, I apply my account to the debates outlined in section 2 and show how this account can better explain, and even resolve, them.

2. An Amorphous Concept? In their account, Cronbach and Meehl (1955) presented construct validity as one of four types of test validity.[3] They defined construct validity as the ability of a test to actually measure the intended construct, that "must be investigated whenever no criterion or universe of content is accepted as entirely adequate to define the quality to be measured" (Cronbach and Meehl 1955, 282). Their account was a rejection of operationalist views like Anastasi's (1950, 67), who claimed that psychological measures could only be validated as measures of specifically defined, observable

3. The other three types of test validity are (1) predictive validity, the degree to which a test can predict future criteria, (2) concurrent validity, the degree to which a test agrees with criteria that are determined at the same time, and (3) content validity, the degree to which the test items are a sample of the universe in which the investigator is interested (Cronbach and Meehl 1955, 282).

https://doi.org/10.1086/705567 Published online by Cambridge University Press

criteria. Construct validity was a concept that allowed for constructs that could not be exhaustively defined by their operations.

Cronbach and Meehl focused more on construct validation practices, rather than defining the concept of construct validity itself. It's unsurprising, then, that ambiguity about construct validity remains. Psychologists following Cronbach and Meehl have attempted to further elucidate what construct validity is, while developing methods of construct validation.

For example, in Campbell and Fiske's (1959) introduction of the first direct statistical measure of construct validity they distinguish between "the adequacy of tests as measures of a construct" and "the adequacy of a construct as determined by the confirmation of theoretically predicted associations with measures of other constructs" (83). Their focus was on the former, rather than the latter, and in making this distinction they imply that they take Cronbach and Meehl's notion of construct validity to encompass both types of adequacy.

While many psychologists have utilized Campbell and Fiske's distinction (e.g., Messick 1995), others have argued against this distinction, insofar as it would be possible to have one type of adequacy without the other. In discussing the construct validity of implicit attitude tests, De Houwer et al. (2009) argue that construct validity requires specifying the causal relationship between construct and measure. Specifying this causal relationship further requires that the construct refers to a psychological entity that exists.

Despite the disagreement and confusion surrounding construct validity, it is an important and useful concept used in both ongoing debates about new and old psychological measures and current cognitive ontology projects in philosophy of cognitive science. In personality psychology, for example, construct validity remains a prescient concern; a search for articles about the construct validity of personality measures shows more than one thousand articles were published on the topic in 2017 alone (Google Scholar, January 17, 2018). In philosophy of cognitive science, philosophers are concerned about construct validity in specific cases—for example, Feest's (2017) work on the construct validity of the IAT—along with more generally how construct validity affects efforts to localize cognitive capacities in the brain (Poldrack 2006; Poldrack and Yarkoni 2016; Sullivan 2016). Thus while construct validity is an important notion in psychological research methods, ambiguity remains: "[construct validity] is now generally recognized as *the central concern* in psychological measurement. . . . Yet construct validity continues to strike many of us, from graduate students to senior professors, as a rather nebulous or 'amorphous' concept" (John and Soto 2007, 475).

3. A Model of Construct Validity. In the previous section I highlighted two types of adequacy underlying Cronbach and Meehl's original definition: adequacy of a test as a measure of a construct, and adequacy of a construct.

https://doi.org/10.1086/705567 Published online by Cambridge University Press

The account of construct validity I give here maintains this distinction, but I define construct validity as referring only to the former type of adequacy and not the latter. The latter type of adequacy I call construct legitimacy. The difference between construct legitimacy and validity is illustrated in figure 1.

The model in figure 1 shows three relationships between the world, a measure, and a construct. Roughly, construct validation is about the justification of inferences from test scores to claims about a construct, while construct legitimization is about the justification of a construct's theory/model. Existence claims are claims about whether the construct refers to an entity out in the world.

Construct validity is a property of the relationship between a measure and a construct, such that a measure has construct validity relative to a specific construct, if we have sufficient evidence that it tracks that construct. Construct validity requires, at bare minimum, that we have evidence we have isolated a distinct construct. The convergent and discriminant validity from Campbell and Fiske's (1959) account is exactly the type of evidence that would indicate that a measure isolates a distinct construct.

Construct validity applies to specific inferences made on the basis of scores from a measure. My definition here follows Messick's (1995) by focusing on the justification of inferences from test scores. However, I take construct validity to apply only to inferences to claims about a construct, rather than any inference. My definition can be formalized as follows:

> **Construct validity:** A measure, *m*, of some construct, *C*, is construct valid if and only if we are sufficiently justified in believing *m* tracks *C*.

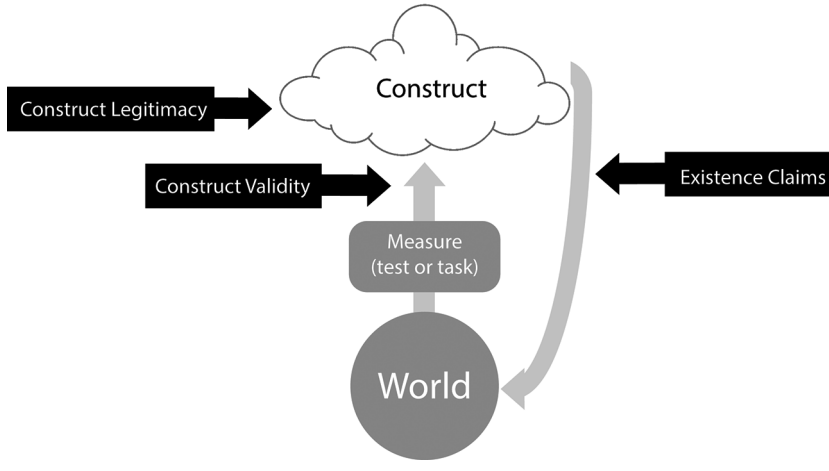


Figure 1.

https://doi.org/10.1086/705567 Published online by Cambridge University Press

Construct legitimacy, in contrast, is a property of a construct itself, relative to the theory or model it is a part of. A construct can be said to be more or less legitimate depending on how well the definition or model of the construct is supported by the available evidence and how useful the construct is in the larger theory. We say that a construct is legitimate, relative to a particular model or theory, when that model or theory is sufficiently justified. My definition of construct legitimacy can be formalized as follows:

> **Construct legitimacy:** A construct, *C*, is legitimate if and only if the theory that posits *C* is sufficiently justified.

The criteria for legitimacy of a construct will depend on both our standards for evaluating theories and our metaphysical commitments about the referents of constructs.

While I disagree with De Houwer et al.'s (2009) criteria as it pertains to construct validity, explicating the causal mechanism underlying a construct provides strong evidence for construct legitimacy. As legitimacy is dependent on the status of the construct's theory, legitimacy claims will often appeal to the particular theory. For example, "Phlogiston is an illegitimate construct because phlogiston theory is not supported by evidence."

Existence claims about a construct are claims about the relationship between the construct, a theoretical concept, and the object(s) in the real world it is intended to pick out. As legitimacy depends on metaphysical commitments about the psychological entities that constructs refer to, a realist about psychological entities may not take construct legitimacy and existence claims to be distinct.

My account here is theory-neutral in that I make no claim for specific criteria that would justify each type of claim. Doing so requires epistemic commitments about theory justification, along with metaphysical commitments about the relationship between constructs and psychological entities. My purpose here is to merely argue that these two concepts, legitimacy and validity, come apart; they can vary independently of each other.

While I will argue here that legitimacy and validity are distinct concepts, they are still partially dependent on each other. Whether a particular test measures an intended construct will depend on what we think that construct is. De Houwer et al. (2009), in contrast, take legitimacy as a necessary prerequisite for construct validity. The consequence of this account is that most psychological measures fail to be construct valid, as we do not know the causal relationship between measure and construct. I propose instead that construct validity can be limited, but not prohibited, by the degree of legitimacy of the target construct. Construct validity requires that the construct a test measures adequately tracks a psychological construct. However, adequately capturing a construct does not necessitate that we have a complete

https://doi.org/10.1086/705567 Published online by Cambridge University Press

model or explanation of that construct, or that the construct refers to an entity that exists. By isolating a distinct construct, I have in mind Campbell and Fiske's (1959) discriminant and convergent validity. A measure then isolates a distinct construct if it converges with other measures of the same construct but discriminates from measures of distinct constructs.

The final part of my account is that construct validity (and construct legitimacy) comes in degrees. A test is more or less valid depending on the evidence for the relationship between the test and the target construct. While not present in the original definition by Cronbach and Meehl, this view is shared by both psychologists and philosophers writing about construct validity more recently (Messick 1995; Smith 2005; Feest 2017). A natural consequence of this view is that both validity and legitimacy are evolving properties, such that a construct or measure may become more or less legitimate/valid over time as new evidence is gathered.

There are good intuitive reasons to distinguish between construct validity and legitimacy. It is easy to find examples of cases in which a test irrefutably has construct validity, even though the construct itself is an illegitimate one using any scientific standard. Take any test of your zodiac sign. Your zodiac sign depends on your birth date, something that we are incredibly accurate at measuring. Thus, zodiac sign measures are as accurate as our measures of birth dates. However, no psychologist (that I know of) would claim zodiac signs are a legitimate psychological construct. In this case, there is sufficient evidence for the construct validity of zodiac sign tests even though zodiac signs are illegitimate.

4. Two Examples from Research Psychology. I now turn to two debates in psychology that illustrate how construct validity and legitimacy come apart in psychological research. These debates show that a measure can have a reasonable degree of construct validity even when the construct it measures has a low degree of construct legitimacy. I will argue that they demonstrate why this distinction is necessary. Trying to understand either case without distinguishing between construct legitimacy and validity would lead to erroneous abandonment of measures that are useful research tools.

4.1. The Implicit Association Test. The IAT (Greenwald, McGhee, and Schwartz 1998) measures implicit attitudes by comparing reaction times between paired concepts. Participants are instructed to sort stimuli that are examples of four concepts; however, in each trial the concepts are split into two pairs (e.g., male-career; female-home), so participants sort four types of stimuli into two categories. Researchers infer a subject's implicit attitudes from a comparison between average reaction times to each association pair. The faster a participant is able to sort stimuli from a paired concept, the stronger that subject's implicit association between the concepts. Thus, if a participant

https://doi.org/10.1086/705567 Published online by Cambridge University Press

sorts male-career stimuli faster than female-career stimuli, they would be judged to have a stronger implicit association between males and careers.

Despite the popularity of the IAT, widespread disagreement remains about what exactly implicit attitudes, the construct the IAT is intended to measure, are. One major source of disagreement is about the process of learning and retrieving implicit attitudes. Nosek and Banaji's (2009) account distinguishes between implicit and explicit attitudes on the basis of retrieval processes. Both are learned the same way, but only implicit attitudes are accessible through an unintentional retrieval process. Gawronski and Bodenhausen (2006), however, take explicit and implicit attitudes to be the result of two different learning processes. Further, Gawronski and Bodenhausen (2006, 696) claim "people have some degree of conscious access to their automatic affective reactions," while Nosek and Banaji (2009) claim people do not.

There is also disagreement about how implicit and explicit attitudes are represented. Nosek and Banaji take implicit and explicit attitudes to be different mental representations, which afford different retrieval processes. However, Fazio and Olson (2014) argue that implicit and explicit attitudes are represented the same way but retrieved differently.

On the basis of this theoretical disagreement, Feest (2017) concludes that implicit attitude is a 'lumpy construct'. There is some agreement (there exists implicit and explicit retrieval processes of attitudes that result in different responses) along with widespread disagreement (whether the different processes retrieve different types of representations, what form the representations take, and how the representations are formed).

The incompleteness of the theories associated with the construct limits the legitimacy of the implicit attitudes. However, the construct does not need to be rejected. Implicit attitudes have an explanatory role, particularly in cases in which explicit attitudes are unable to accurately predict behavior (Feest 2017, 6). Implicit attitudes have this explanatory role even if the theory is incomplete. Rather, implicit attitude's low degree of legitimacy reflects the need for more evidence to arbitrate between the multiple, conflicting hypotheses about the nature of the construct.

The question relevant to my account is whether the construct validity of the IAT as a measure of implicit attitudes is prohibited or merely limited by the relatively low legitimacy of the construct. Some psychologists, most notably De Houwer et al. (2009), have argued that it cannot, because we cannot explicate the causal mechanism underlying implicit attitudes that would explain scores on the IAT. Feest, in contrast, believes the IAT has construct validity despite the issues with implicit attitudes.

In my account, the construct validity of the IAT is not wholly dependent on the legitimacy of the implicit attitudes' construct. Rather, the myriad of other forms of evidence for the construct validity of the IAT can be sufficient in spite of the target construct's relative illegitimacy. Feest points to the explanatory role IAT scores have for certain discriminatory behavior, even

https://doi.org/10.1086/705567 Published online by Cambridge University Press

when predictive validity is low (Greenwald, Banaji, and Nosek 2015). There is also significant convergent evidence for the IAT: IAT measures correlate strongly with explicit attitude measures and with other implicit measures. Importantly, the strength of this association (between implicit and explicit measures of the same attitude) is domain dependent and mediated by an individual's concern with expressing the attitude (Nosek 2005). This means that implicit attitudes usually correspond with our explicit attitudes, except in situations where an individual feels pressure to not express a particular attitude.

Taking this evidence into account, I agree with Feest (2017) that the IAT has construct validity relative to implicit attitudes. Results from the IAT have predictive power, and psychologists have argued that even when effect sizes are small, the practical implications can be large. However, the degree of construct validity could be improved in the future if we develop a better unified model of implicit attitudes that in turn leads to improvements of the IAT. Note that the positive relationship between legitimacy and validity here is indirect. Increasing the legitimacy of the measured construct does not necessarily increase the validity of the construct's measures. Increasing legitimacy increases validity only when such theoretical advances enable improvements in the measure itself—for example, through identification of previously unknown confounding variables that can now be controlled for.

4.2. Implicit Memory. Roediger's (2003) argument against implicit memory demonstrates how a construct can be illegitimate even when the measures of the construct themselves have a significant degree of construct validity. According to Roediger, implicit memory is an illegitimate construct because of a lack of evidence that the construct refers to a uniquely cognitive phenomena, not because the construct is lumpy.

Compared to implicit attitudes, there is a surprising lack of disagreement about how to define implicit memory. *Implicit memory* is a psychological construct that is usually defined in contrast with explicit memory. *Explicit memory* refers to memory that is intentional and therefore can be both recalled and reported by a subject, whereas implicit memory is unintentionally recalled. Implicit memory is commonly measured using motor skill learning tasks or priming tasks (McDermott 2000). In priming tasks, subjects are primed with a set of words, then later tested with both the primed (old) words and new words. Faster reaction times to primed words indicates a priming effect.

Warrington and Weiskrantz (1970) demonstrated that amnesic patients had impaired explicit memory but did not demonstrate impaired performance on word-stem and word-fragment completion tasks. In order to explain the single dissociation, psychologists proposed a new memory capacity: implicit memory.

While this single-dissociation evidence has been replicated numerous times, Roediger (2003, 4) argues that implicit memory is an illegitimate construct because almost any biological system, regardless of whether it involves

https://doi.org/10.1086/705567 Published online by Cambridge University Press

the brain or not, can be said to have implicit memory. For example, the immune system exhibits implicit memory. When the immune system is attacked by a previously encountered pathogen, its response changes, reflecting a 'memory' of the pathogen. Thus, implicit memory is not a uniquely psychological capacity.

This example shows a disconnect between the broad theoretical definition of implicit memory and its narrow operationalization. The problem here is not simply that psychologists have failed to measure the entire set of phenomena that could be considered cases of implicit memory. Rather, the range of phenomena covered by the theoretical definition of implicit memory is so broad that it includes phenomena psychologists would not consider to be cases of implicit memory at all, regardless of their ability to measure them.

Roediger concludes that any commonality between instances of implicit memory exists at a level of general principle rather than at the level of a specific mechanism. Biological systems respond differently to stimuli they have encountered before, but there is no implicit memory capacity explained by some specific cognitive mechanism (Roediger 2003, 13). Put more simply, Roediger believes that legitimate psychological constructs must have a unique cognitive mechanism.

If the implicit memory construct is illegitimate, it cannot be made legitimate by merely redefining the construct to refer to only cognitive phenomena. Roediger views redefining the construct as ad hoc, since without a unique cognitive mechanism, implicit memory is not a uniquely cognitive phenomenon. Rather, it is "a memory function frequently incidental to the operation of the system which subserves a different basic function" (Roediger 2003, 14).

It is hard to make sense of Roediger's argument here unless we distinguish between legitimacy and validity. Roediger has no issue with the actual measures of implicit memory or the effects observed using these measures. His only issue is with the justification for proposing implicit memory as a distinct memory capacity. In other words, Roediger believes implicit memory is an illegitimate construct, even though the measures associated with the construct have construct validity. They measure what we intend for them to measure; however, what they measure is not a uniquely psychological phenomena.

If a unique cognitive mechanism was identified, thereby legitimizing the implicit memory construct, this would only indirectly, and not necessarily, increase the construct validity of its measures. One could imagine a cognitive mechanism that narrowed the scope of implicit memory to only perceptual domains. This would then in turn eliminate some implicit memory measures and potentially increase convergence across the remaining measures.

5. Further Defense of the Distinction. I have claimed that the previous two examples show how construct validity and construct legitimacy can vary independently of each other. I have also claimed that understanding these

https://doi.org/10.1086/705567 Published online by Cambridge University Press

examples requires a graded notion of both construct validity and legitimacy. Both the IAT and priming measures are construct valid, despite the problems with the constructs they measure.

De Houwer et al. (2009), in contrast, denies that any measure of implicit attitudes has construct validity because the construct itself is not well understood (e.g., has low construct legitimacy). Similar to De Houwer et al., Roediger's rejection of the construct "implicit memory" would entail denying that the tasks used to measure implicit memory have any construct validity. My alternative reading of these cases is more permissive. We can talk about the construct validity of measures in cases where the causal mechanism is unknown if we use a notion of construct legitimacy that varies independently of construct validity.

My argument thus far has been conceptual, focusing on the distinction between adequacy and legitimacy and not the strict criteria for construct validity that De Houwer et al. propose. A skeptic may grant this conceptual distinction but still maintain that construct validity requires construct legitimacy. The case against such strict criteria is more pragmatic. If construct validity requires establishing construct legitimacy to the degree that the causal mechanism linking a construct to the results of its measure must be known, construct validity ceases to be a useful concept. De Houwer et al.'s criteria forces us to either accept that most psychological measures are not construct valid or promote "unproductive debates over conceptual identification of attributes underlying psychological measures" (Nosek and Greenwald 2009, 373). Construct validity would be a regulative ideal. We could talk about how a measure is closer to construct validity than another, but absent from such talk is criteria for the actual use of a measure. If no measures, or at least very few measures, are construct valid, what justifies their use in research? Rather than relegating construct validity to an ideal and creating a need for a new standard for test use, I prefer using the concept we already have.

It is not the case then that the IAT has no construct validity because the construct it measures, implicit attitudes, is an ill-defined construct. Rather, the degree to which the IAT has construct validity is in part lessened, but not prohibited, because of our lack of understanding of the construct it is intended to measure. The construct validity of the IAT can be improved by increasing our understanding of implicit attitudes, which in turn would increase the degree of construct legitimacy implicit attitudes have. However there are other types of evidence that would increase the construct validity of the IAT independent of the legitimacy of the implicit attitudes construct. To name just a few: further increasing the IAT's predictive power for future behavior, or decreasing variability between scores from the same participant taking the IAT at different times (also known as test-retest reliability); either improvement would increase the construct validity of the IAT without increasing the legitimacy of the construct it measures. That does not mean

https://doi.org/10.1086/705567 Published online by Cambridge University Press

increasing the construct validity of the IAT would not, in turn, increase the legitimacy of the implicit attitudes construct. However, that improvement would be indirect, by way of more research on implicit attitudes with an improved, more powerful IAT that is subject to fewer confounding variables and sources of score error.

My account here explains current practices in social psychology, where the IAT is an incredibly useful measure. By distinguishing between legitimacy and validity we can explain how the IAT has sufficient construct validity to be used in psychological research even though the construct it measures is ill defined. Moreover, using a graded notion of construct validity allows one to claim the IAT has construct validity even though the degree of construct validity it has could be improved.

If implicit memory is an illegitimate construct, my account allows us to reject the construct without rejecting the tasks used to measure it. Again, my account is able to explain current practices in psychology, where these semantic and visual priming tasks are valid measures in their own right with applications outside of just implicit memory studies.

6. Conclusion. Construct validity is an important and useful concept in research psychology, when understood correctly. As I have argued, criticisms of construct validity have arisen partly because of a failure to adequately distinguish construct validity from construct legitimacy and a failure to use a graded notion of both.

Distinguishing between these two concepts illuminates certain debates in psychology about both measures and constructs. My account differentiates between issues of construct validity, where the measure of a construct is the focus (as in the debate about the IAT), and issues of construct legitimacy, where the usefulness of the construct itself is in question (as in Roediger's argument against implicit memory). By taking both legitimacy and validity as things that come in degrees, my account is able to further explain how measures like the IAT have a degree of construct validity that affords their current use in research, while allowing for future improvement as our understanding of the constructs they measure deepens. While both construct legitimacy and validity are indirectly related, neither is necessary or sufficient for the other.

REFERENCES

Anastasi, Anne. 1950. "The Concept of Validity in the Interpretation of Test Scores." *Educational and Psychological Measurement* 10 (1): 67–78.

Ari, Csilla, and Dominic P. D'Agostino. 2016. "Contingency Checking and Self-Directed Behaviors in Giant Manta Rays: Do Elasmobranchs Have Self-Awareness?" *Journal of Ethology* 34 (2): 167–74.

Brown, Culum. 2014. "Fish Intelligence, Sentience and Ethics." *Animal Cognition* 18 (1): 1–17.

https://doi.org/10.1086/705567 Published online by Cambridge University Press

Campbell, Donald T., and Donald W. Fiske. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin* 56 (2): 81–105.
Cronbach, Lee J., and Paul E. Meehl. 1955. "Construct Validity in Psychological Tests." *Psychological Bulletin* 52 (4): 281–302.
De Houwer, Jan, Sarah Teige-Mocigemba, Adriaan Spruyt, and Agnes Moors. 2009. "Implicit Measures: A Normative Analysis and Review." *Psychological Bulletin* 135 (3): 347–68.
Fazio, Russell H., and Michael A. Olson. 2014. "The MODE Model: Attitude-Behavior Processes as a Function of Motivation and Opportunity." In *Dual-Process Theories of the Social Mind*, ed. Jeffrey W. Sherman, Bertram Gawronski, and Yaacov Trope, 155–71. New York: Guilford.
Feest, Uljana. 2017. "Construct Validity in Psychological tests—the Case of Implicit Social Cognition." Presented at Washington University in St. Louis.
Feighner, John P., and Jessica Herbstein. 1987. "Diagnostic Validity." In *Issues in Diagnostic Research*, ed. Cynthia Last and Michel Hersen, 121–40. Boston: Springer.
Gawronski, Bertram, and Galen V. Bodenhausen. 2006. "Associative and Propositional Processes in Evaluation: An Integrative Review of Implicit and Explicit Attitude Change." *Psychological Bulletin* 132 (5): 692–731.
Gottfredson, Linda, and Donald H. Saklofske. 2009. "Intelligence: Foundations and Issues in Assessment." *Canadian Psychology/Psychologie canadienne* 50 (3): 183–95.
Greenwald, Anthony G., Mahzarin R. Banaji, and Brian A. Nosek. 2015. "Statistically Small Effects of the Implicit Association Test Can Have Societally Large Effects." *Journal of Personality and Social Psychology* 108 (4): 553–61.
Greenwald, Anthony G., Debbie E. McGhee, and Jordan L. K. Schwartz. 1998. "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test." *Journal of Personality and Social Psychology* 74 (6): 1464–80.
John, Oliver P., and Christopher J. Soto. 2007. "The Importance of Being Valid: Reliability and the Process of Construct Validation." In *Handbook of Research Methods in Personality Psychology*, ed. Richard W. Robins, R. Chris Fraley, and Robert F. Krueger, 461–94. New York: Guilford.
McDermott, K. B. 2000. "Implicit Memory." In *The Encyclopedia of Psychology*, ed. A. E. Kazdin, 231–34. New York: Oxford University Press.
Messick, Samuel. 1995. "Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning." *American Psychologist* 50 (9): 741–49.
Nosek, Brian A. 2005. "Moderators of the Relationship between Implicit and Explicit Evaluation." *Journal of Experimental Psychology: General* 134 (4): 565–84.
Nosek, Brian A., and Mahzarin R. Banaji. 2009. "Implicit Attitude." In *The Oxford Companion to Consciousness*, ed. Tim Bayne, Axel Cleeremans, and Patrick Wilken, 84–85. Oxford: Oxford University Press.
Nosek, Brian A., and Anthony G. Greenwald. 2009. "(Part of) the Case for a Pragmatic Approach to Validity: Comment on De Houwer, Teige-Mocigemba, Spruyt, and Moors (2009)." *Psychological Bulletin* 135 (3): 373–76.
Poldrack, Russell A. 2006. "Can Cognitive Processes Be Inferred from Neuroimaging Data?" *Trends in Cognitive Sciences* 10 (2): 59–63.
Poldrack, Russell A., and Tal Yarkoni. 2016. "From Brain Maps to Cognitive Ontologies: Informatics and the Search for Mental Structure." *Annual Review of Psychology* 67:587–612.
Roediger, Henry L., III. 2003. "Reconsidering Implicit Memory." In *Rethinking Implicit Memory*, ed. Jeffrey S. Bowers and Chad J. Marsolek, 3–18. Oxford: Oxford University Press.
Smith, Gregory T. 2005. "On Construct Validity: Issues of Method and Measurement." *Psychological Assessment* 17 (4): 396–408.
Sullivan, Jacqueline Anne. 2016. "Construct Stabilization and the Unity of the Mind-Brain Sciences." *Philosophy of Science* 83 (5): 662–73.
Thünken, Timo, Nadine Waltschyk, Theo C. M. Bakker, and Harald Kullmann. 2009. "Olfactory Self-Recognition in a Cichlid Fish." *Animal Cognition* 12 (5): 717–24.
Warrington, Elizabeth K., and Lawrence Weiskrantz. 1970. "Amnesic Syndrome: Consolidation or Retrieval?" *Nature* 228 (5272): 628–30.

https://doi.org/10.1086/705567 Published online by Cambridge University Press