Construct Validity in Scientific Representation: A Philosophical Tour

Anders Nordahl-Hansen[1]★ and Tone Kvernbekk[2]
[1]*Østfold University College and* [2]*University of Oslo*

ABSTRACT

This article presents a philosophical tour encircling the concept construct validity. Encapsulated by two major perspectives, realism and antirealism, we visit key topics within the philosophy of educational science such as representation and reference, truth, explanation and causation. We discuss how realism and antirealism deal with unobservables through the distinction between appearance and reality. We examine the two perspectives' stance on observational terms (O-terms) and theoretical terms (T-terms), and look at the consequences implied for researchers that reside within the two perspectives. We argue that the understanding of the concept of construct validity is essential for educational researchers and researchers from any scientific discipline. Furthermore, the discussions of construct validity raised here is important beyond the research realm, such as in educational practice and in all everyday inferences people make about theoretical entities. Any researcher and practitioner is free to choose between the -isms, but must be aware that the choice has consequences.

Keywords: *Realism; antirealism; constructs; philosophy of science; educational research*

Introduction

Suppose you grade an exam and make the judgment that the student's understanding of the basic concepts of socio-cultural learning theory is confused and leaves much to be desired. Or, you test a student and make the judgment that her linguistic capacity is excellent. Just how trustworthy are your judgments? Suppose you observe a student's (bad) study habits and explain them due to low motivation; or you watch a change in student achievement over time and explain it with learning taking place. Just how credible are your explanations?

The question concerning trustworthiness and credibility arises because the judgments and explanations in question make references to unobservable entities, attributes or processes such as understanding, ability, motivation or learning. How

---

★Correspondence: Anders Nordahl-Hansen, email: anders.nordahl-hansen@hiof.no

justified are we in such judgments and explanations? In everyday life, we handle such unobservables with ease and generally with a high degree of confidence, even if we occasionally get things wrong. But in research? In research we have to pay close attention to the relation between our judgments about unobservables and the grounds on which we base these judgments – usually rooted in something observable. We should be equally attentive in educational practice, where much can be at stake for the people we make our judgments *about*. This is where construct validity enters the picture.

In this paper, we begin *in medias res* and look at construct validity, what it is and why it is important. This seemingly innocuous point of departure contains some of the biggest issues in science and philosophy: reference, truth, causation, explanation – what we think science is and what it can do for us. Starting from construct validity, we shall embark on a philosophical tour visiting these issues as they arise. Ultimately, we shall argue, both researchers and professional practitioners have some tough personal choices to make here.

But first a little more stage-setting. Samuel Messick (1995) argues that construct validity applies to "all assessments, whether based on tests, questionnaires, behavioral observations, work samples, or whatever" (p. 5). Basically, he says, validity is an evaluative judgment of the degree to which our empirical evidence supports the *appropriateness* of interpretation of test scores and their subsequent *use*. In less technical terms: we validate our constructs to justify the inferences we make about unobservable entities, processes and attributes of people. Am I right to attribute low motivation to this student, based on my observations of his behavior? What happens if my assumption is wrong, but I act toward the student as if it were true? While Messick's discussion of construct validity is highly sophisticated, he does not touch on the adjacent philosophical issues. That seems to be left to philosophers of science. Conversely, philosophers of science do not write about construct validity; that seems to be left to methodologists or measurement specialists. Nor do philosophers of education mention construct validity when they philosophize about educational research. David Bridges (2003), for example, who has written insightfully about educational research, largely focuses on epistemological and ethical assumptions. Richard Pring (2015) surveys different research methods, outlines competing philosophical positions (such as positivism, interpretivism and postmodernism) and discusses action research and practitioner research, yet nowhere does he touch on the issue of construct validity. Pring's book does include a chapter where he addresses such key issues as realism, objectivity, causation, truth, facts, explanation, etc. Each concept receives a short treatment, one by one, but are not made to fit into a bigger picture or aligned under a major perspective. This is what we aim for in this paper: to employ construct validity to illuminate adjacent philosophical issues and make an effort to discover what is at stake, and in the process demonstrate the importance of paying close attention to our own inferential practices, in research and in practice.

Unobservables

Our philosophical tour begins in the fact that construct validity crucially involves unobservables. Lurking in the background here we find the old distinction between appearance and reality. We have the appearance of things, e.g. student behavior, and we have the unobservable, inner entities, properties or mechanisms that we might presume responsibility for what we observe, e.g. motivation, understanding, personality, intention, intelligence, learning, self-formation, and so on. Construct validity plunges us straight into deep metaphysical issues concerning the reality of abstract, unobservable entities. Should we believe what educational research tells us about reality beyond the appearance of things? Is it really the case that we can use motivation to explain behavior? Should we act on our everyday inferences to such unobservables, and thus take measures to increase student motivation? Do people really have the attributes that they appear to have? This is a longstanding philosophical problem: can we in principle distinguish between properties that truly belong to the person or object and properties that they do not really possess but exist only in the mind of the observer (Ladyman, 2002)?

Terms that we use to talk about unobservables are generally called theoretical terms, or T-terms for short. All disciplines make claims about invisible, unobservable, undetectable, theoretical entities, events, states of affairs, properties, processes and qualities. This even holds for everyday and professional talk. So what are these unobservables? How can we know anything about them if we cannot see them? How can I know that a student has low motivation if my claim is justified only by my observations? In other words, what do T-terms refer to, if anything at all?

Realism and antirealism

To investigate the question about what T-terms refer to, we go to the philosophy of science. Arguments over the ontological status of T-terms have developed into two different conceptions of science, commonly known as realism and antirealism (Kvernbekk, 2002, 2005). These conceptions represent the next stop on our philosophical tour. Both realism and antirealism are subtle perspectives encompassing slightly differing versions, but for the sake of clarity of exposition we shall present a standard picture of each of them.

We begin with the antirealists. The best-known antirealists are the logical positivists (but not all antirealists are positivists). James Ladyman (2002) describes the main tenets as follows, "They are empiricist in that they regard observation (experience) as the only source of knowledge; they are anti-theoretical entities; they are anti-causation, they emphasize verification and they downplay explanation" (p. 148). Central to their views is the empiricist criterion of meaning – a criterion that has been formulated in different ways over the years, but which basically says that if a term is to be meaningful, it must be connected to observation or experience. Science would do well to rid itself of pseudo-scientific terms and theories, such as "essence", "thing

in itself" or "superego" and their related theories. According to the logical positivists, such terms are not related to anything that can be observed, and theories using these terms really do not claim anything at all. Unless a claim is empirically verifiable, it is not meaningful. The meaning of a claim is identical to its method of verification; that is, the way we show it to be true by experience. That is to say, scientific theories are to be tested *empirically,* by observation – a viewpoint that is an ingrained part of scientific lore today. At this point we make a short detour to Alfred Ayer, a British logical positivist. Retrospectively reviewing some of the main tenets of logical positivism, he explains that the assumption behind the principle of verification was that everything that could meaningfully be said could be expressed in terms of elementary statements (Ayer, 1959). Such statements were by some positivists taken to be a record of the subject's immediate experiences: "claims about the world had to be verified by *somebody's* experience" (p. 13). Ayer acknowledges that this left science with a subjectivist, even solipsist foundation, and the logical positivists spent much energy trying to remedy this state of affairs. It is worthwhile pointing this out, since positivism is often associated with numbers, quantities and statistics. Underlying, we find people's private sense experiences – and they are surely qualitative in nature.

Not only the logical positivists but all empiricists in general share the view of science that experience is needed to test, confirm or falsify theories. Therefore, they make a clear distinction between the empirical and the non-empirical, between the observational and the theoretical: the empirical has the power to confirm or disconfirm the theoretical, but not vice versa. The empirical, in other words, enjoys a higher epistemic status. We also find here a basic tenet concerning terms. Observable entities, attributes or processes are referred to by observational terms, O-terms, such as blue, lamp, or warmer than (which is a directly observable relation). Unobservables are referred to by T-terms, which must satisfy the empiricist criterion of meaning to be allowed in scientific theorizing.

While antirealists oppose the idea of moving beyond experience, realists stand ready to make a metaphysical commitment to the existence of the unobservables referred to by T-terms, insofar as they are described by correct theories (Hacking, 1983). Intelligence, motivation and abilities exist and can safely be attributed to people and used to explain their behavior. It seems to us that scientific realism is a natural pre-philosophical stance, adopted by most empirical researchers – perhaps because it is a continuation of the commonsensical ways of thinking that we are all socialized into when we grow up. If you are a realist, you interpret theories and claims literally, as telling us something about unobservables and their equally unobservable qualities. You do not think that the meaning of all claims must be tied to observation, like the antirealists do – especially if they adhere to the verification principle described above.

Realism is surely attractive to researchers, because it allows them to think that reference takes place. We are allowed to think that the unobservables exist, and that it is possible to say something about them. We can actually say something about the role of motivation in learning and refer to unobservable entities to explain observable

ones. Moreover, realists think that what we say of the world can be true or false. We shall discuss the ontology of theoretical entities in the subsequent section, but first, a few more words about realism. Despite its attraction, realism faces vast philosophical challenges, and much criticism has been levelled against it. Bas van Fraassen (e.g. 1989) is one of the most prominent and well-known antirealists; by his own description a constructive empiricist and not a logical positivist. One of his reasons for rejecting realism is that he thinks it leads to an inflationary metaphysics – realists commit themselves to the existence of theoretical entities described by scientific theories. Being agnostic about the existence of theoretical entities is a much better bet, van Fraassen claims. This sounds like good advice; we should not willy-nilly commit ourselves to the existence of any theoretical entity, regardless of our judgment of the quality of the theory – better to leave it open. Furthermore, the history of science bears out van Fraassen's point here. One of the best-known examples is phlogiston theory, which in its day was thought to be a highly successful theory in terms of power of explanation and prediction, but eventually turned out to be false. Phlogiston does not exist after all, and so reference fails – and it was clearly a mistake to commit to it. The pessimistic meta-induction is a similar argument, but it makes a stronger claim. This viewpoint is perhaps most forcefully advocated by Larry Laudan (1981). Laudan presents a long list of once successful theories and terms which modern science has found to refer to nothing at all. The history of science thus far gives rise to the inductive inference that even our best theories today will be shown to be false and be replaced by other theories. Therefore, we have no good reason to think that unobservable entities postulated by our current best theories refer to anything at all.

The nature of constructs

Let us now go back and take a closer look at T-terms. One might wonder what antirealists need T-terms or constructs *for*, if they think they have no reference or remain agnostic about it. What is the use of T-terms, if they do not refer to anything? Here is how personality researcher Walter Mischel describes the T-term "trait" – we can substitute a wide variety of educational terms for "trait" (e.g. understanding, diligence or Bildung):

> Traits are constructs that are inferred or abstracted from behavior. When the relations between the observed behavior and the attributed trait are relatively direct, the trait serves essentially as a summary term for the behaviors that have been integrated by the observer … regardless of the exact genesis of trait impressions, trait labels may serve as summaries (essentially arithmetic averages) for observed behavior. (1973, p. 262)

This is the empiricist criterion of meaning at work: T-terms have no meaning unless they are connected to experience. For empiricists of all stripes, O-terms are basic and all other terms must be built on them. Hence, as Stephen Norris (1983) points

out, traits are not attributes possessed by people. Rather, the construct is equated with a class of behavior, and thus reflects the observer's thoughts instead of the actors', Norris suggests. T-terms are nothing but abbreviations of a large class of observations, a form of derived talk. As Norris puts it, "According to logical positivism, theoretical talk in effect is merely a shorthand way of speaking and has no other significant import" (p. 56). This treatment of T-terms presupposes that T-language can be translated into O-language. From the point of view of empirical research, Jack Shonkoff and Deborah Phillips (2000, pp. 82–83) describe the problem(s) as follows:

> In measuring height (or weight or lung capacity, for example), there is little disagreement about the meaning of the construct being measured, or about the units of measurement (e.g., centimeters, grams, cubic centimeters) … Measuring growth in psychological domains (e.g. vocabulary, quantitative reasoning, verbal memory, hand-eye coordination, self-regulation) is more problematic. Disagreement is more likely to arise about the definition of the construct to be assessed. This occurs, in part, because there are often no natural units of measurement. (i.e., nothing comparable to the use of inches when measuring heights)

Disagreement on how to measure constructs is to be expected, they suggest, and it is easy to agree with them on that. But what interests us here is the issue of definition. We have here the problem of operational definitions, which precisely concerns the translation of T into O: some behavioral indicators are picked out from a universe of behaviors to represent the construct under investigation. This fits well with the empiricist criterion of meaning, and is in general the way in which empiricists define terms: by equating them with the procedures for measuring them. But this is not how realists define concepts. For realists, the meaning of the concept (term, construct) is not equal to the procedures for measuring it, but is taken to be the sense or content of the term, which for example might be determined by the place and role of the term in a theory or a domain.

It is worthwhile to ponder the translatability of the theoretical into the observational, the psychological into the physical or the phenomenological, or the educational into classes of behavior. But how should we understand it, Jaegwon Kim asks in his incisive analysis of a certain part of the history of positivism (Kim, 2003). The positivists themselves never quite figured it out – we are not just talking about semantic equivalences here. The strong version of the verification principle demands the complete translation of T into O (a perfect operational definition, T = O), something which, if it were possible, would render the T-term a complete summary of the observations. And, we might note, there would be no construct validity problem, since there would be perfect identity between the T-term and our observational basis. It is also noteworthy that the problem of translatability faces anybody who attempts to make theoretical terms accessible for observation, whether realistically or antirealistically inclined.

Realists, as we have seen, are allowed to think that T-terms have references. They can for example think that abilities are properties of people, not classes of their behavior, or that inner processes such as self-formation actually take place. As we also have seen, they face the problem of justifying that unobservable entities are real, since obviously no observational test can decide this question. Instead, realists hold the view that theoretical entities refer to unobservable entities which have (or can have) observable manifestations. But, as argued in the previous section, we should not be gullible and accept the existence of just any entity described by a T-term. We know even from daily life that many of our terms have no reference; such as "troll" and "unicorn", or idealizations such as "average person" and "rational man".

Truth

What are the consequences of these differing views on the ontology of T-terms? It is time to widen the scope to include theories or claims, and our philosophical tour moves on to the issue of truth.

The question here is how we understand the nature of theories or claims more generally: are they assertoric claims that purport to tell us what phenomena in the world, especially those involving unobservable entities, are like; or are they tools for organizing experience or making predictions? Assertoric claims are capable of being true or false; for example, the claim that feedback on assignments increases student motivation for academic work or that high motivation makes students more diligent. Realists interpret T-terms literally, and they also interpret theories literally – as asserting something about the world and therefore capable of being true or false; truth being understood as correspondence between language and the world. The correspondence theory of truth comes in slightly different versions, but basically it says that a theory is true if the state of affairs it claims to be the case actually is the case; i.e. there are so-called truth makers in the world that determine the truth value of the theory (e.g. Kirkham, 1997). Theories can be true, approximately true or downright false. This does not work unless the terms of the theories, O-terms and T-terms alike, refer to things, properties, attributes or processes in the world. And since realists interpret terms literally, they do so refer, and the theory is capable of truth or falsity. Since we do know at least some of the truths about some phenomenon, for instance the effects of feedback on motivation, we can infer that the T-terms involved successfully refer to entities or attributes in the world.

The antirealist story is different. Leaving observables aside: if the existence of theoretical entities is denied, there is nothing for a theory to be true about, in the correspondence sense of truth. A T-term is derived talk, abstracted from observations, and nothing more than an abbreviation. Theories that employ such terms do not purport to tell us what the world is like, insofar as unobservables are involved – observable phenomena can of course be correctly or incorrectly described. If theories are not

used to provide accounts of phenomena, how then are they used? Like T-terms are shorthand tools to make scientific communication more expedient, theories are also tools and serve to organize our experience and our data in packages that we find convenient. Tools are not true or false, and antirealists therefore reject the correspondence theory of truth. Instead, they make use of the coherence theory of truth or the instrumental (pragmatic) theory of truth. The coherence theory basically defines a claim to be true if it coheres within a system of other statements, and the instrumental theory basically says that a claim is true if it is useful. As Norris (1983) points out, the logical positivists frequently combined the two, in that coherent systems can be used for making predictions and are therefore useful guides to action. Tools are not true or false; they are more or less effective, more or less adequate and fit for their purpose. We should note, however, that a theory can be true, coherent and useful at the same time – these are not exclusive categories.

So what should we do at this point? We can decide to be agnostic about whether theories are true, false, approximately true, adequate or downright fictitious, as long as they enable us to predict, which on an antirealist view is the most important task of science. Theories are confirmed if predictions are realized and disconfirmed if they are not; in accordance with the epistemic status given to experience. But can we, in the long run, be neutral as to treating theories as truths about the world versus treating them as instruments?

Can we avoid taking position on the truth of a theory and deciding whether to believe it or not? We think researchers have a choice to make here, on what they think the basic job of science is.

Causation and explanation

Explanations are generally thought to be (possible) answers to why-questions, although the history of science shows much disagreement over what sort of answers should count as genuine explanations – teleological, functional, theological, astrological, causal or intentional. Whatever we think counts as legitimate explanations, they often contain unobserved processes or mechanisms. This is equally true of all scientific disciplines. We explain observable phenomena by making claims about invisible entities, events, properties, processes or states. This kind of explanation is for realists. So, if you think you can explain bad study habits by referring to low motivation, you show yourself to be a realist. If motivation does not exist and thus cannot be endowed with certain powers to make it responsible for the observed result, then it obviously cannot be used to explain the result in question, for example study habits. The explanatory power of theories resides in the fact that their theoretical terms refer to some unobservable mechanism or process, realists claim, and this stands to reason. If T-terms are abbreviations of O-terms, the theories containing them could not explain observational claims – they could only summarize them, provide shorthand, expedient ways of talking about them.

Realists tend to place great emphasis on the power of theories not only to describe phenomena, but also to explain them. Many take explanation to be the main aim of science (Ladyman, 2002; Norris, 1983). And (approximate) truth, understood as correspondence, is regarded as necessary for a theory to be explanatory. But here the realists might be over-selling their case. As Bas van Fraassen (1989) points out, false theories can actually provide good explanations. For example, Newton's theory is known to be false (on the correspondence theory of truth), but it nevertheless provides a good explanation of the tides, as well as other earthly phenomena. It is not that easy to think of an educational example of this; it could be that we in education are more geared to other uses of theory. Explanatory power, van Fraassen argues, is a pragmatic virtue of theories that comes into play in certain contexts, when and if the researcher takes an interest in it. Antirealists tend to place more emphasis on prediction. Roughly, theories are tested by deriving hypotheses about observable events and then seeing whether the predictions hold – as argued above, on the antirealist view theories are confirmed if the predictions are realized.

We treat causation here for two main reasons. First, because most explanations would seem to be causal, such that causation is involved in the explanation of an event. Second, because while causation is highly contested in education, it is also needed in a discipline which takes *change* as one of its aims, usually in the form of learning. Causation is a dynamic process that involves the generation of some effect, preferably a desired one. "Cause" is a T-term. And again, only realistically interpreted theoretical entities will be able to play the role of a cause and exert an influence, increase student motivation, reduce negativity in interactional patterns, and the like. If your theoretical entity is a summary, a categorization of observations, it does not possess the power to influence other entities and contribute to change. So, here is another choice for researchers to make: whether to think of theoretical entities as possible causes and thus endow them with certain dynamic generative powers; or whether to treat of causation in the usual antirealist manner, which harks back to David Hume and sees causation as constant conjunction. Constant conjunctions are events that occur simultaneously and the causal connection between them is attributed by the observer. In fact, such relationships are best described as absolute correlations, and without a causal relationship between the respective events you cannot use the one to influence the other.

Construct validity: a philosophical bastard?

It is time to end our short tour, and we do so by returning to our point of departure: construct validity, this time armed with a set of central philosophical concepts and perspectives.

Construct validity concerns the legitimacy of the inferences we make about unobservable entities, processes, qualities or states. As we have seen, it is guided by two different perspectives that do not sit well together, and several writers point out the

inconsistencies in the overall theoretical framework (e.g. Borsboom, Mellenbergh & van Heerden, 2004; Michell, 1997; Norris, 1983). The best place to see these inconsistencies is arguably in the works of the classics in the field, for example Lee Cronbach and Paul Meehl. In their 1955 paper, which sets the stage for much of the subsequent discussion about construct validity, they avoid speaking of anything having to do with causation, and there is no discussion of theoretical entities having to exist. Instead, they employ a hypothetico-deductive view of construct validity, a way of thinking which belongs with the antirealist view of testing, confirmation and disconfirmation. This was picked up by Messick (1975), who argued that the major requirement of construct validity is confirmed predictions about how a test's performances should relate to other performances. However, in 1971 Cronbach argues that descriptions that referring to people's internal processes require construct validation – something which implies a realist understanding of constructs. His view, that people's performance can be explained by the process or state which produces the performance, is also clearly realist (Cronbach, 1971, p. 465). On the other hand, he does not commit to the correspondence theory of truth, but explicitly refers to usefulness (p. 477). He argues that construct validation is not about showing constructs to be actual existing things, but showing that they are consistent with the evidence – an antirealist inclination (pp. 482–483). In 1977, Meehl also took a realist turn and argued that research should move beyond co-variations to look for causal relationships and mechanisms that can explain the observed correlations (Meehl, 1977, p. 37).

A final point to be noted in this section: the very definition of construct validation itself. Roughly speaking, the meaning of a term is its sense; the ideas and descriptions associated with the term. Simple O-terms can be defined by pointing to their references, such as "mug" or "book". But since the references of T-terms are invisible, unobservable or undetectable, they cannot be defined in the same way as O-terms. There are different accounts of how T-terms acquire their meaning and how the meaning changes over time (which it undoubtedly does), but these need not concern us here. What now about construct validity? We hint in the introduction that we take the meaning of construct validation to center on the justification of our inferences about unobservable entities, processes or attributes – thus revealing, we presume, realist leanings. Those with antirealist leanings would instead employ the verification criterion of meaning and define construct validation in terms of the procedures or techniques for conducting it. Thus, Messick in 1981 defined construct validation as implying a joint convergent and discriminant strategy (p. 575), which echoes Cook and Campbell's view from 1979 that evaluating construct validity proceeds on testing for convergence and for divergence (p. 61). It will be recalled that Messick in 1995 defined validity as "an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *interpretations* and *actions* based on test scores or other modes of assessment" (p. 5, italics in original). This definition betrays realist inclinations in that the meaning of construct validation clearly is no longer couched in terms of procedures for handling

it – the use of *support* indicates that what he is after, is the justification (adequacy and appropriateness) of our interpretations (inferences). His subsequent treatment of the different aspects of validity is highly sophisticated and shows the complexities of our inferential practices, especially those that concern measurement, testing and assessment. One of these aspects he terms "consequential". This aspect concerns an appraisal of actual and potential consequences of test use, especially in regard to sources of invalidity (such as e.g. bias). This clearly points toward realist inclinations – if you think a construct is a summary of observations, you surely would not worry about sources of invalidity.

Thus, we can see that realism and antirealism follow construct validation even into the concept itself.

Conclusion

This philosophical tour started in construct validity and ends in construct validity, via visits to such central topics as reference, truth, explanation and causation as these are encapsulated by two major perspectives: realism and antirealism. We have kept the tour short for reasons of clarity, in order to highlight main differences and show what is at stake. Hence, we have elected to overlook many details and nuances and possible cross-combinations.

We hope a clear picture has emerged of why construct validity is important in educational research – and not only in research, but in practice and in all everyday inferences where we make use of theoretical entities to make attributions, assessments, evaluations and explanations. Our assessments of people may matter greatly to them, and we should do our best to safeguard and justify our inferences. Interestingly, the field of construct validity is guided by both realism and antirealism. These are -isms and thus not to be understood as empirical theses, but rather as overreaching, comprehensive attitudes toward research in general, theoretical entities, the question of truth, the nature of claims and even definitional issues. They can hardly claim to be true or false in themselves, at least not according to the correspondence theory of truth. Any researcher is free to choose between them, but has to be aware that the choice has consequences.

Author biography

Anders Nordahl-Hansen, PhD, is Professor of Special Education at Østfold University College. His research involves special education, autism and other developmental disorders. He also publishes research within research methods and statistics.

Tone Kvernbekk, PhD, is Professor in educational philosophy at the Department of Education, University of Oslo. She has contributed in various publications within the field of educational philosophy on topics related to theory and praxis as well as evidence-based practice.

References

Ayer, A. J. (1959). Editor's Introduction. In A. J. Ayer (Ed.), *Logical Positivism* (pp. 3–28). New York: The Free Press.

Borsboom, D., Mellenbergh, G. J. & van Heerden, J. (2004). The Concept of Validity. *Psychological Review, 111*(4), 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

Bridges, D. (2003). *Fiction written under Oath? Essays in Philosophy and Educational Research.* Dordrecht: Kluwer Academic Publishers.

Cook, T. D. & Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis for Field Settings.* Chicago: Rand McNally.

Cronbach, L. J. (1971). Test Validation. In R. L. Thorndike (Ed.), *Educational Measurement* (pp. 443–507). Washington, D.C.: American Council on Education.

Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281–302. https://doi.org/10.1037/h0040957

Fraassen, B. C. van. (1989). *Laws and Symmetry.* Oxford: Clarendon Press.

Hacking, I. (1983). *Representing and Intervening. Introductory topics in the philosophy of natural science.* New York: Cambridge University Press.

Kim, J. (2003). Logical Positivism and the Mind-Body Problem. In P. Parrini, W.C. Salmon & M. H. Salmon (Eds.), *Logical Empiricism. Historical and Contemporary Perspectives* (pp. 263–278). Pittsburgh: University of Pittsburgh Press.

Kirkham, R. (1997). *Theories of truth.* Cambridge, MA: The MIT Press.

Kvernbekk, T. (2002). Vitenskapsteoretiske perspektiver. In T. Lund (Ed.), *Innføring i forskningsmetodologi* (pp. 19–78). Oslo: UniPub.

Kvernbekk, T. (2005). *Pedagogisk teoridannelse. Insidere, teoriformer og praksis.* Bergen: Fagbokforlaget.

Ladyman, J. (2002). *Understanding Philosophy of Science.* London: Routledge.

Laudan, L. (1981). A Confutation of Convergent Realism. *Philosophy of Science, 48*(1), 19–49. Retrieved from www.jstor.org/stable/187066

Meehl, P. E. (1977). Specific Etiology and Other Forms of Strong Influence: Some Quantitative Meanings. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy, 2*(1), 33–53. https://doi.org/10.1093/jmp/2.1.33

Messick, S. (1975). The Standard Problem: Meaning and Values in Measurement and Evaluation. *American Psychologist, 30*(10), 955–66. https://doi.org/10.1037/0003-066X.30.10.955

Messick, S. (1981). Constructs and Their Vicissitudes in Educational and Psychological Measurement. *Psychological Bulletin, 89*(3), 575–588. https://doi.org/10.1037/0033-2909.89.3.575

Messick, S. (1995). Standards of Validity and the Validity of Standards in Performance Assessment. *Educational Measurement: Issues and Practice, 14*(4), 5–8. https://doi.org/10.1111/j.1745-3992.1995.tb00881.x

Michell, J. (1997). Quantitative Science and the Definition of Measurement in Psychology. *British Journal of Psychology, 88*(3), 355–383. https://doi.org/10.1111/j.2044-8295.1997.tb02641.x

Mischel, W. (1973). Toward a Cognitive Social Learning Reconceptualization of Personality. *Psychological Review, 80*(4), 252–283. https://doi.org/10.1037/h0035002

Norris, S. P. (1983). The Inconsistencies at the Foundation of Construct Validation Theory. In E. R. House (Ed.), *Philosophy of Evaluation* (53–74). San Francisco: Jossey-Bass.

Pring, R. (2015). *Philosophy of Educational Research.* London: Bloomsbury Academic.

Shonkoff, J. P. & Phillips, D. A. (2000). *From Neurons to Neighborhoods: The science of early childhood development.* Washington, D.C.: National Academies Press.