Citation:
Stone, J., & Zumbo, B.D. (2016). Validity as a Pragmatist Project: A Global Concern with Local Application. In Vahid Aryadoust, and Janna Fox (Eds.), *Trends in Language Assessment Research and Practice* (pp. 555-573). Newcastle: Cambridge Scholars Publishing.

CHAPTER TWENTY-FIVE

VALIDITY AS A PRAGMATIST PROJECT: A GLOBAL CONCERN WITH LOCAL APPLICATION

JAKE STONE
PARAGON TESTING ENTERPRISES
AND BRUNO D. ZUMBO
UNIVERSITY OF BRITISH COLUMBIA, CANADA

Abstract

This chapter will explain the key principles of Pragmatism. As an approach to validity based on foundational statements about human understanding, Pragmatism reflects a potentially global approach to validity concerns. As an approach to validity that is deeply embedded in how test takers, test makers, and score users interact with language assessments, Pragmatism is always essentially a local endeavor. As such, Pragmatism, without self-contradiction, can helpfully inform contemporary validity practice, especially in assessment contexts where the global and the local are intertwined. This chapter will explain how the Pragmatic approach undergirds key ideas in the philosophy of science and also undergirds language test validity practices that involve the use of argumentation and explanation. Perhaps of even greater relevance are the misunderstandings that can arise pertaining to concepts such as “constructs” and "causal attributes" when the Pragmatic approach is neglected.

Keywords: inference to the best explanation; language assessment; Pragmatism; validity; validity argument

> "And the tangible fact at the root of all our thought-distinctions, however subtle, is that there is no one of them so fine as to consist in anything but a possible difference of practice." (James, 1907, p. 18).

Introduction

Does an Asian English language test need to take the same approach to validity as a European or American test? On the one hand, language itself is regionally diverse and the uses of language assessments vary depending on the (often culturally embedded) expectations of score users. On the other hand, the validity of language assessment is a global concern as test scores can be used on a global scale and even if the score use is intended for a local context, the inferences from scores can still be made to something global, i.e., English proficiency. Language assessment specialists may also need to demonstrate the validity of a local assessment to international colleagues. For this purpose, there needs to be a shared perspective as to what constitutes validity. If validity is a global concern, there needs to be a shared perspective on what constitutes a valid assessment and also on the methods through which assertions of validity can be set forth and evaluated.

In this chapter, we discuss some foundational questions that must be addressed when seeking a shared perspective on approaches to validity. How does a phenomenon such as language proficiency become an object of study? How do we make sense of and make claims about such phenomena? How do we advance our understanding and refine our claims about that which we have come to understand?

These questions may appear to be beyond the scope of a short book chapter. Our project, however, is concise. This chapter will explain the key principles of Pragmatism[1], a philosophy developed a century ago by American scholars including William James, John Dewey, and George Herbert Mead. It should be noted at the outset that Pragmatism as a philosophy remains insufficient as a guide for developing validity programs. Nonetheless, as we shall relate in the next section, the Pragmatic approach undergirds key ideas in the philosophy of science, ranging from Popper's falsification theory (2002) to Kuhn's theory of scientific paradigms (1962) to Lakatos' theory of research programs (1975). The Pragmatic approach

[1] Because Kane (2013b) describes himself as pragmatic "with a small p" (p. 120), we will refer to Pragmatism with a capital "P" in order to respect the distinction Kane has drawn between the colloquial use of the word pragmatic and adherence to the philosophy of Pragmatism.

also undergirds test validity practices that involve the use of argumentation (e.g., Kane, 2006, 2013a) and explanation (e.g., Zumbo, 2007, 2009). Perhaps of even greater relevance are the misunderstandings that can arise pertaining to concepts such as "constructs" when the Pragmatic approach is neglected.

The first half of this chapter will focus on a general explanation of Pragmatism and consider some possible objections to the Pragmatic approach. With this in place, with no loss of generality, we will consider Pragmatism's implications on the validity of language assessments.

Pragmatism

William James, a Pragmatist philosopher and a founding father of the American discipline of psychology once wrote that "the tangible fact at the root of all our thought-distinctions, however subtle, is that there is no one of them so fine as to consist in anything but a possible difference of practice." (James, 1907, p. 18).[2]

This statement encapsulates the central tenet of Pragmatism. Pragmatism holds that understanding emerges through action in the world and that all our understandings, no matter how abstract or conceptual they may seem to be, are distinctions in how we act, that is, in how we do things in particular practical contexts. Conversely, if there is no practical distinction in action arising from any two thoughts or concepts, then there is no difference between these two thoughts or concepts. They are, for all intents and purposes, one and the same.

Pragmatism is not, in itself, a theory of scientific knowledge. Pragmatism seeks to account for how people make sense of the world through action in the world. Questions such as how one might get to an appointment on time (Dewey, 2012), why the door won't open, or how I can tell a Bangkok taxi driver to take me to Nonthaburi, are as relevant to an account of how we make sense of the world as is the case of the scientific method. The scientific method is a particular way of making sense of the world through our actions in the world.

[2] This statement by James is, in fact, a paraphrase of James' colleague and (somewhat less eloquent) Pragmatist philosopher Charles Sanders Peirce.

A few brief illustrations may help to clarify this point. Consider a three year-old child who sees a friend "hopping." She too says she is "hopping," but "hops" on two legs (that is, she "jumps"). Her friend corrects her and the child eventually learns to hop. The child has learned a new word by learning a distinction in practice. Now consider a flat cement roof. After a few years, the iron rods have corroded and the strength of the roof is compromised. The owner of the roof then learns, through the practice of repairing the roof, that there is a key distinction between a sealed concrete roof and an unsealed concrete roof. In both cases, a practical problem has arisen and, through the resolution of the problem, refined understandings emerge. The scientific method is also fundamentally pragmatic in this sense. The scientist encounters an anomaly in her observations, and through the practice of science, that is, by acting on the anomaly in a principled manner, she seeks to develop a distinction in her actions that leads to a distinction in outcomes and hence a distinction between the predicted and anomalous situations. In sum, Pragmatism is not a scientific system, but rather it is an approach to developing human understanding that is founded on an account of how we, as humans, orient to the world.

Some key features of the Pragmatic approach are as follows: First, when our actions in the world proceed smoothly, we believe or simply presuppose that our practices sufficiently explain our world because they sufficiently facilitate action in our world. When we encounter an obstacle or hindrance to our intended course of action, we need to make new distinctions about what works and what does not work, that is, we make new distinctions only when we encounter new problems. As an example, that people can do certain things with varying degrees of proficiency is a phenomenon that has been understood since ancient times (Aristotle, 2009). The need to test proficiency in a standardized manner emerged, however, as a solution to a problem, namely, how people can be selected for roles in large organizational structures. The ancient examination system in China sought to solve this problem, as did the officer examinations in the 18th Century British navy. As the problems of selection in large organizations became more prevalent and more varied through the 19th Century, so the practices of assessment became more refined (Rose, 1998; Spolsky, 1990). Language testing only became prevalent, however, when the problem of language proficiency became more apparent in the second half of the 20th century as the number of non-native English speaking students attending English speaking universities increased. It was in the context of solving the problems (or perceived problems) arising from the influx of foreign students that language proficiency testing proliferated.

Second, a Pragmatic approach is an empiricist approach. Empiricism holds that we can only make sense of that which we are able to detect with our five senses (and ideas emerge in our minds as we integrate our sense data). Following a Pragmatic approach involves making sense of our observations of the world. However, the focus of Pragmatism is distinct from the empiricism of the 17th century empiricist philosopher John Locke. Locke conceptualized the world as moments of observation coupled with the cognitive faculties that make sense of observations. The emphasis of Pragmatism is not on observation alone, but rather on the actions, contexts, and circumstances in which the observations took place. Observations can only be understood in terms of the actions that led to the observations and the practical outcomes that followed from the observations. Unlike Locke, with a Pragmatist approach, mental faculties are not autonomous from practical activity. Consider, for example, a multiple choice reading test. Each item contributes a point of data for a final score, and that score is a statement about a person's ability. Such ability is often interpreted as a latent ability construed from the scores. From a more Lockean (and also cognitive) point of view, this latent ability is an attribute of the person and needs to be conceptualized and encapsulated as such. A Pragmatic approach, however, would also focus on the context of the item, the skills the test taker used to complete the item, and the practices of the assessment specialists who developed the item.

Third, Pragmatism is concerned with what works in practice. Unlike many philosophical approaches that seek to explain how ideas represent real things in the real world, Pragmatism holds that concepts, theories and ideas are potential plans of action. The connection between mind and world precedes and belies the separation of idea and actual thing (or "thing in itself" to borrow from the Kantian tradition). Distinctions we make in the world are distinctions that emerge in our encounters with the world, distinctions between "food" and "not food" or "good pronunciation" and "weak pronunciation." The formation of ideas follows action and applies to future action. We can say that we understand something or that something makes sense to the extent that we know the practical implications of engaging with that thing. Truth is the *assertion* that some action, operation, or practice will, in fact, achieve the outcome or consequence we believed would occur. Our theories and concepts are *warrants* that our assertions are correct. As the Pragmatist philosopher, John Dewey (1941), explained, the central question of any inquiry is what practical conditions, actions, or operations warrant an assertion. As we shall discuss shortly, there is a very

close relationship between Dewey's discussion of warrants and assertions and Kane's (2013a) discussion of warrants and claims.

Fourth, Pragmatism holds that theories and concepts are instruments that we use to guide our actions. A theory has no value beyond its practical implications. A concept is not an abstract entity, but rather exists as a way to discuss and guide our engagement with particular aspects of particular situations. Theories and concepts are derived from our actions and function as guides to action. They have no existence beyond their relation to action.

Objections to Pragmatism

There are some objections to Pragmatism that may spring to mind. The first and perhaps the strongest objection to Pragmatism is that the world exists irrespective of our knowledge about it. Relativity theory does not exist because of human activity but because the universe is such that relativity theory applies. Moreover, the contraction of time and space in objects travelling close to the speed of light has always been the case and is altogether unrelated to human action. Contra the Pragmatist claim that theories emerge in practical engagement with the world, one might object that there is indeed knowledge that is purely theoretical. Again, relativity theory is a case in point since Einstein's theory was hotly debated and "theoretical" discussion went on among physicists for over half a century before any empirical tests were successfully carried out.

When we turn our focus away from physical sciences and towards social sciences another objection arises. One might, in addition, object that abstract concepts make sense only because they are abstract and have an existence beyond particular instances. By way of example, language proficiency is a concept that is abstracted from any particular case of using language. It is our conception of what language proficiency is that shapes our assessment of particular instances of language use.

If we consider language proficiency to be a special kind of concept, say, a *construct* or a *causal attribute*, then language proficiency is an unobservable quality or attribute that determines any individual instance of language use (e.g., Borsboom, Mellenbergh, & van Heerden, 2004; Cronbach and Meehl, 1955). Simply put, my language proficiency determines my language use. As Cronbach and Meehl suggested, there are laws of nature that determine the occurrences, actions, and outcomes we see in the world. A construct exists as an unobservable feature of these laws of

nature (or as Cronbach and Meehl suggested, a construct exists within a "nomological net") that determines outcomes in a regular manner. Indeed, one might argue that Pragmatism has things backwards—it is not our actions in the world that determine our concepts or constructs, but rather, it is the constructs that determine the qualities and attributes of our actions. The task of a social science such as language assessment is to capture such unobservable constructs in such a way that they can be measured as accurately as possible.

Addressing Objections to Pragmatism

We will consider these possible objections sequentially. Setting aside the somewhat contrived arguments of solipsists and skeptics, there is little doubt that the universe exists, existed before the emergence of human life and, with the exception of some particulars on Planet Earth (exceptions ranging from nation states to office parties), existed independently of human action. Why, though, is there little doubt that the universe exists and that it exists in such a way that the theory of relativity applies? There is little doubt because we assert its existence and we provide warrants for that assertion. Over the last three centuries, we have developed a sense that our assertions are becoming increasingly precise and that we are developing an increasingly refined understanding of our universe. These understandings emerge through our practices of inquiry, the ongoing refinement of the distinctions we make and the application of our findings to develop new ways of doing things and new things that we can do.

It is also the case that theories can precede their empirical demonstration or perhaps exist without even the possibility of empirical demonstration. It is not the case, however, that theories exist in a manner that is autonomous from human action or that theories exist in a way that has no influence or bearing on human action. Theory generation and theory testing are human activities, activities that depend on how we use the language available to us and how we generate new, shared ways of using our language. This point was addressed by Dewey's colleague, George Herbert Mead (1932, 1934), in his account of symbolic interactionism.

In very brief, symbolic interactionism can be explained as follows: Humans are social organisms. We attune to the activity of those around us and are ready to engage with our surroundings as others engage. A special characteristic of the human form of life is that a person not only attunes to what others are doing but also attunes to the way others are attuning to her.

If we return to our earlier illustration of the child learning to hop, she attuned to what her friend was doing and was also aware that her friend was attuning to her. She was able to act as her friend acted and to use the word "hop" as her friend used the word. Her learning the word "hop" was (1) embedded in the activity of hopping and (2) involved her and her friend engaging both with the activity and the word *in the same way*. It is through such cooperative, shared activity that language emerges.

Theory building is a case (perhaps a special case) of shared, symbolic interaction with others. Theory building involves people attuning to and engaging in the same way with the symbols they have created. The extent to which we can assert that a theory is sound, valid, or correct, is the extent to which we can provide warrants for our assertions. Our warrants depend on our mutually agreed, cooperative engagement with the practices and symbols through which the theory was developed. To borrow from Quine (1951), "[t]he totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges" (p. 459).[3] Our theories certainly become a rich network of explanations about our experience and our anticipations of how activity in the world will unfold. Indeed, the way we interact with our theories takes up a life of its own, a life of human interaction in universities, research centers, test development units and so on. These theories, however, are part of a man-made fabric that emerges in human interaction and persists through our coordinated, mutually understood, shared engagement with symbolic, linguistic activity.

Hence, in response to the second objection, concepts as simple as redness or as nebulous as language proficiency certainly exist independently of particular instances of red things or proficient utterances. However, such concepts do not exist independently of our symbolically grounded interactions with the concepts. A concept can be very precise. We might agree for example that redness is identified with a very specific wavelength of light or that language proficiency is identified by a very specific assessment instrument. To muster such agreement, however, depends on the warrants that stand behind our assertions and the shared practices and methods used to develop the warrants that we use.

[3] Although Quine is not always considered to be a Pragmatist philosopher, he was influenced by Pragmatism and instrumental in reviving an interest in Pragmatism in the latter decades of the 20th Century.

It should also be noted that the notion that constructs are unobservable entities determining observable actions is not generally accepted among validity theorists (see Slaney & Racine, 2013, for discussion) nor was this characterization of constructs posited as more than a possibility by Cronbach and Meehl (1955). In fact, Cronbach and Meehl also recognized that constructs emerge in the practices of collaborative inquiry. Construct validity, they noted, depended on the degree of agreement among researchers which in turn depended on the specificity of the theory or nomological net articulated by a construct's proponents. This argument is very close to Dewey's point that our assertions depend on the warrants we are able to set forth and test in practical contexts.[4] We will return to this issue later in the chapter.

Why Pragmatism is Relevant to Validity but Insufficient for Validity

Pragmatism is relevant to developing valid language assessments because it highlights two critical aspects of the assessment process. First, an assessment is always an assessment of practical activity. This activity may be the activity involved in using a language, flying a plane or doing metaphysics, but it is always something that people do.[5] Second, developing an assessment tool is also a form of human activity in which we attempt to draw refined distinctions among the things that a person does or says in order to make some assertion about that person. In short, assessment is the activity of refining distinctions about some form of human activity.

The validity of any form of assessment is at risk when the Pragmatic nature of assessment is neglected or replaced with notions of unobserved entities *detached from human action*. An assessment is valid to the extent that

[4] This is not to say that Cronbach and Meehl were Pragmatists. Rather, the qualifications they made to their proposed characterization of "construct" opened the door for a Pragmatic interpretation of what a construct is.

[5] One might object that there are assessments of traits or personality which are not things that people do. Such constructs are, nonetheless, created in human activity and embedded in the context of human activity. Typically such assessments ask what people do or ask people to reflect on how they feel (and reflecting is doing something with words). Moreover, as discussed in the previous section, engaging with a concept such as personality is engaging with shared symbolic, linguistic activity. In addition, taking an assessment is, in itself, an organized form of human activity.

(1) it observes or measures human activity,
(2) the action of observation leads to our making reliable distinctions in that activity,
(3) distinctions we make are supported by an extensive range of warrants, and
(4) that we can extrapolate and generalize from that activity.

As Kane (2013a) has also argued, the warrants, in turn, may take the form of theories, related observations, or reasoned arguments. Such warrants are also embedded in and emerge within our practices as people and as researchers.

This point can be illustrated by returning to the earlier example of latent ability. Latent ability is construed from test performance, which, in the case of a multiple choice test, is a series of empirical points of data. Each point of data (i.e., response), however, was generated by a person's engagement in activity. To describe ability as anything other than an individual's ability to answer a particular item is to make a generalization from particular moments of practical activity. We may reasonably claim that such a generalization is informative. However, such claims can only make sense if warranted by theoretical and empirical claims about what the test taker was doing when engaged in the practical task of answering each test item. We are making probabilistic claims about what a person may or may not be able to do in any given moment of communicative interaction. However, if we interpret a latent ability as an encapsulated and autonomous entity or attribute of a person and, at the same time, neglect the activity through which we conceptualized that attribute or entity, we are at risk of reifying such attributes, that is, treating them as having a concrete reality within a person, or assuming a person possesses a single thing called language ability. This is a very different claim to one where we make probabilistic predictions about what a person may or may not be able to do.

It is important to note that Pragmatism undergirds a century of developments in the philosophy of science. The philosophical works of Popper (2002), Kuhn (1962), and Lakatos (1975), for example, are all premised upon an understanding of science as a collective activity of human inquiry. Despite sharing this premise, however, these three approaches to the philosophy of science arrived at distinct conclusions as to how we derive valid, shared understandings of scientific knowledge. Pragmatism, then, remains insufficient if utilized as the sole guide to validating the outcomes of our inquiries or assessments. Pragmatism provides an essential premise for our work in language assessment, but does not offer a clear formula or

methodology for validation. As we shall discuss, Pragmatism nonetheless forms the foundation for contemporary movements in language assessment validity and warns us against false assumptions and unwarranted assertions.

In the remaining sections of this chapter we will consider how Pragmatism undergirds some contemporary approaches to test validation and how Pragmatism can inform our practices of validation in language assessment. We will consider the relation of Pragmatism to Kane's (2006, 2013) approach to validity as a form of argument and Zumbo's (2007, 2009) approach to validity as a form of explanation. Finally, we will return to our use of constructs.

Validation through Argument and Explanation.

In recent years, developing a validity argument has become a mainstream approach to validating language assessments (see, for example, Chapelle, Enright, & Jamieson, 2008). Following Kane (2013), validating an assessment involves two steps. First, we need to develop an argument as to how the test scores can be used. We develop very explicit claims as to what kinds of inferences are being made and the justifications that support these inferences. This is known as the Inference Use Argument (IUA). Second, we need to consider the validity of the proposed uses for the assessment and the inferences that support these proposed uses (see Kane, 2013, for a comprehensive discussion). As Kane notes, the distinction between determining the inferences and uses of a test on the one hand and determining its validity on the other is conceptual. It is also, we would note, methodological, encouraging the developers and users of an assessment to carefully review the assumptions and inferences made during the development of the assessment.

The IUA and validity argument share key features. In very brief, an argument commences with *data* and moves to a *claim*. By way of illustration, that Abdurahman scored 27/38 on Form 5 of a reading test is datum. That Abdurahman has intermediate level reading proficiency is a claim. To make our way from datum to claim, we need to develop coherent, plausible, and clear *warrants*. In this case, our warrants will include evidence for standard setting, equating of test forms, and domains within which the claim of reading proficiency applies. Each of these warrants involves multiple supporting claims. The warrant for standard setting, for example, involves claims that the standard setting panel was qualified, the standard setting design was relevant and the standard setting practices were

carried through with appropriate accuracy and rigor. Hence, each warrant will depend on some form of *backing* that provides support for the warrant.

Prima facie, this method of inquiry is analogous to that set forth by Dewey (2012). Indeed, all we need do is substitute the term *claim* for *assertion* and we seem to have Dewey's Pragmatist model for enquiry set forth. Moreover, as with Dewey, Kane (2013) avoids reference to truth on the one hand and our representations of the truth (i.e., knowledge) on the other. Rather, Kane focuses on the reasonableness of the claims that we make and the efficacy of the actions (i.e., uses of test scores) that follow from the claims. However, we need to do our due diligence before placing Kane's approach to validity in the Pragmatist camp. The central influence on Kane's approach to validity is not Dewey, but Toulmin (2003). Indeed, Kane's work can be readily interpreted as the application of Toulmin's theory of argumentation to the discipline of testing and assessment.

Toulmin is generally recognized as an ordinary language philosopher who was influenced by the philosophers Ludwig Wittgenstein and Gilbert Ryle (Hitchcock, 2010). The later philosophy of Wittgenstein shares several features with Pragmatism (Rorty, 1982). [6] For Wittgenstein and his followers, we make sense of our world through our actions. Language is something that we as people do. Language is always involved in a form of life and only makes sense as part of that form of life. When I say, for example, that Abdurahman has intermediate reading proficiency, this only makes sense in a human form of life in which people read; their reading ability matters in varying practical contexts; and we engage in the practice of assessing people's reading.

Of greater relevance, Toulmin himself was familiar with the work of Dewey. It seems that his only objection to Pragmatism was that it did not go far enough. Following Wittgenstein (1953), Toulmin had a keen interest in refuting analytical philosophy, a philosophy of logical propositions that can demonstrate necessary truths when the premises are clearly specified. In the concluding chapter of *The Uses of Argument* (the central influence on Kane's work), Toulmin (2003) argues that the Pragmatic approach is the most fruitful approach to resolving problems. His objection that Pragmatism

[6] Wittgenstein's work is typically divided into an earlier and later period, with each period being characterized by a distinct philosophical approach.

fails to refute analytical philosophy is more of a concern in the work of philosophers than it is in the work of assessment professionals.[7]

To be clear, no assertion is being made herein that Toulmin and Kane were primarily influenced by the Pragmatist philosophers nor that developing an IUA and validity argument is a methodology set forth by Dewey. Rather, the assertion here is that Toulmin, Kane, and those who work with IUA and validity arguments are engaged in a validation practice that coheres with the principles of Pragmatism and makes assertions in a manner that coheres with a Pragmatic approach to human inquiry.

Similarly, approaches to validity based on inferring to the best explanation (e.g., Zumbo, 2007, 2009) also utilize a Pragmatic approach to enquiry. Simply put, a test score is valid if we can (a) identify the question that the test score is intended to answer, and (b) identify the best possible explanation for the score in response to the question. In Abdurahman's case, what question was being asked when he took the test and received a score? What is the best explanation for that score and how far reaching is that explanation?

Inference to the best explanation is a Pragmatic approach to enquiry. The approach starts with a problem that needs resolving and concludes with an explanation that is testable and applicable in particular contexts. Moreover, the process of developing an explanation is very much a human practice involving the development and refinement of tools and methods that help us to clarify our theories and martial our arguments. Indeed, perhaps the key distinction between an argumentation approach to validation and an explanatory approach to validation is that the explanatory approach is premised on the development of validity arguments but then switches the focus to how we then decide which is the best argument or the best explanation.

The Pragmatic Approach in Language Assessment

Merleau-Ponty once wrote: "Since explanation is not discovered but created, it is never given with the fact, but is always simply a probable

[7] The objection is also problematic. While the key features of Pragmatism presented herein are shared by the best known Pragmatist Philosophers, Dewey, James, Mead and Pearce, there are also differences among their philosophical approaches. In G. H. Mead's later work (1932) symbolic interactionism had evolved into a philosophical approach that was incompatible with analytic philosophy.

interpretation" (2003, p. 133). If we agree with Merleau-Ponty, and if we accept that the Pragmatic approach to inquiry that undergirds validation by both argument and explanation remains our best possible approach, then it seems that any use of any score in a language assessment is never based on anything other than a probable interpretation. This may seem disconcerting as we surely owe our test takers certainty. It is unsettling to think that each week people receive letters that tell them they have been denied work in Canada or denied a place in a university based on a probable interpretation of their test scores! And yet, following both Pragmatism and subsequent approaches to the philosophy of science, this is what we do.

What, then, does Pragmatism tell us about how we should provide the best and most probable explanations? Although Dewey (1941, 1938, 2012) had much to say about inquiry and logical inference, returning to Dewey to seek methodological guidance would be of little use. As Dewey himself pointed out, our inquiries are embedded within particular contexts and involve particular problems. The methods we use to validate our language assessments will depend on these contexts. It is for this reason, that we turn to contemporary theorists (e.g., Borsboom et al., 2004; Kane, 2006, 2013; Zumbo, 2007, 2009) to guide our validation efforts, follow the examples of other language assessment professionals when developing methods of validation (e.g., Chappelle et al., 2008), and, most importantly, critically examine our own assumptions and ensure that each assumption has been tested as a practical, live, and probable possibility.

Crucially, being Pragmatic in our validation practices does not mean muddling through our validation efforts in an ad hoc manner nor does it mean that we pragmatically (with a small "p") do whatever seems to work. On the contrary, Pragmatism shares the foundational practices of science, namely careful observation, testing, and refinement of methods.

Moreover, Pragmatism is a "mythbuster." Following Pragmatism (and contra the assumptions of much Western Enlightenment philosophy), there is no gulf between the world as it really is on the one hand and our knowledge of the world on the other. There is no gap between a person's true language proficiency and our measure of that person's language proficiency because there is no true language proficiency that exists autonomously from how we engage in the practices of language assessment. Pragmatism, as Toulmin (2003) pointed out, is one small step from skepticism, but it is a crucial step. Skepticism holds that we can never know about the world beyond our sense data (Hume, 2003). Pragmatism holds

that traditional ideas of knowledge are misplaced. We are not pursuing knowledge about a world "out there," but rather we are refining our understandings of how we can interact with the world. This is why understanding the Pragmatist roots of contemporary validation theories and practices is important. Pragmatism provides caveats that must apply when we choose what we are trying to validate and what claims we can reasonably make.

Does it follow that adopting a Pragmatic approach to language assessment validation precludes the possibility of construct validation or the validation of causal attributes? As noted earlier, constructs have been understood by some as entities or processes that determine outcomes. Borsboom et al. (2004) argued that demonstrating the existence of causal attributes is the only form of validity.[8] Language proficiency, for example, can be considered as a construct (or perhaps as a causal attribute) determining performance in a language assessment. As Pragmatists, we are refining our understandings of how we can interact with the world. We are looking at what people do and what they are likely to do. How, then, should we understand constructs or causal attributes? Can we say that constructs are real or that causal attributes exist? Following the philosopher Hacking (1998), such a question is misplaced. We can never answer an interrogative about whether something is real or whether it exists. Rather we need to ask; a construct is a real *what*? *How* does this attribute exist?

Maybe, as some hold (e.g., Borsboom et al., 2004), there are real, unobservable attributes that determine the performance, attributes that we are able to observe and directly measure, a performance such as responses in a language assessment. Maybe such causal attributes are embedded in a nomological net (Cronbach & Meehl, 1955); neither Kane nor Zumbo preclude this possibility. If, however, we are to hold that there are real, unobservable, determining attributes or processes, then we have two possible choices (or "living options" to adopt a Pragmatist term). These options were set forth by William James (1912). On the one hand, we can hold that there is no more to the world than we can possibly discover through our practical interactions. Hence, the best possible explanation is the explanation that utilizes the best possible practices of practical inquiry. On the other, we can hold that our practical interactions with the world will

[8] Borsboom et al. (2004) distinguished such causal attributes from "constructs" as they argued that there is no need for a nomological net. Either the attribute has a measurable effect or it does not. An attribute's interaction with a nomological net, they argue, is not relevant to its measurement.

never bring about an understanding of how the world works. We can believe that there is more to the world than we can possibly find out through our best scientific practices - James proffered this second option as a plea for faith in God. The second option is perhaps less of an option for claiming unobservable causal attributes or constructs. Hence, the validity of a construct or causal attribute depends on James' first live option, the Pragmatist option.

Validating an assessment by utilizing constructs or causal attributes as the explanandum for a test score is fundamentally a Pragmatic endeavor, depending on data, warrants, backing, and finally assertions that are testable and consistently useful. On the one hand, Borsboom et al.'s (2004) argument for causal attributes depends on their specification through the practices of measurement. On the other hand, as Cronbach and Meehl (1955), Kane, (2013), and Zumbo (2007) observe, construct validity depends on the development of an extensive, well supported argument. Even then, construct validity may not be the best possible explanation for a test score. In language assessment, for example, time spent studying a language, how a person uses a language on a day-to-day basis, whether a person uses that language at work and other such factors may offer alternative, competing explanations. In short, as both Kane and Zumbo have recognized, construct validity can play a role in developing the validity argument for a language assessment, but it may not be the only role.

Conclusion

To conclude, when we assess language and when we seek to justify the product of our assessment (typically a numeric score), we are engaging in a Pragmatic endeavor. Pragmatism does not dictate the specific methods we use or theories we develop, but simply explains how we are able to make sense of our theories and understand our professional practices and then put them to good use with a clear anticipation of the consequences of our actions. As Zumbo (2009) points out:

> It is rare that anyone measures for the sheer delight one experiences from the act itself. Instead, all measurement is, in essence, something you do so that you can use the outcomes… (Zumbo, 2009, p. 66)

As Zumbo highlights, measurement is always for a purpose. In language assessment, the purpose of measurement is for the perceived good of the person being assessed and/or for the perceived good of the institution using the assessment score. Pragmatism is an orienting point of view that

undergirds the work of language assessment professionals and brings a central focus to the outcomes and consequences of measurement. Indeed, following a Pragmatist point of view, the efficacy and accuracy of a measurement is substantively determined by its practical outcomes.

References

Aristotle. (2009). *Nicomachean ethics*. Oxford, UK: Oxford University Press.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological review*, *111*, 1061.

Chapelle, C., Enright, M., & Jamieson, J. (Eds.). (2008). *Building a validity argument for the Test of English as a Foreign Language*. New York, NY: Routledge.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

Dewey, J. (1938). *Logic: The theory of inquiry*. New York: Henry Holt and Company.

—. (1941). Propositions, warranted assertibility, and truth. *The Journal of Philosophy*, *38*(7), 169–186.

—. (2012). *How we think*. Mineola, NY, Courier Dover Publications.

Hacking, I. (1998). *Rewriting the soul: Multiple personality and the sciences of memory*. Princeton, NJ: Princeton University Press.

Hitchcock, D. (2010). Obituary: Stephen Edelston Toulmin. *Argumentation*, *24*, 399–401.

Hume, D. (2003). *A treatise of human nature*. Mineola, NY, Courier Dover Publications.

James, W. (1907). *Pragmatism: A new name for some old ways of thinking*. New York: Longman Green and Co.

—. (1912). *The will to believe: And other essays in popular philosophy*. New York: Longmans, Green, and Company.

Kane, M. (2006). Validation. In R. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Washington, DC: American Council on Education and National Council on Measurement in Education.

Kane, M. T. (2013a). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*, 1–73.

—. (2013b). Validation as a Pragmatic, Scientific Activity. *Journal of Educational Measurement*, *50*, 115–122.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lakatos, I. (1975). Falsification and the methodology of scientific research programmes. In S. G. Harding (Ed.), *Can theories be refuted?* (pp. 205–259). Dordrecht, Netherlands: Springer.

Mead, G. H. (1932). Scientific objects and experience. In A. Murphy (Ed.), *The philosophy of the present* (pp 140–160). LaSalle, Illinois: Open Court.

—. (1934). *Mind, self and society: From the standpoint of a social behaviorist*. Chicago: University of Chicago Press.

Merleau-Ponty, M. (2002). *Phenomenology of perception* (C. Smith Trans.). New York: Routledge.

Popper, K. R. (2002). *The logic of scientific discovery*. New York: Routledge.

Quine, W. V. O. (1951). Two dogmas of empiricism, *Philosophical Review, 60*, 20–43.

Rorty, R. (1982). *The consequences of pragmatism: Essays, 1972–1980*. University of Minnesota Press.

Rose, N. (1998). *Inventing ourselves: Psychology, power, and personhood*. Cambridge, UK: Cambridge University Press.

Slaney, K. L., & Racine, T. P. (2013). What's in a name? Psychology's ever evasive construct. *New Ideas in Psychology, 31*, 4–12.

Spolsky, B. (1990). The prehistory of TOEFL. *Language Testing, 7*, 98–118.

Toulmin, S. E. (2003). *The uses of argument*. Cambridge, UK: Cambridge University Press.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe Trans.). Oxford, UK: Blackwell.

Zumbo, B. D. (2007). Validity: Foundational Issues and Statistical Methodology. In C.R. Rao and S. Sinharay (Eds.) *Handbook of Statistics, Vol. 26: Psychometrics* (pp. 45–79). Amsterdam, Netherlands: Elsevier Science.

—. (2009). Validity as contextualized and pragmatic explanation, and its implications for validation practice, in R. Lissitz (ed.), *The concept of validity: Revisions, new directions and applications* (pp. 65–82), Charlotte, NC: Information Age Publishing Inc.