

Clearing a path toward evaluation autonomy: Minimizing measurement disjuncture and  
improving the validity of findings through culturally specific assessment

David A. Sul

Sul & Associates International

November 15, 2019

Paper presented at the American Evaluation Association Conference, Minneapolis, MN.

### **Abstract**

Educational research findings often rely on evidence obtained through the use of standardized assessment instruments. This presentation draws a distinction between Western and Indigenous worldviews in noting that all assessment instruments are constructed within a particular worldview. Measurement disjuncture (Sul, 2019) refers to the misalignment that occurs when elements of an instrument development process from one worldview are applied to the instrument development process of another worldview. In manners that are both qualitative and quantitative, measurement disjuncture negatively affects the establishment of measurement validity. Further, as a result of measurement disjuncture, researchers are less likely to acknowledge that educational programming has had an impact when, in fact, it may have (Type II error). This paper focuses on the development of culturally specific assessments, the Papakū Makawalu Competency Assessment and the Dibishgaademgak Anishinaabemowin (language) assessment, as a means to minimize measurement disjuncture. The paper concludes with a declaration of educational and assessment autonomies.

*Keywords:* Measurement validity, measurement disjuncture, culturally specific assessment, culturally responsive assessment, language revitalization, indigenous knowledge

### Background

Measurement has been a practice of Indigenous<sup>1</sup> peoples for millenia. Whether constructing baskets or charting paths to navigate the natural environment, Indigenous people have used practical tools to conduct measurements, often using various parts of their bodies when appropriate for the situation as described by Emma Meawasige (Anishinaabe) from the Serpent River First Nation:

*“Well first of all our people didn’t have rulers. They didn’t have rulers to measure with. They measured by parts of the body... and everything is very visionary. You use your thumb, you used your fingers, your hand... like your hand full, fistful, arm length. Your arm length was here, (top of shoulder to tip of finger) the foot is used by your foot and it was always approximate. From what I saw, the way they worked from measurements, they always used hand-spans (from tip of thumb to middle finger with hand stretched out).”* (Day-Murdoch, 2018, p. 2)

In addition, for generations, Indigenous peoples have utilized performance-based assessment practices to determine how individuals could best contribute to the society (Bordeaux, 1995). Adults observed children exhibiting varying degrees of skill in tasks such as “hunting, running, consensus building, healing, and spiritual leadership” (Bordeaux, 1995, p. 3) and those who demonstrated superior performance were the ones who later led hunting parties,

---

<sup>1</sup> The term “Indigenous” is used throughout this paper to refer to the people and peoples who identify their ancestry with the original inhabitants of North America. It is also used as a modifier to describe such aspects of the study such as “Indigenous language.”

provided spiritual guidance, served as orators for the people, and performed other necessary tasks for the group. To this day, observation, assessment, and feedback practices remain present within Indigenous communities and are used by parents, elders, teachers, master craftspeople, and ceremonial leaders. These practices can play a critical role in cultural and linguistic revitalization efforts.

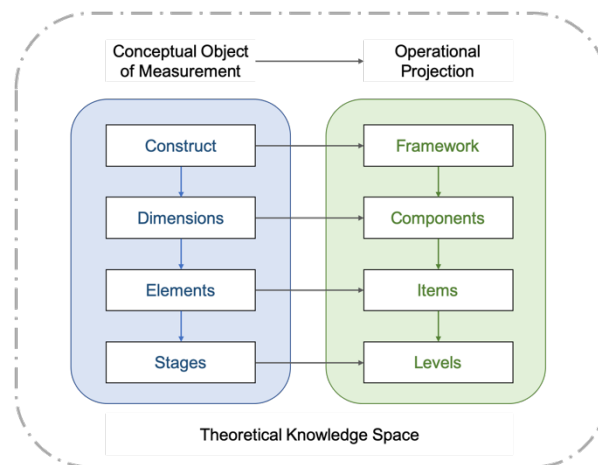
### **Problem Statement**

As with any other product of human activity, tests are cultural artifacts (Solano-Flores, 2011, p. 3) existing within a given worldview. As such, elements of the instrument-development process are prescribed necessarily by the cultural worldview under which they are presented. The cultural validity of tests is the degree to which they address sociocultural influences such as values, beliefs, experiences, and epistemologies inherent within cultures as well as the socioeconomic conditions under which cultural groups exist (Solano-Flores & Nelson-Barber, 2001). In the construction of assessments, it is important to maintain alignment amongst all of these elements.

### **Forms of assessment alignment**

Educational assessments function within a theoretical knowledge space that is comprised of both conceptual and operational components. The operational aspects of the assessment – the framework, components, items, and item levels – are constructed as projections of their counterparts in the conceptual space - the construct, dimensions, elements, and stages of learning. Educational instrument developers seek a formal structure that maintains alignment both within and between the conceptual and their operational projections present within the

theoretical knowledge space. The conceptual and operational aspects of an assessment and their position within each half of the knowledge space are presented in the figure below.



*Figure 1.* Alignment within and between the conceptual and operational elements of an instrument development process.

Educational instrument developers also seek a formal structure that maintains alignment amongst the constructs of knowledge, learning expectations, the educational framework, adopted curriculum, methods of instruction, and forms of assessment. There can be multiple educational settings, and, thus, an issue of validity arises when assessment instruments are developed within one worldview and applied inside of another. This issue addressed by this research occurs when assessment instruments are developed within Western worldviews and applied within an Indigenous worldview. This is summarized in the figure below.

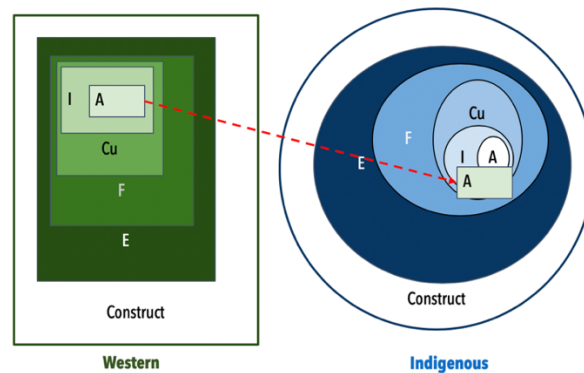


Figure 2. Assessment applied across Western and Indigenous worldviews.

An instrument-development process is comprised of an array of conceptual and operational components and is advanced through thoughtful consideration and decision-making. In the case where assessment instruments are developed within Western worldviews and applied within an Indigenous world view, since each element, consideration, and decision is influenced by the worldview in which it exists, a multitude of opportunities exist for misalignment between the two worldviews. This led to the posing of three fundamental questions about this form of misalignment that exists within instrument-development processes: *What do we call this? Why is this a problem? What do we do about it?*

### **What do we call this?**

Measurement validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). Key elements of this definition are addressed by the terms “evidence,” “theory,” “interpretations,” “scores,” “uses,” and “tests.” The meaning of these terms within the

very definition of measurement validity is grounded in and influenced by the worldview under which the instrument development occurs.

Misalignment that is grounded in cultural and linguistic differences has been referred to as “disjuncture” (Appadurai, 1996; Meek, 2010; Wyman et al., 2010) or “discontinuity” (Bougie, Wright, & Taylor, 2003; Brown-Jeffy & Cooper, 2011; Edwards, 2006; Meek, 2007). Cultural discontinuity in school settings has been defined conceptually as “a school-based behavioral process where the cultural value-based learning preferences and practices of many ethnic minority students—those typically originating from home or parental socialization activities—are discontinued at school” (Tyler et al., 2008). The cultural-discontinuity hypothesis posits that culturally-based differences in the communication styles of minority students’ home and the Anglo culture of the school lead to conflicts, misunderstandings, and, ultimately, failure for those students (Ledlow, 1992). Cultural discontinuity arises for students when their personal values clash with the ideals that shape their school system (Wiesner, 2006). Ladson-Billings (1995) described the “discontinuity” problem as the gap between what students experience at home and what they experience at school with respect to their interactions of speech and language with teachers.

*Measurement disjuncture* is the misalignment that occurs when elements of an instrument-development process from one worldview are applied to the instrument-development process of another worldview (Sul, 2019). Although measurement disjunctures can occur across multiple worldviews, this research will center on the measurement disjuncture that exists across Western and Indigenous worldviews.

**Why is this a problem?**

When assessment instruments are developed within a Western worldview and are applied within an Indigenous setting, measurement disjuncture results. Measurement disjuncture affects the establishment of measurement validity, and, hence, the inferences made based on the scores derived from such assessments. This is primarily due to the introduction of measurement error caused by the misalignment.

In many educational settings, assessment instruments are often used to determine the proficiency of a learner. A Type I error is made when an individual who is not actually proficient is deemed to be proficient. In this instance, the assessment has overestimated the ability of the learner. A Type II error is made when an individual who is actually proficient is deemed not to be proficient. In this case, the assessment has underestimated the ability of the learner. Whether it has been through the use of voting literacy tests or through English-only state standardized tests, measurement disjuncture has contributed to the underestimation of the abilities of examinees.

These same type errors also negatively affect the conclusions drawn from quantitative research designs when attempting to determine the impact of educational programs. The figure below presents a typical analysis of variance (ANOVA) table used to interpret the differences between groups in a controlled quantitative study. Three elements of the table that are influenced by measurement disjuncture are noted. In reference to the figure below, when the error term (a) increases, the mean square error (b) increases causing the value of the  $F$  statistic (c) to decrease. With a smaller than expected  $F$  statistic, researchers are less likely to acknowledge that the treatment has had an effect when, in fact, it has, which represents a Type II error.



SOURCE	df	SS	MS	F
Factor A (between)	k-1	$\sum_{i=1}^n n_i (\bar{x}_i - \bar{x})^2$	$MSA = \frac{SSA}{df_A} = \frac{SSA}{k-1}$	$F = \frac{MSA}{MSE}$
<sup>a</sup> Error (within)	n-k	$\sum_{i=1}^n (n_i - 1) s_i^2$	<sup>b</sup> $MSE = \frac{SSE}{df_e} = \frac{SSE}{n-k}$	
Total	n-1			

Figure 3. The impact of measurement disjuncture on the interpretation of quantitative research design conclusions.

When measurement disjuncture exists within assessment instruments used by educational researchers and is unaccounted for within the research design, the influence of interventions may end up being undervalued. In practical terms, researchers evaluating programs to improve educational outcomes for Indigenous people through the application of assessment instruments developed within a Western worldview may end up undervaluing the influence of such programs.

### What do we do about it?

Although the term *measurement disjuncture* is presented here, attempts to both describe and address the disjuncture within broader educational environments are not new. Au and Jordan (1981) described as “culturally appropriate” the incorporation of “talk story” into a program of reading instruction for Native Hawaiian students that improved upon expected scores on standardized reading tests. Mohatt, Erickson, Trueba, and Guthrie (1981) used the term “culturally congruent” to describe teachers’ use of interaction patterns that simulated Native American students’ home cultural patterns to produce improved academic performance. Jordan (1985) defined educational practices as “culturally compatible” when the culture of students is used as a guide in choosing aspects of the educational program to maximize academically desired behaviors and minimize undesired behaviors. Researchers beginning in the 1980s used

the term “culturally responsive education” to describe the language interactions of teachers with linguistically diverse and Native American students (Cazden & Leggett, 1981; Erickson & Mohatt, 1982). Erickson and Mohatt (1982) suggested their notion of culturally responsive teaching could be seen as a beginning step for bridging the gap between home and school. Ladson-Billings (1995) claimed the term culturally responsive represented a more expansive, dynamic, and synergistic relationship between the culture of the school and that of the home and greater community.

Attempts to address the problem of measurement disjuncture from a practical standpoint have included the use of cultural decentering (Werner & Campbell, 1970), a technique whereby English language survey items are translated from English to Spanish and then translated back into English. When confronted with a set of learning objectives from different perspectives, the technique of “cross-walking” or mapping of one set of learning objectives from one worldview to another has been applied. Quantitative researchers may opt to statistically correlate the results of one assessment to those of another. Some assessment development processes involve the use of post-hoc linguistic or cultural (or both) review panels. Other assessment researchers have adopted culturally relevant or culturally responsive practices. Here, the adoption of culturally specific assessments is offered as an additional approach to the minimization of measurement disjuncture.

### **Towards culturally specific assessment**

Ladson-Billings (1995) conducted a significant qualitative study on the teaching methods of teachers who demonstrated consistent academic success with African American students. Her work launched the movement towards the acknowledgement and identification of a “culturally

relevant pedagogy.” Ladson-Billings (1995), grounded in Black feminist thought, introduced the theory of “culturally relevant pedagogy” to emphasize the significance of teaching to and through the cultural strengths of ethnically diverse students. Ladson-Billings (1995) and Jordan (1985) argued for the use of culturally relevant pedagogy to engage actively and motivate students from ethnically diverse backgrounds to improve their academic achievement. Ladson-Billings (1995) established three criteria for a culturally relevant pedagogy that could be used to address the “discontinuity” problem: (a) an ability to develop students academically; (b) a willingness to nurture and support cultural competence to help students to maintain their cultural integrity while succeeding academically; and (c) the development of a sociopolitical or critical consciousness. In a culturally relevant classroom, a child’s culture is not only acknowledged but seen as a source of strength that can be utilized to attain academic success. Sociopolitical consciousness has been described as an individual’s ability to critically analyze the political, economic, and social forces shaping society and one’s status in it (Seider et al., 2018). For the last definitional criterion, Ladson-Billings (1995) borrowed from Freire and acknowledged that students must develop a broader sociopolitical consciousness and the skills to critique the cultural norms, values, mores, and institutions that produce and maintain social inequities (Freire, 1970). The development of a sociopolitical or critical consciousness within students allows them to acknowledge and act on historical circumstances that affect their current reality (Freire, 1970; Ladson-Billings, 1995).

Researchers in the field of program evaluation began to utilize the term “responsive evaluation” in the early 1970s in reference to a focus on issues of practical importance to program managers and developers (Stake, 2011). Stake (1973) sought to remove the emphasis on static program objectives developed by those furthest from the delivery of program services and

stressed the importance of being responsive to situational realities in the management of programs and to the reactions, concerns, and issues of participants. This represented a dramatic departure from the emphasis on the use of evaluation plans that relied on preconceived notions of program expectations. Stake (1973) believed that the ultimate test of the validity of an evaluation is the extent to which it increases the audience's understanding of the program. Stake's (1973) work led to the stream of responsive evaluation research and practices that exist today.

Drawing upon the lineage of research in responsive evaluation and culturally relevant pedagogy, Hood (1998) argued that student learning is more effectively assessed through the use of assessment approaches that are culturally responsive. Combining the ideas of Ladson-Billings (1995) and Stake (1973), Hood (1998) promoted the development of such performance-based assessments as a means of achieving equity for students of color. Hood (1998) noted that there were to be challenges and difficulties in the development of both performance tasks and scoring criteria that would be "responsive to cultural differences and adequately assess the content-related skills that are the focus of the assessment." Culturally specific assessment (Sul, 2019) represents an extension of Hood's (1998) culturally responsive assessment onto a named worldview through the addition of an additional criterion: the assessment development process functions within a system of knowledge that exists within a named worldview. Thus, the formal definition of culturally specific assessment that will be utilized throughout this document is (a) assessment that supports the academic development of students; (b) is inclusive of a willingness to nurture and support cultural competence; (c) aims to support the development of a sociopolitical or critical consciousness within students; (d) is focused on constructs and measures of importance to educational practitioners and other key stakeholders; and (e) functions within a

system of knowledge that exists within a named worldview. The table below summarizes the progression of the terminologies towards this definition of culturally specific assessment.

*Table 1.* Progression of terminologies towards culturally specific assessment.

<b><u>Culturally</u></b>	<b><u>Description</u></b>	<b><u>Proponents</u></b>
Responsive evaluation	Evaluation focuses is on issues of practical importance to program managers and developers	(Stake, 1973)
Culturally appropriate instruction	Early attempt to describe efforts to address the discontinuity problem	(Au & Jordan, 1981)
Culturally congruent instruction	Teachers' use of interaction patterns that simulated the Native American students' home cultural patterns	(Mohatt et al., 1981)
Culturally responsive education	Involves language interactions of teachers with linguistically diverse students	(Cazden & Leggett, 1981; Jordan, 1985; Mohatt et al., 1981)
Culturally compatible instruction	Culture of students is used as a guide in choosing aspects of the educational program	(Vogt, Jordan, & Tharp, 1987)
Culturally relevant pedagogy	(1) an ability to develop students academically; (2) a willingness to nurture and support cultural competence; and (3) the development of a sociopolitical or critical consciousness.	(Ladson-Billings, 1995)
Culturally responsive assessment	Supports the academic development of students; inclusive of a willingness to nurture and support cultural competence; aims to support the development of a sociopolitical or critical consciousness; Focused on constructs and measures of importance to educational practitioners and other key stakeholders	(Hood, 1998)
Culturally specific assessment	Supports the academic development of students; Inclusive of a willingness to nurture and support cultural competence; Aims to support the development of a sociopolitical or critical consciousness; Focused on constructs and measures of importance to educational practitioners and other key stakeholders; Functions within a system of knowledge that exists within a named worldview	(Sul, 2019)

Throughout the transition of terminologies from culturally appropriate instruction to culturally responsive assessment, researchers have focused their attention on the improvement of academic performance of learners within educational settings or environments that are grounded in the worldview of the dominant culture. The transition toward culturally specific assessment described here represents an attempt to do the same within the worldview of cultures functioning within a named worldview. This represents the distinction between culturally specific assessment and culturally responsive assessment.

### **Culturally specific assessment development**

Culturally specific assessment is suggested as a potential solution to the problem of measurement disjuncture. In order to determine whether the minimization of measurement disjuncture can be achieved through the employment of culturally specific assessments, educational environments that meet the criteria for culturally specific assessment are sought. Such environments do exist. Indigenous knowledge, inclusive of language, cultural knowledge, and wisdom, are being promoted throughout Aotearoa (New Zealand), Hawai‘i, tribal communities within North America, and First Nations communities within Canada. Two culturally specific assessment development cases, the Papakū Makawalu Competency Assessment (PMCA) and the Dibishgaademgak Anishinaabemowin (Measuring Anishinaabemowin), are described below.

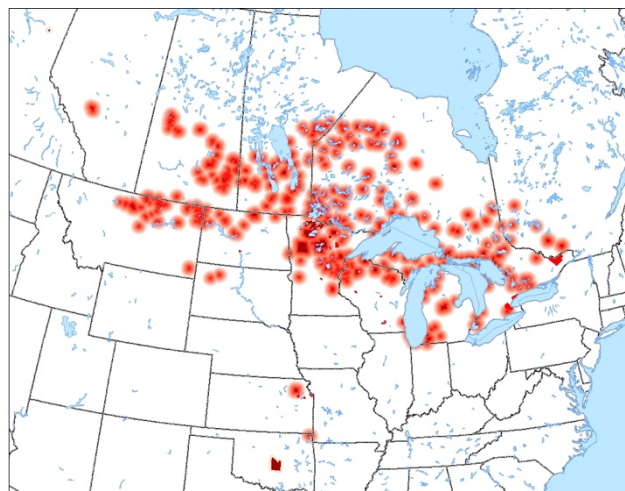
#### **Papakū Makawalu Competency Assessment (PMCA)**

Papakū Makawalu is the traditional means by which Native Hawaiians observed, interpreted, organized and categorized the natural world around them, and conveys its

practitioner's heightened level of awareness, knowledge, understanding, and acknowledgement of mankind's own innate power of observation and ability to process information. Papakū Makawalu is a comprehensive educational initiative that incorporates a conceptual framework, student learning expectations, program of assessment, and models for the teaching of environmental science. Assessment of student learning is a core element of the Papakū Makawalu methodology. The Papakū Makawalu Competency Assessment (PMCA) was developed as the formal science assessment for the Ka'Umeke Kā'eo Hawaiian language immersion HFCS located in Keaukaha, Hilo, Hawai'i Island. The PMCA assists in monitoring the transition from the novice stage of learning to the proficient stage of learning described by the Dreyfus Five-stage Model of Adult Skill Acquisition (Dreyfus & Dreyfus, 1980). The assessments are based on Papakū Makawalu grade level expectations developed during the 2014-15 academic school year. Development and administration of the PMCA began in 2014-15 with the Papa 1 and Papa 5 assessments administered to students in grades 1 and 5, respectively. From 2015-16 through 2017-18 academic school years, the PMCA was comprised of a series of grade level assessments administered annually. Each of the grade level assessments focus on the domains of Wae'ano (categorization), Kilo (keen observation) and Makawalu (analysis and synthesis of information). In July 2016, Ka'Umeke Kā'eo contracted with the author to assist with the psychometric validation of the suite of Papakū Makawalu assessments. Development work on the PMCA was halted in 2018-19 academic year but has been reactivated for the 2019-20 academic year.

**Dibishgaademgak Anishinaabemowin (Measuring Anishinaabemowin)**

Anishinaabemowin, an Algonquian language, is spoken widely throughout Canada by approximately 20,000 Anishinaabe people (Statistics Canada, 2017). In Canada, Anishinaabemowin communities are found in southwestern Quebec, Ontario, southern Manitoba and parts of southern Saskatchewan. In the United States, Anishinaabemowin communities exist along the northern border from Montana to Michigan and as far south as Oklahoma (see Figure 1 below). Considered “endangered” in the United States, there are an estimated seven hundred speakers of Anishinaabemowin across the United States (Hermes, Bang, & Marin, 2012).



*Figure 4.* Location of all Anishinaabe Reservations/Reserves in North America, with diffusion rings about communities speaking an Anishinaabe language. Cities with Anishinaabe population also shown (Lippert, 2007).

This research supports an Anishinaabemowin language assessment initiative sponsored by Kenjgewin Teg, an educational institution located in M’Chigeeng on Mnidoo Mnising (Manitoulin Island), Ontario, Canada. Kenjgewin Teg is governed by the United Chiefs and Councils of Mnidoo Mnising who represent eight First Nations: Sagamok Anishnawbek First Nation, Sheguiandah First Nation, Aundeck Omni Kaning First Nation, M’Chigeeng First



Nation, Zhiibaahaasing First Nation, Sheshegwaning First Nation, Whitefish River First Nation, the Mamawmatawa Holistic Education Centre, and the Constance Lake First Nation. On October 14, 2011, these eight First Nations established the Anishinabek Language Declaration that asserted their right to: “revitalize, use, develop and transmit to future generations their histories, languages, oral traditions, philosophies, writing systems and literatures, and to designate and retain their own names for communities, places and persons.” Included within the Anishinaabek Language Declaration was the expectation that employees of Kenjgewin Teg will “provide all work and service functions in their ancestral language by 2030” (United Chiefs and Councils of Mnidoo Mnising, 2011).

The eight First Nations sought the design and development an Anishinaabemowin language assessment that could be used to support the Kenjgewin Teg in meeting this long-term goal. In 2014, the Kantaa-Anishinaabemi language assessment was created to assist in determining Kenjgewin Teg employees’ proficiency in Anishinaabemowin. It was based on the Common European Framework of Reference for Languages (2001) and the 2012 proficiency guidelines established by the American Council on the Teaching of Foreign Languages (2012). In December 2017, Kenjgewin Teg contracted with the author to update the assessment to align better with Anishinaabek principles of learning and overall worldview. The revised assessment, known as the Dibishgaademgak Anishinaabemowin (Measuring Anishinaabemowin) is expected to provide a better understanding of the performance of Anishinaabemowin learners than that obtained through the prior language assessment.

### **Measurement approach**

The development of these assessments is guided by three broad questions:

1. What indigenous knowledge is being assessed?
2. What dimensions of the indigenous knowledge are being assessed?
3. Are the indigenous knowledge dimensions structured unidimensionally or multidimensionally?

In addition, four aspects of the structure of these assessments are considered prior to the selection of the measurement approach. The student ratings on both assessments, are obtained through a learner performance rating process. Second, these performance ratings adhere to the ordinal level of measurement. Third, the number of rating options within each of the assessment items can vary throughout the instrument. Finally, the Papakū Makawalu utilizes a multidimensional construct framework whereas the Dibishgaademgak Anishinaabemowin utilizes one that is unidimensional. The internal structure of these assessments make them an ideal candidate for the application of the multidimensional random coefficients multinomial logit model or MRCMLM (Adams, Wilson, & Wang, 1997). The approach integrates the Partial Credit Model (Masters, 1982) and is applied when multiple dimensions are present within a single overarching construct. MRCMLM is grounded in the 1-parameter logistic (1PL) Item Response Theory (IRT) model, commonly referred to as the Rasch Model (Rasch, 1960).

### **Significance**

This research introduces two important concepts for the field of educational measurement. “Measurement disjuncture” is defined here as the misalignment that occurs when elements of an instrument-development process from one worldview are applied to the instrument-development process of another worldview. Measurement disjuncture affects the

establishment of measurement validity, and, hence, the inferences made based on the scores derived from such assessments. This is primarily due to the introduction of measurement error caused by the misalignment. In quantitative educational research studies, the presence of measurement disjuncture leads to an increase in the Type II error rate.

“Culturally specific assessment” is defined as (a) assessment that supports the academic development of students; (b) is inclusive of a willingness to nurture and support cultural competence; (c) aims to support the development of a sociopolitical or critical consciousness within students; (d) is focused on constructs and measures of importance to educational practitioners and other key stakeholders; and (e) functions within a system of knowledge that exists within a named worldview. Culturally specific assessments are offered as a possible solution to the problem of measurement disjuncture.

The developers of the Papakū Makawalu Competency Assessment and the Dibishgaademgak Anishinaabemowin aspire to approach measurement from a stages- and performance-based perspective that aligns well with their Indigenous notions of knowledge attainment. The articulation of the domains of knowledge and stages of learning are offered by Papakū Makawalu experts and master speakers and teachers of Anishinaabemowin.

As with other educational assessments, these assessments exist within a self-determined worldview. The proposed culturally specific instrument development process is not a significant directional shift for the field of assessment. Rather, the *group* identifying the system of knowledge and naming the worldview is the significant directional shift. This has broader implications for the establishment of educational and assessment autonomy as exemplified by the statements below:

*We assert the right to educate ourselves within our own worldview. This is the declaration of our educational autonomy.*

*We assert the right to develop assessments within our own worldview. This is the declaration of our assessment autonomy.*

This research will continue with a formal description and validation of the methodology used to develop both the Papakū Makawalu Competency Assessment and the Dibishgaademgak Anishinaabemowin. A long-term research agenda will focus on defining, validating, and utilizing a culturally specific assessment-development process in support of educational efforts of Indigenous people while minimizing measurement disjuncture thereby increasing the measurement validity of the resultant assessment scores. To support this broader agenda, outreach efforts to Indigenous communities seeking to develop culturally specific assessments are being conducted.

### References

- Adams, R. J., Wilson, M., & Wang, W. (1997). The Multidimensional Random Coefficients Multinomial Logit Model. *Applied Psychological Measurement*, 21(1), 1–23.
- American Council for Teaching Foreign Languages. (2012). *ACTFL proficiency guidelines 2012* (pp. 1–24). Alexandria, VA: American Council on the Teaching of Foreign Languages.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Appadurai, A. (1996). *Modernity at large: Cultural dimensions of globalization*. Minneapolis, MN: University of Minnesota Press.
- Au, K. H.-P., & Jordan, C. (1981). Teaching reading to Hawaiian children: Finding a culturally appropriate solution. In H. Trueba, G. Guthrie, & K. H.-P. Au (Eds.), *Culture and the bilingual classroom: Studies in classroom ethnography* (pp. 216–242). Rowley, MA: Newbury House.
- Bordeaux, R. (1995). *Assessment for American Indian and Alaska native learners*. Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools, Appalachia Educational Laboratory.
- Bougie, É., Wright, S. C., & Taylor, D. M. (2003). Early heritage-language education and the abrupt shift to a dominant-language classroom: Impact on the personal and collective esteem of Inuit children in Arctic Québec. *International Journal of Bilingual Education and Bilingualism*, 6, 349–373.
- Brown-Jeffy, S., & Cooper, J. E. (2011). Toward a conceptual framework of culturally relevant pedagogy: An overview of the conceptual and theoretical literature. *Teacher Education Quarterly*, 65–84.
- Cazden, C., & Leggett, E. (1981). *Culturally responsive education: Recommendations for achieving Lau remedies II*. Washington, DC: U.S. Department of Health, Education & Welfare, National Institute of Education.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment* (pp. 1–260). Press Syndicate of the University of Cambridge.

- Day-Murdoch, I. (2018). *Indigenous knowledge and our connection to the land*. Retrieved from Kenjgewin Teg website:  
[http://www.kenjgewinteg.ca/uploads/1/4/7/8/1478467/indigenous\\_knowledge\\_volume1\\_print\\_version\\_.pdf](http://www.kenjgewinteg.ca/uploads/1/4/7/8/1478467/indigenous_knowledge_volume1_print_version_.pdf)
- Dreyfus, S. E., & Dreyfus, H. L. (1980). *A Five-Stage Model of the Mental Activities Involved in Directed Skill Acquisition*: (pp. 1–18). <https://doi.org/10.21236/ADA084551>
- Edwards, J. (2006). Language revitalization and its discontents: An essay and review of saving languages: An introduction to language revitalization. In L. Grenoble & Whaley, Lindsay (Eds.), *Saving languages: An introduction to language revitalization* (pp. 101–120). Cambridge, England: Cambridge University Press.
- Erickson, F., & Mohatt, C. (1982). Cultural organization and participation structures in two classrooms of Indian students. In G. Spindler (Ed.), *Doing the ethnography of schooling* (pp. 131–174). New York, NY: Holt, Rinehart & Winston.
- Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Continuum.
- Hermes, M., Bang, M., & Marin, A. (2012). Designing Indigenous language revitalization. *Harvard Educational Review*, 82, 381–402.
- Hood, S. (1998). Culturally responsive performance-based assessment: Conceptual and psychometric considerations. *The Journal of Negro Education*, 67, 187–196.
- Jordan, C. (1985). Translating culture: From ethnographic information to educational program. *Anthropology & Education Quarterly*, 16, 105–123.
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, 32, 465–491.
- Ledlow, S. (1992). Is cultural discontinuity an adequate explanation for dropping out? *Journal of American Indian Education*, 31, 21–36.
- Lippert, C. J. (2007). *Location of all Anishinaabe reservations/reserves in North America, with diffusion rings about communities speaking an Anishinaabe language*. Retrieved from <https://upload.wikimedia.org/wikipedia/commons/c/c1/Anishinaabewaki.jpg>
- Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika*, (2), 149–174.  
<https://doi.org/10.1007/BF02296272>

- Meek, B. A. (2007). Respecting the language of elders: Ideological shift and linguistic discontinuity in a Northern Athapascan community. *Journal of Linguistic Anthropology*, 17, 23–43.
- Meek, B. A. (2010). *We are our language: An ethnography of language revitalization in a Northern Athabascan community*. Tucson, AZ: University of Arizona Press. cat00548a.
- Mohatt, G., Erickson, F., Trueba, H., & Guthrie, G. (1981). Cultural differences in teaching styles in an Odawa school: A sociolinguistic approach. In *Culture and the bilingual classroom: Studies in classroom ethnography*. Rowley, MA: Newbury.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Danmark: Paedagogiske Institut.
- Seider, S., Graves, D., El-Amin, A., Soutter, M., Tamerat, J., Jennett, P., ... Johannsen, J. (2018). Developing sociopolitical consciousness of race and social class inequality in adolescents attending progressive and no excuses urban secondary schools. *Applied Developmental Science*, 22, 169–187.
- Solano-Flores, G. (2011). Assessing the cultural validity of assessment practices: An introduction. In M. del R. Basterra, E. Trumbull, & G. Solano-Flores (Eds.), *Cultural validity in assessment: Addressing linguistic and cultural diversity* (pp. 3–21). New York, NY: Routledge.
- Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching*, 38, 553–573. <https://doi.org/10.1002/tea.1018>
- Stake, R. (1973, October). *Program evaluation particularly responsive evaluation*. Presented at the New Trends in Evaluation, Goteborg, Sweden.
- Stake, R. (2011). Program evaluation particularly responsive evaluation. *Journal of MultiDisciplinary Evaluation*, 7(15), 180–201.
- Statistics Canada. (2017). Aboriginal languages in Canada, 2016 Census of population. Retrieved from Aboriginal languages in Canada, 2016 Census of Population website: <https://www150.statcan.gc.ca/n1/pub/11-627-m/11-627-m2017035-eng.htm>
- Sul, D. A. (2019, March). *Reclaiming educational autonomy and minimizing measurement disjuncture through a culturally specific assessment development process*. Presented at the Culturally Responsive Evaluation and Assessment Conference, Chicago, IL.

- Tyler, K. M., Uqdah, A. L., Dillihunt, M. L., Beatty-Hazelbaker, R., Conner, T., Gadson, N., ... Stevens, R. (2008). Cultural discontinuity: Toward a quantitative investigation of a major hypothesis in education. *Educational Researcher*, 37, 280–297.
- United Chiefs and Councils of Mnidoo Mnising. (2011). Anishinabek language declaration. Retrieved from Anishinabek Language Declaration website:  
<http://www.uccmm.ca/anishinabek-language-declaration.html>
- Vogt, L. A., Jordan, C., & Tharp, R. G. (1987). Explaining school failure, producing school success: Two cases. *Anthropology & Education Quarterly*, 18, 276–286.
- Wiesner, J. L. (2006). *School climate interventions for Native American students: Minimizing cultural discontinuity in public schools*. University of Wisconsin-Stout, Menomonie, WI.
- Wyman, L., Marlow, P., Andrew, F. C., Miller, G. S., Nicholai, R. C., & Rearden, N. Y. (2010). Focusing on long-term language goals in challenging times: A Yup'ik example. *Journal of American Indian Education*, 49, 28–49.