Can global learning metrics be culturally responsive?

Stafford Hood, University of Illinois at Urbana-Champaign

I am largely in agreement that the questions being raised at the symposium about global learning metrics and approaches to assessment are indeed some of the “right” questions. I have come to accept the personal reality that one cannot always be expected to have the right answer because it is too often the case that our reasonably good answers typically only raise more questions. As I think it should be. However, it is imperative that we have the right questions. I concur that the question as to whether global learning metrics can be culturally responsive, (within the context of hopefully some consensual agreement about what constitutes good and quality education) is one of the right questions. My initial reflex answer to this question is: I hope so; it depends; and why should it?

I was pleased that the conference organizers realistically acknowledged that any consideration about global learning metrics must accept the pervasive and long standing pattern to ignore/neglect any meaningful consideration of culture and cultural contexts that continues to exist in too many places. With this in mind, the expectation for any reasonably acceptable answer about whether global learning metrics can be culturally responsive has multiple layers of complexity, least of all being what different cultural groups believe should be good or quality education-- that is in their best interest. At the same time, of equal importance is what each of these groups may consider to be acceptable evidence that their children have received a good or quality education-- that is in their best interest.

I continue to believe that the assessment tools and systems we use should be responsive to the cultural background and experiences of our traditionally disenfranchised racial minorities, even if some are culturally specific. I also believe that the assessment devices and systems we use should be culturally valid to increase the possibility that they provide accurate and useful information to effectively educate these students. Validity is at the core of any consideration about assessment instruments we use to measure student learning. However, this simplicity becomes a bit more complex when the question is whether the instrument is measuring student learning equally well for all students, particularly those who have been traditionally disenfranchised, by their educational, economic, social and political systems. This is where it starts to get interesting.

It is the case that large and historical patterns of differences in the test performance of groups based on race, gender, and income have been too often the rule rather than the exception. Therefore, forcing one to wonder how much can we trust the accuracy and legitimacy of interpretations based on these scores. For example, the historical disproportionately low performance of African American students on standardized achievement tests, when compared to their white counterparts, continues to be prevalent in our conversations. This pattern makes it difficult to easily dismiss Keena Arbuthnot’s (2014) question, as to whether there is “…some issue surrounding Black students or their culture that can explain why the tests could possibly be measuring something different for them in comparison to White students”?

My early thinking on culturally responsive assessment in the late 1990’s was informed by the clarity that was emerging from the work of Gloria Ladson-Billings, Jaqueline Jordan Irvine, Carol Lee, and others around culturally relevant/responsive pedagogy. I pondered, with a couple of others about how to intervene in the test development and validation process through statistical and judgmental bias review approaches. I argued for the development of assessments that:

“…are grounded in the cultural contexts of examinees of color. If it is true that culturally responsive forms of assessment can quite possibly improve the assessment of what examinees of color know and can do relative to specific learning outcomes, then it is incumbent upon us to begin discussing how these devices can be developed and validated” (Hood, 1998 p.. 188).

Solano-Flores and Nelson-Barber (2001) provided a clearer and more substantive articulation on cultural validity in assessment. They defined cultural validity as:

“…the effectiveness with which…assessment addresses the socio-cultural influences that shape student thinking and the ways in which students make sense of …items and respond to them. These socio-cultural influences include the sets of values, beliefs, experiences, communication patterns, teaching and learning styles, and epistemologies inherent in the students' cultural backgrounds, and the socioeconomic conditions prevailing in their cultural groups” (as cited in Solano-Flores 2011, p.3)

Solano-Flores (2011) would later remind us that any product produced by a human being will be imperfect. Therefore tests are also imperfect and are in fact “cultural artifacts” reflecting the dominant culture that uses them for inclusion and exclusions to societal benefits. He asserted that tests are:

“…created with the intent to meet certain social needs or to comply with the mandates and legislation established in a society; they are written in a language (dialect of that language) used by those who develop them; their content is a reflection of the skills, competencies, forms of knowledge, and communication styles valued by a society—or the influential groups of that society;

He would further state that tests “… assume among test takers full familiarity with the contexts used to frame problems, the ways in which questions are worded and the expected ways to answer those questions”

As I maintain a certain level of optimism, it remains difficult not to hear the contrarian voices. For example, Mehan (2008) rightfully reminds us that students from poor, ethnic, and linguistic minority backgrounds are restrained in the starting blocks before they allowed to run the “race for success”. Whether this is by accident or design I leave for each of us to individually decide. Still I think we can generally agree that in too many of our classrooms, racial minority and poor students find themselves in situations that are “toxic or disabling” with the prospect of learning a distant possibility with assessment seemingly being used “…for purposes of behavioral control and for coercion to learn” (Boykin, 2013). Believing this to be the case, I agree with Boykin that the success of our traditionally underserved students in the classroom and their performance on high stakes tests is a social justice issue. He states:

“In all, the social justice argument is that the bias, problematic or illegitimacy of such measures must be exposed and rectified since in their present forms these assessments do not allow the abilities and propensities of certain groups to be accurately or adequately ascertained… (p.33)

He goes on further to say:

“we need to determine how such assessments can more truly discern the performance capabilities of certain groups without undermining the psychometric quality of such assessments; without unduly penalizing those who fare well on such measures in their present forms; and without inappropriately undermining the pursuit of excellence in the standards of the outcomes sought” (ibid.)

Indeed, we must build better assessments that allow for the abilities of “certain groups” to be accurately measured while also building these instruments so that the psychometric quality is not compromised. This will not be an easy task but it certainly must be done.