



OPEN ACCESS

EDITED BY
Richard James Wingate,
King's College London,
United Kingdom

REVIEWED BY
David Hay,
King's College London,
United Kingdom

*CORRESPONDENCE
Rodica Ivan
rivan@altusassessments.com

SPECIALTY SECTION
This article was submitted to
Language, Culture and Diversity,
a section of the journal
Frontiers in Education

RECEIVED 05 August 2022
ACCEPTED 06 October 2022
PUBLISHED 20 October 2022

CITATION
Mortaz Hejri S, Ivan R and Jama N
(2022) Assessment through a
cross-cultural lens in North American
higher education.
Front. Educ. 7:1012722.
doi: 10.3389/feduc.2022.1012722

COPYRIGHT
© 2022 Mortaz Hejri, Ivan and Jama.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Assessment through a cross-cultural lens in North American higher education

Sara Mortaz Hejri, Rodica Ivan* and Nimo Jama

Altus Assessments, Toronto, ON, Canada

KEYWORDS

culture, assessment, cultural sensitivity, cultural competence, equity, education

Introduction

Cross-cultural perspectives are of paramount importance for educational institutions in increasingly diverse communities. Although this notion has been discussed extensively in teaching and learning contexts (e.g., Sugahara and Boland, 2010; Gay, 2013; Cortina et al., 2017), it has often been overlooked when assessing learners (Solano-Flores, 2019). For that reason, in this paper, we focus on the intersection of assessments and cross-cultural perspectives and review three of its interrelated facets. For reasons of space, we restrict our discussion to the context of higher education in North America, but these findings may also inform other culturally diverse educational environments. First, we discuss cross-cultural assessments and highlight some of the challenges they face. Then, we turn to the development of culturally sensitive assessments and present some strategies that can be used for this purpose. Finally, we shift toward discussing how cross-cultural perspectives have been incorporated into competency-based education and assessments as cultural competence, which professionals need to demonstrate alongside specialized knowledge and practical skills.

Challenges to cross-cultural assessments

Cross-cultural assessments generally refer to the use of standardized tests for culturally and linguistically diverse populations (Ortiz and Lella, 2005). A starting point for cross-cultural assessments was the language classroom, where individuals would have to learn English, for instance, as well as adapt to Anglo-Saxon and/or North American culture (Upshur, 1966; Phillipson, 1992). However, any assessment administered to students from diverse cultural backgrounds, is in effect a cross-cultural assessment (Lyons et al., 2021). Thus, assessments are required to be sensitive to cultural

differences¹ and free from cultural bias [American Psychological Association (APA), (2002); Solano-Flores, 2019]. While a crucial goal to strive toward, the process of achieving a cross-cultural assessment faces a number of challenges.

One difficulty cross-cultural assessments face is the very definition of culture, which is by no means fixed (Lang, 1997). An individual's cultural background is not merely a matter of race or language, but at the intersection of heritage, language, beliefs, knowledge, behavior, common experience, and self-identity (gender, sexual orientation, etc.) (Ortiz and Lella, 2005; Montenegro and Jankowski, 2017). A culturally sensitive assessment must consider all the aspects that make up cultural diversity, as well as their complex interactions.

An additional challenge is that, like all human artifacts, assessments are affected by the cultural background of their developers (Cole, 1999; Solano-Flores, 2019). For instance, tests developed in North America or in the UK will invariably be imbued with content that reflects mainstream North American / Western European values (Phillipson, 1992; Ortiz and Lella, 2005), which cater to White Western conceptions of learning and assessments rather than to cross-cultural pedagogies and ways of knowing (Graham, 2020). Individuals who do not adhere to mainstream views, learning strategies, and life experiences are likely to be disadvantaged by such assessments. Not only does socioeconomic privilege lead to higher scores for racial majority applicants (Smith and Reeves, 2020; Whitcomb et al., 2021), but students from different cultural backgrounds have also been shown to prefer different methods of learning (Oxford, 1996; Hong-Nam and Leavell, 2007; Sugahara and Boland, 2010; Arbuthnot, 2020; Habók et al., 2021). In this sense, it is unlikely for any standardized test to be devoid of demographic group differences (Ortiz and Lella, 2005; Lyons et al., 2021), either when comparing applicants from different countries, or different demographic subgroups within the same country.

Societal inequities reflected in assessment scores

Assessments tend to reflect existing systemic inequities. To illustrate, we analyzed publicly available aggregate data for a standardized test commonly used in graduate school admissions in the US, namely the Graduate Record Examination (GRE) General Test. Here we focus on scores from the 2020 to 2021 application cycle (GRE Snapshot, 2022) for each of the three sections of the GRE test: Verbal Reasoning, Quantitative

¹ Various terminology is used to reference this concept: individuals and tools must be culturally "sensitive", "informed", "responsive", "aware", etc. While these terms come with different shades of meaning, they greatly overlap in their use (Montenegro and Jankowski, 2017; Frawley et al., 2020; Vasquez Guzman et al., 2021; a.m.o).

Reasoning, and Analytical Writing. We investigated race and citizenship, as two major demographic variables for which data was available². We selected the three racial subgroups with the largest sample sizes in the GRE Snapshot report (2022): Asian, Black, and White, and two major citizenship subgroups: US citizens and non-US citizens. We compared subgroup average scores using pairwise *t*-tests and reported the results of our analysis in Cohen's D effect size estimates (Cohen, 1992; Lakens, 2014) alongside descriptive statistics from the GRE Snapshot report (2022) in Table 1. Pairwise comparisons revealed moderate to large differences for all GRE sections between White and Black applicants, as well as between Black and Asian applicants; the Black applicant subgroup was associated with lower average scores. Additionally, large differences were found between US citizens and non-US citizens for Quantitative Reasoning and Analytical Writing; non-US citizens scored higher for Quantitative Reasoning and lower for Analytical Writing, on average. These results highlight significant demographic differences between racial minority and majority subgroups within the US, as well as between subgroups with different US citizenship statuses. Although the analysis was performed on individual demographic variables (as opposed to the intersection of multiple demographic variables, such as race and income), these differences are nevertheless reflective of the difficulty of designing a culturally sensitive assessment for all applicant populations, irrespective of race, nationality, or cultural background.

Steps toward mitigating inequities in cross-cultural assessments

Given the challenges discussed in the sections above, for an assessment to be designed through a cross-cultural lens, several measures are recommended.

A crucial component is inviting diverse voices at all stages of assessment development: when creating the assessment, determining its efficacy, and interpreting its results (Lyons et al., 2021). This is achieved by building multidisciplinary teams of professionals from diverse backgrounds and identities and considering multiple perspectives when developing and rating an assessment.

While diversifying the cultural make-up of assessment teams is one potential strategy, research has shown that including

² The three race subcategories are broadly used in research on demographic differences in assessments. We have also selected US citizenship status to showcase the disadvantages faced by test takers who are not native to the country where the assessment is designed. As mentioned above, these are only two aspects of cultural background, and there are many different other aspects (e.g., gender, socioeconomic status, etc.) and their complex interactions which we do not discuss here due to space constraints and unavailability of data.

TABLE 1 Descriptive statistics and Cohen's d estimates for GRE 2020–2021 scores for subgroup variables of interest.

Descriptive statistics	Number of applicants	Verbal Reasoning mean (SD)	Quantitative Reasoning mean (SD)	Analytical Writing mean (SD)
Race / ethnicity (US citizens)				
Asian	15,937	153.5 (8.0)	154.9 (8.5)	4.1 (0.8)
Black	13,364	147.4 (7.6)	144.6 (7.4)	3.4 (0.9)
White	98,851	153.4 (7.4)	151.1 (7.5)	4.0 (0.8)
Country of citizenship				
US citizens	180,924	152.6 (7.9)	150.7 (8.2)	4.0 (0.8)
Non-US citizens	186,357	150.3 (8.7)	160.7 (8.6)	3.3 (0.8)
Cohen's D estimates	Comparison group	Verbal Reasoning d [effect size]	Quantitative Reasoning d [effect size]	Analytical Writing d [effect size]
Race / ethnicity (US citizens)				
Asian	Black	0.78*** [moderate]	1.28*** [large]	0.83*** [large]
Black	White	−0.81*** [large]	−0.87*** [large]	−0.74*** [moderate]
White	Asian	−0.01 [negligible]	−0.50*** [moderate]	−0.12*** [negligible]
Country of citizenship				
US citizens	Non-US citizens	0.28*** [small]	−1.19*** [large]	0.87*** [large]

Cohen's d estimates are positive ($d > 0$) when the subgroup in the first column has a higher average than the Comparison Group, and negative ($d < 0$) when the Comparison Group has a higher average. Cohen's d effect size interpretation for absolute values (Cohen, 1992; Lakens, 2014): $|d| > 0.8$ large; $0.5 > |d| > 0.2$ moderate; $0.2 > |d| > 0.1$ small; $|d| < 0.1$ negligible. Levels of significance for the associated t-tests are flagged as follows: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

items which allow students to connect the content to their lived experiences leads to improved performance (Solano-Flores and Nelson-Barber, 2001, Mislevy and Oliveri, 2019). Additionally, student feedback can be used to inform the suitability of specific measures within assessments (Montenegro and Jankowski, 2020), such as item phrasing (i.e., how pieces of specific content might be phrased for each question). Actively seeking students' feedback and incorporating their perspectives can be done in a variety of ways. For instance, a study conducted with the Centre for Global Programs and studies at Wake Forest University (Brocato et al., 2021) convened a student advisory board with students from a range of backgrounds to gather data on their perceptions of "culture" via semi-structured interviews.

Another critical aspect of designing culturally sensitive assessments is creating opportunities for meaningful student contribution to their own assessment and inviting them to showcase their strengths and display their learning outside of standardized testing. An example includes the assessments carried out longitudinally throughout the 4-year degree program at Portland State University via electronic portfolios (Carpenter et al., 2020). For these

portfolios, students would submit assignments from the course that showed evidence of their learning as well as reflections on their progress during the academic year. In turn, faculty would provide feedback on the portfolios while collaboratively reviewing student work within the context of the course.

In order for an assessment to be culturally sensitive, contextual and structural factors must also be considered. Systemic issues related to culture, bias, power, and oppression influence society as a whole, including the institutions which carry out education and assessments. Such factors are also reflected in institutional norms and resource constraints, which may affect the interpretations of student learning outcomes. Thus, the ultimate objective is to understand not only how students are performing, but to also explore the underlying structures that students perform in and those that affect their learning (Montenegro and Jankowski, 2020). One US based example of an organizational effort to achieve this goal is the National Institute for Learning Outcomes Assessment (NILOA), who support the creation and use of culturally responsive

assessments which take into account students' needs and the context in which the assessment takes place. Aiming to foster equitable outcomes, NILOA also conducts case studies with various institutions, such as those mentioned above (Carpenter et al., 2020; Brocato et al., 2021).

Assessment of cultural competence

In addition to investing in designing culturally sensitive assessments, training programs can also focus on developing cultural competence³ in students, i.e., the ability to effectively interact with people of various cultural backgrounds (DeAngelis, 2015). Aligning with increasingly diverse populations (Mills, 2016), cultural competence has been introduced as an important skill set among professionals that enables them to work effectively across cultural boundaries (Office of Minority Health (OMH), 2000; The Royal Australasian College of Physicians (RACP), 2018; Chun and Jackson, 2021). The notion of cultural competence is currently adapted across different fields; yet, the initial idea seems to be rooted in the healthcare system (Cross et al., 1989; Frawley et al., 2020). Healthcare has become increasingly culturally diverse over the last decades, and clinicians are expected to demonstrate cultural awareness, sensitivity, and competence as they encounter patients with a variety of perspectives, beliefs, and behaviors (Betancourt, 2003; Elminowski, 2015). In the 2000s, medical organizations and accrediting bodies, such as the Association of American Medical Colleges (AAMC) and the Liaison Committee on Medical Education (LCME), brought together experts to develop new standards regarding cultural competence. Since then, a large number of training programs have been designed and delivered for health professions trainees to foster the development of knowledge, skills, and attitudes required to care for culturally diverse patients (Gozu et al., 2007).

Despite the interest toward cultural competence training (Gozu et al., 2007; Vasquez Guzman et al., 2021), the assessment of cultural competence has remained one of the main challenges (Blue Bird Jernigan, 2016). One issue concerns the risk of

including test content based on societal stereotypes (Campinha-Bacote, 2018). In addition, most instruments used for measuring cultural competence within health professions education have not been rigorously validated (Gozu et al., 2007), and the measurement of knowledge has been overemphasized (Blue Bird Jernigan, 2016). Researchers have also found it difficult to determine if a student is truly culturally competent either by observing their performance in simulated settings (Chun, 2010), or by administering attitude surveys (Gozu et al., 2007).

An additional point of contention is whether the assessment of cultural competency and professionalism overlap (Chun, 2010). While some view these two as independent concepts, others argue that there is no need for a separate measurement of cultural competence (Chun, 2010). One example of specialized assessments for cultural competence is the Tool for Assessing Cultural Competence Training (TACCT) (Lie et al., 2008). Although initially developed for curriculum development, TACCT also provides a guide on the assessment of cultural competency. Alternatively, collective tools that assess several aspects of professional performance, such as situational judgment tests (SJTs), could be used to measure cultural competence as an integrated part of a broader construct (i.e., professionalism). Such SJTs could highlight culture as one aspect of doctor-patient interactions, and also provide a significant practical contribution as an alternative performance-based assessment tool across different fields, including higher education, management, military, and engineering (Biga, 2007; Rockstuhl et al., 2015; Reinerman-Jones et al., 2016; Jesiek et al., 2020). Given the nuanced complexity of cultural competence and its various elements, one tool might not cover all aspects; however, SJTs provide an opportunity for students to express the rationale behind their behaviors and decisions. Additionally, open-response format SJTs, as opposed to closed-response format, also allow students to connect their answers to their lived experiences, and thus, allow raters to gain a deeper understanding of students' perspectives when assessing their responses.

Summary

We identified three focal points for developers of cross-cultural assessments that intend to be sensitive to individuals' diverse cultural perspectives. Although our source was North American higher education, these insights can extend to multicultural environments more broadly. We highlighted how cultural backgrounds and societal privilege are reflected in assessment scores and reflected on the difficulty of designing a culturally sensitive assessment. Then, we discussed how the inclusion of multiple perspectives at different stages in the assessment process can help alleviate differences in performance between students from different cultures. These

³ In this paper we employ the term "cultural competence" due to its wide usage in the literature. More recently, however, this term has raised some concerns, given that becoming fully competent in other cultures is next to impossible (Chun, 2010; Blue Bird Jernigan, 2016). Instead, some alternative and complementary approaches have been introduced that could offer real potential to mitigate biases and create structural changes. This includes, but is not limited to the notion of cultural humility (which encourages lifelong commitment to reflective practices and continuous learning) and the notion of structural competency (which promotes efforts aiming to eliminate racial and ethnic disparities in the healthcare system).

issues also point to the need for moving beyond designing culturally sensitive assessments and toward also incorporating measures of students' cultural competence. Fostering the ability to work effectively across diverse communities and cultures is a prerequisite toward achieving a more equitable and inclusive society.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Acknowledgments

We are grateful to the research team at Altus Assessments for all of their insightful feedback

References

- American Psychological Association (APA) (2002). Ethical principles of psychologists and code of conduct. *Am. Psychol.* 57, 1060–1073. doi: 10.1037/0003-066X.57.12.1060
- Arbuthnot, K. (2020). Reimagining assessments in the postpandemic era: creating a blueprint for the future. *Educ. Measur. Issues and Pract.* 39, 97–99. doi: 10.1111/emip.12376
- Betancourt, J. R. (2003). Cross-cultural medical education: conceptual approaches and frameworks for evaluation. *Acad. Med.: J. Assoc. Am. Med. Coll.* 78, 560–569. doi: 10.1097/00001888-200306000-00004
- Biga, A. (2007). *Measuring Diversity Management Skill: Development and Validation of a Situational Judgment Test*. USF Tampa Graduate Theses and Dissertations. Available online at: <https://scholarcommons.usf.edu/etd/633>
- Blue Bird Jernigan, V. (2016). An examination of cultural competence training in US medical education guided by the tool for assessing cultural competence training. *J. Health Disparities Res. Pract.* 9, 150–167. Available online at: <https://digitalscholarship.unlv.edu/jhdp/vol9/iss3/10/>
- Brocato, N., Clifford, M., Brunsting, N., and Villalba, J. (2021). *Wake Forest University: Campus Life and Equitable Assessment*. National Institute for Learning Outcomes Assessment (Equity Case Study), 3–6. Available online at: <https://www.learningoutcomesassessment.org/wp-content/uploads/2021/02/EquityCase-WFU-2.pdf> (accessed August 1, 2022).
- Campinha-Bacote, J. (2018). Cultural competency: a paradigm shift in the cultural competence versus cultural humility debate – part I. *OJIN: Online J. Issues Nurs.* 24. doi: 10.3912/OJIN.Vol24No01PPT20
- Carpenter, R., Reitenauer, V., and Shattuck, A. (2020). *Portland State University: General Education and Equitable Assessment*. National Institute for Learning Outcomes Assessment (Equity Case Study), 3–4. Available online at: <https://www.learningoutcomesassessment.org/wp-content/uploads/2020/06/Portland-Equity-Case.pdf> (accessed August 1, 2022).
- Chun, M. B. (2010). Pitfalls to avoid when introducing a cultural competency training initiative. *Med. Educ.* 44, 613–620. doi: 10.1111/j.1365-2923.2010.03635.x
- Chun, M. B. J., and Jackson, D. S. (2021). Scoping review of economical, efficient, and effective cultural competency measures. *Eval. Health Professions* 44, 279–292. doi: 10.1177/0163278720910244
- Cohen, J. (1992). A power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037/0033-2909.112.1.155
- Cole, M. (1999). “Culture-free versus culture-based measures of cognition” in *The Nature of Cognition*, ed R. J. Sternberg (Cambridge, MA: The MIT Press), 645–664.
- Cortina, K. S., Arel, S., and Smith-Darden, J. P. (2017). School belonging in different cultures: the effects of individualism and power distance. *Front. Educ.* 2, 56. doi: 10.3389/feduc.2017.00056
- Cross, T., Bazron, B., Dennis, K., and Isaacs, M. (1989). *Towards A Culturally Competent System of Care, Volume I*. Washington, DC: Georgetown University Child Development Center, CASSP Technical Assistance Center.
- DeAngelis, T. (2015). In search of cultural competence. *Monit. Psychol.* 46, 64. Available online at: <https://www.apa.org/monitor/2015/03/cultural-competence> doi: 10.1037/e520422015-022
- Elminowski, N. S. (2015). Developing and implementing a cultural awareness workshop for nurse practitioners. *J. Cult. Diversity* 22, 105–113.
- Frawley, J., Russell, G., and Sherwood, J. (2020). “Cultural competence and the higher education sector: a journey in the academy” in *Cultural Competence and the Higher Education Sector*, eds J. Frawley, G. Russell, J. Sherwood (Singapore: Springer), p. 3–11. doi: 10.1007/978-981-15-5362-2_1
- Gay, G. (2013). Teaching to and through cultural diversity. *Curr. Inq.* 43, 48–70. doi: 10.1111/curi.12002
- Gozu, A., Beach, M. C., Price, E. G., Gary, T. L., Robinson, K., Palacio, A., et al. (2007). Self-administered instruments to measure cultural competence of health professionals: a systematic review. *Teach. Learn. Med.* 19, 180–190. doi: 10.1080/10401330701333654
- Graham, E. J. (2020). “In Real Life, You Have to Speak Up”: civic implications of no-excuses classroom management practices. *Am. Educ. Res. J.* 57, 653–693. doi: 10.3102/0002831219861549
- GRE Snapshot (2022). *A Snapshot of the Individuals who took the GRE Test*. Available online at: <https://www.ets.org/s/gre/pdf/snapshot.pdf> (accessed August 1, 2022).
- Habók, A., Kong, Y., Ragchaa, J., and Magyar, A. (2021). Cross-cultural differences in foreign language learning strategy preferences among Hungarian, Chinese and Mongolian University students. *Heliyon* 7, e06505. doi: 10.1016/j.heliyon.2021.e06505
- Hong-Nam, K., and Leavell, A. G. (2007). A comparative study of language learning strategy use in an EFL Context: monolingual Korean and bilingual Korean-Chinese university students. *Asia Pacific Educ. Rev.* 8, 71–88. doi: 10.1007/BF03025834
- Jesiek, B., Woo, S. E., Parrigon, S., and Porter, C. M. (2020). Development of a situational judgment test for global engineering competency. *J. Eng. Educ.* 109, 470–490. doi: 10.1002/jee.20325

and valuable suggestions on the earlier versions of this manuscript.

Conflict of interest

Authors SH, RI, and NJ were employed by Altus Assessments.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *Eur. J. Soc. Psychol.* 44, 701–710. doi: 10.1002/ejsp.2023
- Lang, A. (1997). Thinking rich as well as simple: Boesch's cultural psychology in semiotic perspective. *Cult. Psychol.* 3, 383–394. doi: 10.1177/1354067X9733009
- Lie, D. A., Boker, J., Crandall, S., Degannes, C. N., Elliott, D., Henderson, P., et al. (2008). Revising the tool for assessing cultural competence training (TACCT) for curriculum evaluation: findings derived from seven US schools and expert consensus. *Med. Educ. Online* 13, 1–11. doi: 10.3402/meo.v13i.4480
- Lyons, S., Johnson, M., and Hinds, B. F. (2021). *A Call To Action: Confronting Inequity in Assessment*. Lyons Assessment Consulting. Available online at: https://www.lyonsassessmentconsulting.com/assets/files/Lyons-JohnsonHinds_CalltoAction.pdf (accessed August 1, 2022).
- Mills, S. (2016). *Cross-Cultural Measurement*. The Score. April 2016. Available online at: <https://www.apadivisions.org/division-5/publications/score/2016/04/culturally-fair-tests> (accessed August 1, 2022).
- Mislevy, R. J., and Oliveri, M. E. (2019). Digital module 09: sociocognitive assessment for diverse populations <https://ncme.elevate.commpartners.com/Educ.Measur.:IssuesandPract.38,110-111>. doi: 10.1111/emip.12302
- Montenegro, E., and Jankowski, N. A. (2017). *Equity and Assessment: Moving Towards Culturally Responsive Assessment*. (Occasional Paper No. 29). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA). Available online at: <https://files.eric.ed.gov/fulltext/ED574461.pdf> (accessed August 1, 2022).
- Montenegro, E., and Jankowski, N. A. (2020). *A New Decade for Assessment: Embedding Equity into Assessment Praxis* (Occasional Paper No. 42). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA). Available online at: <https://www.learningoutcomesassessment.org/wp-content/uploads/2020/01/A-New-Decade-for-Assessment.pdf> (accessed August 1, 2022).
- Office of Minority Health (OMH) (2000). *Assuring Cultural Competence in Health Care: Recommendations for National Standards and an Outcomes-Focused Research Agenda*. Rockville, MD: Office of Minority Health.
- Ortiz, S. O., and Lella, S. A. (2005). *Cross-Cultural Assessment*. Encyclopedia of School Psychology. SAGE Publications Inc.
- Oxford, R. L. (1996). *Language Learning Strategies Around the World: Cross-Cultural Perspectives*. Manoa: University of Hawaii Press
- Phillipson, R. (1992). *Linguistic Imperialism*. Oxford: Oxford University Press
- Reinerman-Jones, L., Matthews, G., Burke, S., and Scribner, D. (2016). A situation judgment test for military multicultural decision-making: initial psychometric studies. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 60, 1482–1486. doi: 10.1177/1541931213601340
- Rockstuhl, T., Ang, S., Ng, K.-Y., Lievens, F., and Van Dyne, L. (2015). Putting judging situations into situational judgment tests: Evidence from intercultural multimedia SJTs. *J. Appl. Psychol.* 100, 464–480. doi: 10.1037/a0038098
- Smith, E., and Reeves, R.V. (2020). *SAT Math Scores Mirror and Maintain Racial Inequity*. US Front, Brookings. December 2020. Available online at: <https://www.brookings.edu/blog/up-front/2020/12/01/sat-math-scores-mirror-and-maintain-racial-inequity> (accessed August 1, 2022).
- Solano-Flores, G. (2019). Examining cultural responsiveness in large-scale assessment: the matrix of evidence for validity argumentation. *Front. Educ.* 4, 43. doi: 10.3389/feduc.2019.00043
- Solano-Flores, G., and Nelson-Barber, S. (2001). On the cultural validity of science assessments. *J. Res. Sci. Teach.: Official J. National Assoc. Res. Sci. Teach.* 38, 553–573. doi: 10.1002/tea.1018
- Sugahara, S., and Boland, G. (2010). The role of cultural factors in the learning style preferences of accounting students: a comparative study between Japan and Australia. *Accounting Educ.: Int. J.* 19, 235–255. doi: 10.1080/09639280903208518
- The Royal Australasian College of Physicians (RACP) (2018). *Aboriginal and Torres Strait Islander Health Position Statement*. Available online at: https://www.racp.edu.au/docs/default-source/advocacy-library/racp-2018-aboriginal-and-torres-strait-islander-health-position-statement.pdf?sfvrsn=cd5c151a_4 (accessed August 1, 2022).
- Upshur, J. A. (1966). Cross-cultural testing: what to test. *Language Learning, vol. XVI, Nr. 3, 4*. doi: 10.1111/j.1467-1770.1966.tb00820.x
- Vasquez Guzman, C. E., Sussman, A. L., Kano, M., Getrich, C. M., and Williams, R. L. (2021). A comparative case study analysis of cultural competence training at 15 U.S. Medical Schools. *Acad. Med.* 96, 894–899. doi: 10.1097/ACM.0000000000004015
- Whitcomb, K. M., Cwik, S., and Singh, C. (2021). Not all disadvantages are equal: racial/ethnic minority students have largest disadvantage among demographic groups in both STEM and non-STEM GPA. *AERA Open* 7, 23328584211059823. doi: 10.1177/23328584211059823