

Research in Counseling: Are we enabling the bias that we advocate against?

Courtney Heard, M.A., ShinHwa Lee, M.S., Richard Balkin, Ph.D.

AACE 2011 Conference

Fort Worth, Texas

http://web.me.com/rsbalkin/Site/Conference_Handouts_Presentations.html

Presentation Objectives

- Discuss the differences between Test Bias and Test Fairness
- Discuss assessment bias from a Multicultural perspective
- Discuss Construct Irrelevant Variance and Construct Invariance
 - Definition
 - Prevalence in the counseling literature
- Discuss future implications for assessment and construct definition
- Discuss future implications for Multicultural literature and assessment development

Bias in Assessment

- Controversial issue of debate among many researchers
 - Writings and research on assessment bias boomed in the late 1960's to early 1970's
 - Socio-historical change in the United States
 - Civil change in academic and occupational settings
 - Ethnic integration of institutions (e.g. academic environments)
 - Instruments developed to assess various constructs (e.g. intelligence and academic achievement/ability)
 - Normed using predominately Caucasian samples of middle to higher socio-economic status.
 - Leads to advocacy for diverse sample use and fair treatment in assessment development and score interpretation

Bias in Assessment

- Example: Sedlacek & Kim (1995)
 - “Using measures normed on White populations to assess non-White people”
 - A goal in assessment is to develop an instrument that is cross-culturally valid
 - If there are cross-cultural differences in cultural and racial experiences, it is unlikely that one instrument will work equally well for all
 - This speaks to Test Fairness and potential cultural biases in assessment
 - Does bias in assessment equate to unfairness?

Bias in Assessment

- Test Bias vs. Test Fairness
 - Test Bias- Occurs when groups or subgroups experience differences in scores or score interpretations on an instrument
 - Various types of bias in assessment
 - Current presentation is covering multicultural bias in assessment
 - Test bias is a component of test fairness
 - Test Fairness- Refers to the way in which group or subgroup differences in scores are interpreted and used to make decisions about the group
 - Example of Test bias vs. Test fairness
 - Ethnic comparisons and the achievement gap
 - Have achievement tests been reliably and validly shown to be accurate measures of achievement?-"Yes"
 - Is the variable ethnicity considered an important element in defining the construct achievement- "No"

Bias in Assessment

- Construct-Irrelevant Variance
 - Occurs when there are differences in test scores based on factors that are unrelated to the construct being measured
 - Construct-irrelevant variance exists when the "test contains excess reliable variance that is irrelevant to the interpreted construct" (Messick, 1989, p.34).
 - Two types of construct-irrelevant variance:
 - Construct-irrelevant difficulty: "aspects of the task that are extraneous to the focal construct make the test irrelevantly more difficult for some individuals or groups" (Messick, 1989, p. 34).
 - Construct-irrelevant easiness: "extraneous clues in item or test formats permit some individuals to respond correctly in ways irrelevant to the construct being assessed" (Messick, 1989, p. 34).

Bias in Assessment

- An Example: Rabiner, Murray, Schmid, & Malone (2004)
 - Examined the relationship between inattention and achievement between Caucasian and African American students
 - One hypothesis was that attention problems may be a factor in the achievement gap between African American and Caucasian students.
 - The authors reported that being African American was a significant positive predictor of inattention.
- The constructs were achievement and attention
 - There were reported ethnic differences in achievement and attention
 - Reporting that ethnicity is a potential variable of influence to achievement and attention is a potentially prejudicial statement
 - Ethnicity is an irrelevant factor to attention and achievement as ethnic group membership is not part of the definition of either construct

Bias in Assessment

- Construct-Irrelevant Variance
 - Presence of factors unrelated to the construct being measured which facilitates differences in test scores
 - Example: Rabiner, Murray, Schmid, & Malone (2004)
 - Factors included: The use of first graders as a sample (age)
 - The majority of the sample were students of lower socio-economic status
 - Teachers were asked to rate students inattentive behaviors using the Conners' Teacher Ratings Scale
 - Teacher's scores may be influenced by various systemic issues
 - Teachers were asked to rate students academic achievement based on their assessment of performance over the year (subjective reports)
 - Each factor a potential influence to differences in test scores between the groups

Bias in Assessment

- Rabiner, Murray, Schmid, & Malone (2004)
 - The teacher's scoring may have been influenced by factors including but not limited to:
 - Racial or ethnic group biases
 - Without the use of concrete evidence of academic achievement in courses
 - Based on perceptions of inattentiveness, these children may have been scored lower in the area of achievement
 - May lead to incorrect inferences about a specific ethnic group (e.g. in this case regarding inattention and its relationship to achievement)
 - Diminishes the validity of the scores on the assessment and may potentially impact the decisions made about a particular group

Bias in Assessment

- How can test bias be minimized?
 - Assessment Development
 - Clearly define the construct
 - Item construction
 - Use language that can be generalized across diverse groups
 - Ensure that items are related to the construct being measured
 - Develop a large number of items so that item analysis does not significantly reduce the number of valid items
 - Testing for invariance

Bias in Assessment: Testing for Invariance

- Invariance is a fancy word for equivalence.
- Dimitrov (2010) advocated for factor invariance testing to ensure that a construct being measured has the same theoretical structure for each group.
- How is this done?
 - Use EFA to identify across two different samples to ensure that the same factor pattern exists for each group.
 - Test CFA models to establish whether known factor structures fit for different groups.

Bias in Assessment

- Study Design
 - Ensure the instrument utilized for the research assesses the intended construct
 - Researchers should be aware of the norming of the assessments chosen (e.g. ethnic diversity, or, socio-economic status)
 - Ensure that the sample is representative of the larger population and results can be generalizable

Minimizing Cultural Bias in Assessment

- Be aware of the use of ethnicity as an independent variable
 - Essentially, this may imply that changes in the dependent variable are the result of ethnic group membership
 - Works against advocacy efforts of diverse populations and may support discriminatory beliefs about multicultural groups
- When choosing an assessment
 - Be aware of the normed group as it may not have included a diverse sample
 - Results may not be able to be generalized to minority groups
 - Be aware of the cultural limitations of the assessment
 - Test administration, scoring, and interpretation
- When utilizing assessments with diverse populations
 - Operate from a multicultural perspective
 - Consider assessment results in addition to other information before making recommendations for client treatment

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.
- Educational Testing Service. (2009). ETS guidelines for fairness review of assessments. Retrieved from http://www.ets.org/Media/About_ETS/pdf/overview.pdf
- Helms, J. E. (2006). Fairness is not validity or cultural bias in racial-group assessment: A quantitative perspective. *American Psychologist*, 61(8), 845-859. doi: 10.1037/0003-066X.61.8.845
- Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan
- Rabiner, D. L., Murray, D. W., Schmid, L., & Malone, P. S. (2004). An exploration of the relationship between ethnicity, attention problems, and academic achievement. *School Psychology Review*, 33(4), 498-509.
- Sedlacek, W. E., & Kim, S. H. (1995). Multicultural assessment. *ERIC Digest* (ERIC Digest No.: EDO-CG-95-24).
- Society for Industrial and Organizational Psychology. (2003). Principles for the validation and use of personnel selection procedures. Retrieved from http://www.siop.org/_Principles/principles.pdf