# CULTURAL BIAS IN ASSESSMENT: CAN CREATIVITY ASSESSMENT HELP?

KYUNG HEE KIM
DARYA ZABELINA

*Abstract*

Culture and background may lead to the inaccuracy of assessments, including traditional tests and alternative assessments. Standardized tests intend to measure intelligence and general knowledge, but they are normed based on the knowledge and values of the majority groups, which can create bias against minority groups, including gender, race, community status, and persons with different language backgrounds, socioeconomic status, and culture. Although alternative assessments are considered to be more culturally fair, they are still not completely fair. Creativity is as important as intelligence, and creativity assessment measures are normed on individual cultures. To reduce bias, we suggest that creativity assessments be added to traditional tests and alternative assessments.

*Key Words:* assessment, cultural bias, creativity, standardized tests, cultural validity

## STANDARDIZED TESTS

Standardized tests assert that they are fair and impartial measures of academic performance and have become the most prevalent measures of the quality of educational programs. The U.S. No Child Left Behind Act (NCLBA) required standardized testing of students at all levels of education: elementary, middle, and high school. The NCLBA implemented a federally –mandated testing protocol that continues under the Race to the Top (RttT). Government agencies use these tests to determine that the various school systems in the U.S. equally serve every child, no matter their cultural background. However, performance gaps on standardized tests between ethnic minority groups (African American,

Latino, and Native American) and non-minority groups (mostly white/European American, but lately also Asian American students) in the U.S. prevail despite the government's effort to reduce them (Arbuthnot, 2009; Forum for Education and Democracy, 2008; Salinas & Garr, 2009; U.S. Department of Education, 2005).

Method bias may contribute to performance gaps between ethnic minority groups and non-minority groups in the U. S. due to the norming process of standardized tests. In general, most tests are normed using the scores of majority group populations. It may be inappropriate to use the same assessments with individuals of various racial/ethnic minority groups without norming the instrument to reflect those groups. If the cultural or linguistic backgrounds of the individuals being tested are not adequately represented in the norming group, the validity and reliability of the test are questionable when used with such individuals (Padilla & Borsato, 2008). For example, the Scholastic Aptitude Test (SAT) used an analogy with the word "regatta" that few African American students knew while many Caucasian counterparts were familiar with the word. Such groups may be denied access to educational and career opportunities if these tests produce inaccurate scores of knowledge and ability. In this review, we explore the possibility of expanding current testing practices to include creativity assessment. Such tests produce more uniform scores across ethnic and racial groups because they are normed for specific groups (Kim, in press; Torrance, 1977).

## VARIOUS CULTURAL BIASES

### *Cultural Biases in Interpretation and Meaning of Words in Assessment*

What is considered wise in one society may not be considered wise in another; the value and meaning of intelligence depends on cultural norms. Demonstrating the culturally-specific nature of knowledge and intelligence, Cole, Gay, Glick, and Sharp (1971) conducted an experiment in which Western participants and Kpelle participants from Liberia were given an object-sorting task. Participants were asked to sort twenty objects that were divided evenly into the linguistic categories of foods, implements, food containers, and clothing. Westerners tended to sort these objects into the groups for food and implements, while Liberian participants would routinely pair a potato with a knife because, they reasoned, the knife is used to cut the potato. When questioned, Liberian participants justified their pairings by stating that a wise person would group the items in this way. When the researchers asked them to show what an unwise person would do, they did the taxonomic sort that is more familiar to the Western culture.

In addition to biases due to culturally specific interpretation of test items, language impacts the understanding of test items. Cultural and social norms affect

how test-takers understand and interpret the wording of test questions. How they make sense of the test items can be influenced by their values, beliefs, experiences, communication patterns, teaching and learning styles, and epistemologies of their cultures and societies (e.g., Solano-Flores & Nelson-Barber, 2001).

Further, test item interpretation can be affected by test questions written in a language other than the native language of the test taker. It is important to consider a non-native English speaker's language proficiency before deciding whether to test her/him in English or the native language (Geisinger, 2003). For example, a Latino might appear acculturated to the test administrators, but may be nonetheless more proficient in Spanish than in English (Padilla, 1992). If the test is administered in English and is timed, it is likely that such a student will have more difficulties than if the test is in English, but untimed. To overcome these testing biases, researchers suggest that acculturation should be measured in addition to psychometric tests (Gopaul-McNicol & Armour-Thomas, 2002); however to date, there lacks a consensus on the most effective ways to measure acculturation (Cabassa, 2003).

Sometimes even the same word may have different meanings for different cultures. For instance, the meaning of *educacion* in Spanish is different from that of *education* in English (Reese, Balzano, Gallimore, & Goldenberg, 1995). The social skills of respectful and correct behavior are important to the Spanish when they state *educacion*, whereas only cognitive processing is important to e*ducation* in many Western societies. Another well-documented instance is when Native American students are asked, "Who is the son of your aunt?" "Brother" is selected by all of the Native American students in the fifth grade (even though the officially accepted/ expected answer would be "cousin") because all relatives of the same generation are called "brothers" in Native American culture (Shields, 1997).

The way students use English sentence structure may also depend on cultural backgrounds. Native American students use a different sentence structure from that of English speakers when forming negative questions. When they are asked, "You don't like eating this, [do you]?" they respond, "Yes" while they actually mean "Yes, you are right, I don't like eating it" (Shields, 1997). Asian students use the same sentence structure as Native Americans. Thus, tests must take into account students' ways of knowing and demonstrating their knowledge through their use of language (Solano-Flores & Trumbull, 2003; Swisher & Deyhle, 1992).

Several culturally sensitive tests have been developed to address the issue of this type of bias. Williams' (1972) Black Intelligence Test of Cultural Homogeneity (BITCH) is a better predictor of learning ability for African American students than other ethnicity groups. The results of the BITCH show that African American students perform better on the test than Caucasian students (Williams, 1975). In addition, the Naglieri Nonverbal Ability Test (NNAT) (Naglieri, 1991) and the Comprehensive Test of Nonverbal Intelligence (CTONI) (Ham-

mill, Person, & Wiederhold, 1997) have been developed as culture-free ability assessments. The NNAT does not use words or language in any of the items on the test, and the figures that make up each item are not specific to any particular culture. In addition, the NNAT has been standardized on an English-speaking sample and a Spanish-speaking sample. The CTONI provides oral or pantomime instructions, and the examinee answers by pointing to the response that he or she thinks is correct. A review of the mean standard scores on the CTONI shows that all African American, Asian American, Hispanic American, Native Indian, and Caucasian students scored well within the normal range (Hammill et al., 1997). This may be because both the NNAT and the CTONI have used standardization samples with ethnic profiles that mirror the ethnic profile of the 1990 U.S. Census (Zurcher, 1998).

### *Cultural Biases of Differential Effect of Acculturation*

Acculturation is the result of cultural adaptation due to intercultural contact, which is mainly considered as immigrants' adapt to the host (dominant or mainstream) culture. Less acculturated immigrant students from cultures with emphasis on strong ties to family may experience more difficulty learning to read English than students whose culture lacks emphasis of strong family ties (Portes & Zady, 2001). As such, students who identify with their families and intend to remain close to the family tend to have lower reading achievement in English, whereas those who report willingness to move away exhibit higher achievement. For example, reading achievement of Asian immigrant students is negatively influenced by the strength of their identification with their native culture. However, reading achievement of Hispanic counterparts is negatively influenced by the strength of their identification with American culture (Portes & Zady, 2001). This shows that Hispanic students do well if they are attached to their own culture, rather then to their host culture. This indicates that achievement is negatively influenced by perceptions and experiences of discrimination for both Hispanic and Asian immigrant students, but in different ways.

Acculturation to the culture of a particular school also creates negative effects in academic achievement. In schools in which Standard English is valued and students' home languages are devalued, whether they be different dialects of English or other languages, a student's adoption of the school's valued language may feel like a rejection of one's home community, which may be a difficult and painful decision (J. Baker, 2002). White middle-class students whose first language is English and who behave as teachers expect them to behave are provided more opportunities to learn than their peers, whereas students who do not embody these privileged ideals are often positioned as being deficient and difficult to teach (Walker, 2006).

### *Cultural Biases in Achievement Motivation*

Self image and motivation can also vary greatly by culture. Americans have a tendency to see themselves positively, which is critical for emotional well-being (Baumeister, Tice, & Hutton, 1989; Taylor & Brown, 1988). U.S. culture highlights the importance of self-esteem and emphasizes it in everyday life as a key to success and well-being (Ng, Pomerantz, & Lam, 2007). However, studies suggest that Asian students tend to be self-critical, whereas European , American, and Canadian students are typically self-enhancing (Heine, 2005; Heine, Lehman, Markus, & Kitayama, 1999). East Asian students are in general more sensitive to failure than to success, whereas European students are more sensitive to success than to failure (Heine et al., 2001; Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997). Overall, East Asian children persist longer than American students in the face of difficulty or failure (e.g., Blinco, 1992; Heine et al., 2001) because achievement is highly valued. Positive emotional function is not as important, emphasizing students' failures and downplaying their successes. In contrast, in America, positive emotional well-being is highly valued (Eid & Diener, 2001), downplaying students' failures and highlighting their successes. This may be detrimental to American students' achievement.

The type of achievement feedback thought to be most beneficial to children's advancement differs between various cultural groups. Researchers suggest that East Asian parents tend to believe that their role is to train their children to improve their performance, whereas American parents tend to believe that their role is to protect and build their children's self-esteem by providing positive feedback (Miller, Wang, Sandel, & Cho, 2002). Thus, East Asian parents downplay children's success and highlight children's failure by setting even higher standards, whereas American parents highlight children's success by praising or rewarding children (Hess, Chih-Mei, & McDevitt, 1987).

Studies on educational aspirations have demonstrated that in American school settings, the roles of teachers and of peers have greater effects than family on educational attainment and achievement (Lee, 2007; Lee & Smith, 1999; Raudenbush, 1984). However, in East Asian families, parental expectation and pressures have greater effects than school settings on children's educational attainment and achievement (Ellinger & Beckham, 1997; Lee, 2002; Lee, 2007). In East Asia, parents, families, and communities take responsibility for teaching and disciplining children, and thus achievement motivation is socially oriented (Holloway, 1988). However, in America a child's achievement is not considered to necessarily reflect the achievement of the parents, the family, or the community, and thus achievement is motivated by independent or individual goals (Bempechat & Drago-Severson, 1999).

### Cultural Biases of Placement in Special Education

Key educational and life decisions, such as placement in special education classes, are based on the results of standardized tests (Gopaul-McNicol & Armour-Thomas, 2002). Unjustifiable reliance on IQ and other evaluation tools (Losen & Orfield, 2002) has been cited as one of the factors contributing to the over-representation of minority children in special education classes. A central issue is administering assessments in English to non-native speakers (Padilla, 1988). Several lawsuits have centered on this issue, the most notable of which is *Diana v. California State Board of Education (1970).* This case was favorably settled for the student, however the ruling had little impact on professional practice (Padilla, 1988). Hispanic students who are non-native English speakers are still disproportionally represented in special education (Artiles, Rueda, Salazar, & Higareda, 2002). The unfortunate outcome of this might be that such students are likely to drop out of school at higher rates than their peers not labeled as "disabled" (Ferri & Connor, 2005).

### Other Sources of Bias

While significant progress has been made in development of unbiased tests, other sources of bias may present an obstacle to educating culturally diverse students. Teachers' different expectations for students of various racial or social classes are correlated with less effective instructional practice (Leacock, 1969; Murray, 1996). Teachers lack the knowledge and skills to successfully interact with students who are different from themselves (Ladson-Billings, 1995). Teachers often fail to identify potentially qualified students for selective programs in schools, especially when students are culturally different from them (Peterson, 2000).

It is important for teachers to familiarize themselves with common phrasings of test questions and to be more aware of the various linguistic patterns of the many ethnic/cultural groups represented by the students in their classrooms. Teachers forced to use standardized tests can help students practice their language development resulting in improved test taking skills (Shields, 1997). Native American students whose teachers integrated local standards with state standards and aligned curriculum with assessment through a portfolio process have made achievement gains (Koelsch & Trumbull, 1996). Demmert (2001) suggested that teachers maintain linguistic and cultural congruence between home and school, educate students in their native language, and use local knowledge and culture in the curriculum to improve the academic performance of Native American students. For English language learners, extra time and/or dictionaries should be provided or the language can be modified (Nelson-Barber & Trumbull, 2007). Teachers should pay attention to students' test-taking needs and accommodate their cultural backgrounds in order to increase the validity of test scores.

Educators should be aware of all potential sources of cultural bias so that they can support students from different cultures academically and psychologically. If test results are interpreted without consideration of cultural and educational factors of certain groups' the scores may inaccurately reflect the actual ability and knowledge of those students. The results may be biased even if the tests themselves appeared to be unbiased (Skiba, Knesting, & Bush, 2002). Therefore, teacher training in culturally responsive pedagogy is necessary (Klingner et al., 2005).

## ALTERNATIVE ASSESSMENTS

Alternative assessments, in contrast to traditional or standardized tests, are considered more culturally fair than paper and pencil tests. Power tests that measure how much the students know are considered more culturally fair than speed tests that measure how much the students answer correctly within a time limit. Oral instructions are more culturally fair than written instructions, non-verbal contents are more culturally fair than verbal contents, and familiar contents to the students are more culturally fair than unfamiliar contents. Therefore, among alternative assessments, the most culturally fair assessments would be non-timed, hands-on performance assessment on familiar contents with oral instructions, such as group projects, informal assessments, or interviews.

### *Group Projects*

Group projects are often used as one alternative assessment method. However, the fairness of group assessment results may depend on students' cultures: African American (Aronson & Bridgeman, 1979; Berry, 2003) and Mexican American (Aronson & Bridgeman, 1979) students tend to learn well through collaborative learning methods, whereas Euro-American students do not. A meta-analysis (Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003) found that peer-assisted learning is most effective with young, urban, low-income, and minority students. However, female students, working-class students, students of color, and students who are considered low achievers by their teachers and peers tend to be marginalized and prevented from engaging in meaningful discussions with their groups (Esmonde, 2009). Further, Native American students' willingness to participate is diminished by the competitive nature of group assessments (Nelson-Barber & Estrin, 1995; Swisher & Deyhle, 1992).

### *Informal Assessment*

Teachers ask students questions during class as a way of informal assessment. However, cultural differences exist in the function of questions. In African and Asian cultures, children are expected to listen and not to ask questions (Ner-

love & Snipper, 1981). In these cultures, students are not expected to respond to questions to which the teacher already knows the answer (Heath, 1983). Many students in some these cultures will not respond to questions at all (More, 1989; Rhodes, 1988). In addition to these African and Asian students, Native American students' seeming reluctance to participate verbally in response to teacher questions in the classroom causes the students to appear to be nonverbal or silent (Dumont, 1972); however, these students have cultural expectations to behave as such.

### *Interviews*

Examiners and interviewers can also affect students' performance. Communication among strangers is contrary to the core beliefs of collectivist cultures (Kim & Choi, 1994). Thus, students from these cultures may need to establish a personal relationship with their examiner before an interview or an oral test can be determined to be a valid assessment. Fuchs and Fuchs's (1986) meta-analysis found that the effects of examiner unfamiliarity have a significant impact on students' standardized test scores, especially for low socioeconomic students. Examiner effects may lead to a bias, especially in behavior ratings (e.g., Skiba, Knesting, & Bush, 2002). Therefore, teachers conducting traditional or behavioral assessments require adequate training in culturally competent practices (Castillo, Quintana, & Zamarripa, 1999).

## ADDING CREATIVITY ASSESSMENT

Expanding assessment practices to include creativity testing might address cultural biases (Kaufman & Agars, 2009). Creativity is defined as producing something that is novel and useful (Runco & Jaeger, 2012). Observations of eminent creators show that eminence in any field requires specific domain knowledge to be joined with creativity, and high IQ alone does not lead to superior creative achievements. To maximize the impact of high IQ, creativity must be identified and cultivated (Wallach & Wing, 1969; Kim, 2005, in press). Lifetime creative accomplishments are related to creativity in childhood (Plucker, 1999, Torrance, 2002). Creativity assessment shows few differences across gender or ethnicity (Kaufman & Agars, 2009; Torrance, 1977). Further, evidence from data collected statewide on the effects of the Georgia multiple criteria rule for identifying students (Georgia Department of Education, 2010) supports the effectiveness of adding creativity assessments for identifying gifted students, especially those from underserved populations (Williams, 2000). The addition of creativity assessment as an option to meet the standards has facilitated identification of gifted students from underserved populations (Krisel & Cowan, 1997).

### *The Torrance Test of Creative Thinking*

The best option for creativity assessment currently may be to employ the Torrance Test of Creative Thinking (TTCT) (Torrance, 2008). The TTCT includes two forms (A and B) of the TTCT-Verbal and two forms (A and B) of the TTCT-Figural. The TTCT can be administered as an individual test or group test for any age and developmental level, beginning with kindergarten. The tests require 30 to 45 minutes working time, so speed is relevant. The tests require some drawing ability, but artistic quality is not necessary to receive credit. Torrance recommended the creation of a game-like, "fun" atmosphere to avoid the threatening situations associated with testing.

The TTCT –Figural consists of three activities: picture construction, picture completion, and repeated figures of lines or circles. The TTCT-Figural is comprised of five norm-referenced measures so that the numbers of points earned are relative to the norm group. The measures are Fluency, Originality, Elaboration, Abstractness of Titles, and Resistance to Premature Closure. In addition, there are thirteen criterion-referenced measures of Creative Strengths so that credit is given depending on whether the criterion appears in the responses. Fluency measures the ability to produce many ideas; Originality measures the ability to produce unique ideas; Elaboration measures the ability to produce a number of ideas beyond the minimum details; Abstractness of Titles measures the degree a title is expressed beyond obvious labeling of the pictures drawn; and Resistance to Premature Closure measures the degree of psychological openness. The thirteen Creative Strengths measure various creative thinking and personality constructs including: Emotional Expressiveness, Storytelling Articulateness, Movement or Action, Expressiveness of Titles, Synthesis of Incomplete Figures, Synthesis of Lines or Circles, Unusual Visualization, Internal Visualization, Extending or Breaking Boundaries, Humor, Richness of Imagery, Colorfulness of Imagery, and Fantasy (see Kim, 2006 for details).

The TTCT-Figural is particularly well known for being fair in terms of gender, race, community status, and for persons with different language backgrounds, socioeconomic status, and culture (Cramond, 1993; Torrance, 1977). Researchers found that in most situations there are no statistically significant differences in performance on the TTCT due to race or socioeconomic status, and in some cases, the TTCT favors African American children and children of low socioeconomic backgrounds (Torrance, 1971; Torrance & Torrance, 1972). The TTCT has been translated into over 35 languages, is the most widely used test of creativity, and its 40 years of longitudinal studies support its predictive validity (Kim, 2007).

### *The Rainbow Project*

Sternberg (2009) suggested it is possible to increase excellence and diversity simultaneously in higher education admissions. Sternberg's theory of successful intelligence (Sternberg, 1997, 1999) postulates that intelligence has three components: creative skills in generating novel ideas, analytical skills in discerning whether they are good ideas, and practical skills in implementing the ideas and persuading others of their worth. The Rainbow project measures creativity by multiple-choice items and by performance-based items (Sternberg, Grigorenko, & Jarvin, 2006). The addition of creativity measures to standardized admissions tests provided better predictive power for first-year college academic performance and reduced ethnic-group differences compared to standard admission tests in the U.S. (Sternberg et al., 2006). As a result, Tufts University now includes creativity assessment as part of the college admission procedures (Sternberg, 2008).

School assessments, like standardized tests, often emphasize analytical and memory-based skills. Success in life depends of a broader range of abilities than conventional tests can measure. For example, memory and analytical abilities may be sufficient to produce high grades in science courses, but are probably not sufficient to produce outstanding research. In particular, outstanding research must be creative in generating ideas for theories and/or experiments, analytical in discerning whether ideas are good, and practical in getting ideas funded and accepted by competitive refereed journals. Sternberg's theory of successful intelligence (Sternberg, 1997, 1999) provides one basis for improving the prediction of students' success and possibly for establishing greater equity and diversity (Bowen, Kurzweil, & Tobin, 2006). It suggests that broadening the range of skills tested in order to go beyond analytic skills and to include practical and creative skills as well, might significantly enhance the prediction of undergraduate performance beyond current levels.

## CONCLUSIONS

Creativity leverages intelligence and is a better predictor of creative accomplishments than is IQ (Kim, 2008b). Including creativity as an additional criterion to standardized tests and alternative assessments could benefit minority groups by reducing or eliminating the various cultural biases present in traditional standardized tests. Creativity assessment may allow students to be assessed based on their actual cognitive ability rather than their ability to adapt to the culture of the majority, especially when the assessment minimizes verbal components (Jellen & Urban, 1989; Torrance, 1977). Creativity plays a role in and is related to

intelligence; however, Kaufman (2010) states that ethnic differences in creativity are rare. He argues that the use of creativity in college admission assessment may reduce ethnic bias resulting in an increase in fairness in admission decisions (Kaufman, 2010). The use of creativity in standardized testing may place minorities on equal ground with the majority culture in ways that traditional standardized tests have failed or are incapable of doing.

Developing culturally competent assessments should not be the only answer to cultural fairness. In addition to culturally fair assessment, other accommodations should be made for individuals with culturally diverse backgrounds to minimize the distortions that arise from their specific disadvantages. Culturally competent assessment is more than just culturally fair testing (Skiba et al, 2002). Addressing test bias is only the tip of the iceberg. Culturally competent assessment requires the interpretation of test results to inform educators' and administrators' identification of educational contexts that may methodically increase the disadvantages of these students and then develop interventions to alleviate these disparities (Skiba et al, 2002). Moreover, considering cultural background should not be focused only on race, ethnicity, or language. Broader cultural factors should also be considered including, but not limited to, religion, gender, age, social class, sexual orientation, and others. This requires not only culturally fair assessment, but infusion of multicultural issues in teaching methods and curriculum development to attend to the special needs of culturally diverse students (Villegas & Lucas, 2002). Therefore addressing cultural bias in testing requires a multimodal intervention.

## REFERENCES

Arbuthnot, K. (2009). The effects of stereo type threat on stnadardized mathematics test performance and cognitive processing. *Harvard Educational Review, 79*, 448-472.

Aronson, E., & Bridgeman, D. (1979). Jigsaw groups and the desegregated classroom: In pursuit of common goals. *Personality and Social Psychology Bulletin, 5*, 438-446.

Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2002). English language learner representation in special education in California urban school districts. In D. J. Losen, & G. Orfield (Eds.), *Racial inequity in special education* (pp. 117-136). Cambridge, MA: Harvard Education Press.

Baker, J. (2002). Trilingualism. In L. Delpit & J. K. Dowdy (Eds.), The skin that we speak: Thoughts on language and culture in the classroom (pp. 49–62). New York:New Press.

Baumeister, R. F., Tice, D. M., & Hutton, D. G, (1989). Self-presentational motivations and personality differences in self-esteem. *Journal of Personality, 57*, 547-579.

Bempecht, J., & Drago-Severson, E. (1999). Cross-national differences in academic achievement: Beyond etic conceptions of children's understandings. *Review of Educational Research, 69*, 287-314.

Berry, R. Q. (2003). Mathematics standards, cultural styles, and learning preferences: The plight and the promise of African American students. *The Clearing House, 76,* 244–249.

Blinco, P. M. A. (1992). A cross-cultural study of task persistence of young children in Japan and the U.S. *Journal of Cross-Cultural Psychology, 23*, 407-415.

Bowen, W. G., Kurzweil, M. A., & Tobin, E. M. (2006). *Equity and excellence in American higher education.* Charlottesville, VA: University of Virginia Press.

Brown, R. T., Reynolds, C. R., & Whitaker, J. S. (1999). Bias in mental testing since Bias in Mental Testing. *School Psychology Quarterly, 14*, 208-238.

Cabassa, L. J. (2003). Measuring acculturation: Where we are and where we need to go. *Hispanic Journal of Behavioral Sciences, 25,* 127-146.

Castillo, E. M., Quintana, S. M., & Zamarripa, M. X. (1999). Cultural and linguistic issues. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Conducting school-based assessments of child and adolescent behavior* (pp. 274-308). New York: Guilford.

Cole, M., Gay, J., Glick, J., & Sharp, D. W. (1971). *The cultural context of learning and thinking*. New York: Basic Books.

Cole, N. S. (1981). Bias in testing. *American Psychologist, 36*, 1067-1077.

Cramond, B. (1993). The Torrance Tests of Creative Thinking: From design through establishment of predictive validity. In R. F. Subotnik & K. D. Arnold (Eds.), *Beyond Terman: Contemporary longitudinal studies of giftedness and talent* (pp. 229-254). Norwood, NJ: Ablex.

Demmert, W. (2001). *Improving academic performance among Native American students: A review of the research literature*. Charleston, WV: ERIC Clearinghouse on rural Education and Small Schools (ERIC Document reproduction Service No. ED463917).

Dumont, R. (1972). Learning English and how to be silent: Studies in Sioux and Cherokee classrooms. In C. Cazden, V. John, & D. Hymes (Eds.), *Functions of language in the classroom* (pp. 344-369). New York: Teachers College Press.

Eid, M., & Diener, E. (2001). Norms for experiencing emotions in different cultures: Inter- and intranational differences. *Journal of Personality and Social Psychology, 81*, 869-885.

Ellinger, T. R., & Beckham, G. M. (1997). South Korea: Placing education on top of the family agenda. *Phi Delta Kappan, 78*, 624-625.

Esmonde, I. (2009). Ideas and identities: Supporting equity in cooperative mathematics learning. Review of Educational Research, 79, 1008-1043. DOI: 10.3102/0034654309332562

Fagan, J. F. (2000). A theory of intelligence as processing: Implications for society. *Psychology, Public Policy, and Law, 6,* 361-387.

Fagan, J. G., & Holland, C. R. (2002). Equal opportunity and racial differences in IQ. *Intelligence, 30,* 361-387.

Fagan, J. F., & Holland, C. R. (2009). Culture-fair prediction of academic achievement. *Intelligence, 37*, 62-67.

Feldhausen, J. F. (2001). *Talent development in gifted education* (EDO-EC-01-5). Arlington, VA: ERIC Clearinghouse on Disabilities on Gifted Education Digest, E610, June 2001.

Ferri, B. A., & Connor, D. J. (2005). In the shadow of *Brown:* Special education and overrepresentation of students of color. *Remedial and Special Education, 26*, 93-100.

Forum for Education and Democracy. (2008). *Democracy at risk: The need for a new federal policy in education.* Washington, DC: The Forum for Education and Democracy.

Fuchs, D., & Fuchs, L. S. (1986). Test procedure bias: A meta-analysis of examiner familiarity effects. *Review of Educational Research, 56*, 243-262.

Geisinger, K. F. (2003). Testing students with limited English proficiency. In J. E. Wall, & G. H. Walz (Eds.), *Measuring up: Assessment issues for teachers, counselors, and administrators* (pp. 147-159). Greensboro, NC: CAPS Press.

Georgia Department of Education. (2010). *Resource manual for gifted education services.* Retrieved March 4, 2010 from http://public.doe.k12.ga.us/DMGetDocument.aspx/gifted_regulations.pdf?p=4BE1EECF99CD364EA5554055463F1FBB77B0B70FECF5942E12E123FE4810FFF53501CAAE8CB828386A1B54D8AFDA9790&Type=D

Gopaul-McNicol, S., & Armour-Thomas, E. (2002). *Assessment and culture: Psychological tests with minority populations*. Orlando, FL: Academic Press.

Hammill, D. D., Pearson, N. A., & Wiederhold, J. L. (1997). *Comprehensive Test of Nonverbal Intelligence (CTONI).* Austin: Pro-Ed.

Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms.* New York: Cambridge University Press.

Heine, S. J. (2005). Where is the evidence for pancultural self-enhancement? A reply to Sedikides, Gaertner, & Toguchi (2003). *Journal of Personality and Social Psychology, 89*, 531-538.

Heine, S. J., Kitayama, S., Lehman, D. R., Takata, T., Ide, E., Leung, C., et al. (2001). Divergent consequences of success and failure in Japan and North America: An investigation of self-improving motivations and malleable selves. *Journal of Personality and Social Psychology, 81,* 599-615.

Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review, 106*, 766-794.

Hess, R. D., Chih-Mei, C., & McDevitt, T. M. (1987). Cultural variations in family beliefs about children's performance in mathematics: Comparisons among People's Republic of China, Chinese-American, and Caucasian-American families, *Journal of Educational Psychology, 79*, 179-188.

Holloway, S. (1988). Concepts of ability and effort in Japan and the United States. *Review of Educational Research. 58*, 327-345.

Jensen, A. (1980). *Bias in mental testing*. New York: Free Press.

Jellen, H., & Urban, K. (1989). Assessing creative potential world-wide: The first cross-cultural application of the Test for Creative Thinking – Drawing Production (TCT-DP). *Gifted Education International,* **6**, 78–86.

Kaufman, J.C. (2010). Using creativity to reduce ethnic bias in college admissions. *Review of General Psychology*, 14(3), 189-203.

Kaufman, J. C., & Agars, M. D. (2009). Being creative with the predictors and criteria for success. *American Psychologist, 64*, 280-281.

Kidder, W. C., & Rosner, J. (2002). How the SAT creates 'Built-In-Headwinds' An educational and legal analyses of disparate impact. *Santa Clara Law Review, 43*, 131-212.

Kim, K.H. (2005). Can only intelligent people be creative? A meta-analysis. *Journal of Secondary Gifted Education*, 16, 57-66.

Kim, K. H. (2006). Can we trust creativity tests? A review of The Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal*, 18, 3-14.

Kim, K. H. (2007). The two Torrance creativity tests: The Torrance Tests of Creative Thinking and Thinking Creatively in Action and Movement. In T. Ai-Girl (Ed.), *Creativity: A handbook for teachers* (pp. 117-141). Hackensack, NJ: World Scientific.

Kim, K. H. (2008b). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. *Journal of Creative Behavior*, *42*, 106-130.

Kim, K. H. (in press). Why we need creativity tests: Creative CAT and the Torrance Tests of Creative Thinking-Verbal and Figural. *Creativity Research Journal.*

Kim, U., & Choi, S-H. (1994). Individualism, collectivism, and child development: A Korean perspective. In P. M. Greenfield & R. R. Cocking (Eds.), *Cross-cultural roots of minority child development* (pp. 227-257). Hillsdale, NJ: Erlbaum.

Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the U.S. and self-criticism in Japan. *Journal of Personality and Social Psychology*, *72*, 1245-1267.

Klingner, J. K., Artiles, A. J., Kozleski, E., Harry, B., Zion, S., Tate, W. et al. (2005). Addressing the disproportionate representation of culturally and linguistically diverse students in special education through culturally responsive educational systems. *Education Policy Analysis Archives, 13*, 1-39.

Koelsch, N., & Trumbull, E. (1996). Portfolios: Bridging cultural and linguistic worlds. In R. C. Calfee & P. Perfumo (Eds.), *Writing portfolios in the classroom: Policy and practice promise and peril* (pp. 261-284). Mahwah, NJ: Erlbaum.

Krisel, S. C., & Cowan, R.S. (1997). Georgia's journey toward multiple-criteria identification of gifted students. *Roeper Review, 20*(2), A1-A3.

Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. *Theory into Practice, 34*, 159-165.

Leacock, E. B. (1969). *Teaching and learning in city schools: A comparative study*. New York: Basic Books.

Lee, G. (2002). The role of Korean parents in the literacy development of their children. *International Journal of Early Childhood, 34*, 1-8.

Lee, J. (2007). Two worlds of private tutoring: The prevalence and causes of after-school mathematics tutoring in Korea and the United States. *Teachers College Record, 109,* 1207-1234.

Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: The role of school academic press. *American Educational Research Journal, 36,* 907-945.

Losen, D. J., & Orfield, G. (2002). Introduction. In D. J. Losen & G. Orfield (Eds.), *Racial inequity in special education* (pp. xv-xxxvii). Cambridge, MA: Harvard Education Press.

Miller, P. J., Wang, S., Sandel, T., & Cho, G. E. (2002). Self-esteem as folk theory: A comparison of European American and Taiwanese mothers' beliefs. *Parenting: Science and Practice, 2*, 209-239.

More, A. (1989). Native Indian learning styles: A review for researchers and teachers. *Journal of American Indian Education*, *27*, 15-28.

Murray, C. B. (1996). Estimating achievement performance: A confirmation bias. *Journal of Black Psychology, 22*, 67-85.

Naglieri, J. A. (1991). *Naglieri Nonverbal Ability Test*. San Antonio, TX: Harcourt Assessments.

Nelson-Barber, S., & Estrin, E. (1995). Bringing Native American perspectives to the teaching of mathematics and science. *Theory into Practice, 34*, 174-185.

Nelson-Barber, S., & Trumbull, E. (2007). Making assessment practices valid for indigenous American students, *Journal of American Indian Education, 46*, 132-47.

Nerlove, S., & Snipper, A. S. (1981). Cognitive consequences of cultural opportunity. In R. H.

Munro, R. L. Munro, & B. B. Whiting, (Eds.), *Handbook of cross-cultural human development* (pp. 423-474). New York: Garland.

Ng, F. F-Y., Pomerantz, E. M., & Lam, S-F. (2007). European American and Chinese parents' responses to children's success and failure: Implications for children's responses. *Developmental Psychology, 43*, 1239-1255.

Padilla, A. M. (1988). Early psychological assessment of Mexican-American children. *Journal of the History of the Behavioral Sciences, 24*, 113-115.

Padilla, A. M. (1992). Reflections on testing: Emerging trends and new possibilities. In K. F. Geisinger (Ed.), *Psychological testing of Hispanics* (pp. 271-284). Washington, DC: American Psychological Association.

Padilla, A. M., & Borsato, G. N. (2008). Issues in culturally appropriate psychoeducational assessment. In L. A. Suzuki, & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment* (pp. 5-21). San Francisco, CA: John Wiley & Sons, Inc.

Peterson, J. (2000). Valuing the values: Moving from tolerance to affirmation. *Reclaiming Children and Youth, 9*, 36-40.

Plucker, J. A. (1999). Longitudinal data is the proof in the pudding? Reanalyses of Torrance's (1958 to present) longitudinal data. *Creativity Research Journal*, *12*, 103-114. doi: 10.1207/s15326934crj1202.

Portes, P. R., & Zady, M. F. (2001). Determinants of reading achievement of immigrant adolescents: The role of demographic and psycho-cultural factors in school adaptation. Paper presented at the annual meeting of the American Educational Research Association in Seattle, WA, April 10-14.

Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology, 76*, 85-97.

Reese, L., Balzano, S., Gallimore, R., & Goldenberg, C. (1995). The concept of education: Latino family values and American schooling. *International Journal of Educational research, 23*, 57-81.

Rhodes, R. (1988). Native American learning styles: Implications for teaching and testing. In Arizona Department of Education (Ed.), *Proceedings of the eighth annual Native American Language Issues Institute* (pp. 11-21). Choctaw, OK: Native American Language Issues Institute.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peerassisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95,* 240–257.

Rosenbach, J. H., & Mowder, B. A. (1981). Test bias: The other side of the coin. *Psychology in the Schools, 18*, 450-454.

Runco, M. A., & Jaeger, G. J. (2012) The standard definition of creativity. *Creativity Research Journal, 24*, 92-96. DOI: 10.1080/10400419.2012.650092

Salinas, M. F., & Garr, J. (2009). Effect of learner-centered education on the academic outcomes of minority groups. *Journal of Instructional Psychology, 36*, 226-237.

Shepard, L. A. (1987). The case for bias in tests of achievement and acholastic aptitude. In S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and controversy* (pp. 177-190). New York: Falmer Press.

Shields, C. M. (1997). Learning about assessment from Native American schools: Advocacy and empowerment. *Theory into Practice, 36*, 102-109.

Skiba, R. J., Knesting, K., & Bush, L. D. (2002). Culturally competent assessment: More than nonbiased tests. *Journal of Child and Family Studies, 11*, 61-78.

Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching, 38*, 553-573.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English language learners. *Educational Researcher, 32*, 3-13.

Sternberg, R. J. (1997). *Successful intelligence.* New York: Plume.

Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology, 3*, 292-316.

Sternberg, R. J. (2008). Applying psychological theories to educational practice. *American Educational Research Journal, 45*, 150-165.

Sternberg, R. J., Grigorenko, E. L., & Jarvin, L. (2006). Identification of the gifted in the new millennium: Two assessments for ability testing and for the broad identification of gifted students. *KEDI Journal of Educational Policy, 3*(2), 7-27.

Sternberg, R. J., & The Rainbow Project Collaborators (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence, 34,* 321–350.

Swisher, K., & Deyhle, D. (1992). Adapting instruction to culture. In J. Reyhner, J. (Ed.), *Teaching American Indian students* (pp. 81-95). Norman, OK: University of Oklahoma Press.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.

Torrance, E. P. (1971). Are the torrance tests of creative thinking biased against or in favor of 'disadvantaged' groups? *Gifted Child Quarterly, 15*, 75-80.

Torrance, E.P. (1977). *Discovery and nurturance of giftedness in the culturally different*. Reston, VA: Council on Exceptional Children.

Torrance, E. P. (2002). *The Manifesto: A guide to developing a creative career.* West Westport, CT: Ablex Publishing.

Torrance, E. P. (2008). *The Torrance Tests of Creative Thinking Norms-Technical Manual Figural (Streamlined) Forms A & B*. Bensenville, IL: Scholastic Testing Services, Inc.

Torrance, E. P., & Torrance, P. (1972). Combining creative problem-solving with creative expressive activities in the education of disadvantaged young people, *Journal of Creative Behavior, 6*, 1-10.

US Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2005). *NAEP – The Nation's Report Card.* Downloaded from http://nece.ed.gov/nationsreportcard.

Valencia, R. R., & Suzuki, L. A. (2000). *Intelligence testing and minority students: Foundations, performance factors, and assessment issues*. Thousand Oaks, CA: Sage.

Villegas, A. M., & Lucas, T. (2002). Preparing culturally responsive teachers: Rethinking the curriculum. *Journal of Teacher Education*, 53(1), 20-32.

Walker, E. N. (2006). Urban high school students' academic communities and their effects on mathematics success. *American Educational Research Journal,* 43(1), 43–78.

Wallach, M.A., & Wing, C.W. (1969). *The talented student: A validation of the creativity-intelligence distinction*. New York: Holt, Rinehart and Winston, Inc.

Williams, R. L. (1972). *Black Intelligent Test for Cultural Homogeneity*. St. Louise, MO: Williams and Associates.

Williams, R. L. (1975). The BITCH-100: A culture specific test. *Journal of African Issues, 3*, 103-116.

Williams, E. (2000). *The history of the evolution of gifted identification procedures in Georgia.* (Doctoral dissertation, University of Georgia, 2000). *Dissertation Abstracts International, 160,* 153.

Zurcher, R. (1998). Issues and trends in culture-fair assessment. *Intervention in School and Clinic, 34*, 103-106.