

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

The Nebraska Educator: A Student-Led Journal Department of Teaching, Learning and Teacher Education

10-5-2021

Multiracial Individuals and Educational Testing

Karen Alexander

University of Nebraska-Lincoln, kalexander@huskers.unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/nebeducator>



Part of the [Bilingual, Multilingual, and Multicultural Education Commons](#), [Quantitative Psychology Commons](#), and the [Teacher Education and Professional Development Commons](#)

Alexander, Karen, "Multiracial Individuals and Educational Testing" (2021). *The Nebraska Educator: A Student-Led Journal*. 61.

<https://digitalcommons.unl.edu/nebeducator/61>

This Article is brought to you for free and open access by the Department of Teaching, Learning and Teacher Education at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in The Nebraska Educator: A Student-Led Journal by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Multiracial Individuals and Educational Testing

Karen Alexander

Department of Educational Psychology
University of Nebraska-Lincoln

Abstract

A literature review focused on quantitative measures and methods regarding multiracial individuals and educational testing revealed that multiracial individuals are uniquely different than monoracial individuals in terms of their racial identity and these unique identities interact with test scores. Until recently, this uniqueness has been ignored by institutions and within the field of educational testing. The uniqueness of multiracial identity should be taken into consideration when using test measures to make decisions for selection and when comparing group outcomes. The review provides a brief picture regarding the history of categorization of multiracial individuals and current research which connects the multiracial experience to test score performance, followed by information on data collection, data coding, data analysis, implications, and recommendations. Suggested methods to address the methodological and analytical challenges of how to categorize multiracial individuals for purposes of group comparisons are challenging and frankly, unsatisfying. Yet, there are some clear recommendations such as allowing individuals to check as many racial/ethnic categories that apply to their identity versus forcing a choice of one race or using “Other” as an option. The limited research regarding multiracial individuals and educational tests supports the need for further research in this field.

Keywords: multiracial identity, educational tests, racial identity, standardized tests, psychometrics

In 1997 Tiger Woods, who is one-fourth Black, one-fourth Thai, one-fourth Chinese, one-eighth White and one-eighth American Indian, identified himself as “Cablinasian” (AP News, 1997). Tiger’s statement made national news due to his unique descriptor of how he identifies racially/ethnically, and because it brought to light the challenge that many, if not all, multiracial people in the United States have in terms of racial/ethnic identity. Multiracial individuals do not fit neatly into the demographic boxes that have been used in the past by the U.S. Census Bureau (USCB) (Jones & Smith, 2001), in surveys, and on standardized testing materials, etc. The complexity of racial/ethnic identity for people like Tiger Woods has ramifications on the interpretations and use of test scores categorized by race/ethnicity in terms of validity and reliability. The purpose of this review is to summarize the literature regarding the educational testing of multiracial individuals in terms of measurement practices, and implications for research as well as policy, and recommendations, particularly in terms of categorization of race/ethnicity. Because educational test scores are used for many decisions such as college admittance, scholarship offers, course options for students, etc., it is important to understand the effects of test scores on multiracial individuals. The main body of this manuscript is preceded by a framework that includes some historical background, definitions of terms and a review of research demonstrating that identifying multiracially is quantitatively different than identifying as a monoracial individual and these differences are meaningful within the context of educational testing.

Framework

Definitions and Historical Background

According to the U.S. Census Bureau (2017), race is a person’s social group(s) self-identification, while ethnicity refers to whether someone’s origin is Hispanic or Latino. Based

upon these definitions, a multiracial person would be someone who would identify with at least two racial social groups. These definitions indicate that the USCB has chosen to use social construction to define race. Others have proposed using biology to define race, but genetics does not support this definition (Sternberg et al., 2005). The social construction definition of race is what is used within this manuscript. It is important to note that race and ethnicity are not mutually exclusive, which makes research involving race somewhat challenging and it means that at times race and ethnicity may be used interchangeably within this manuscript.

It was not until 1997 that the USCB gave the option to individuals to mark more than one race, so the first U.S. census that included this option was conducted in the year 2000. Options for racial social groups on the 2000 U.S. Census were: “white”, “black or African American”, “American Indian and Alaska Native”, “Asian”, and “Native Hawaiian and Other Pacific Islander”. Prior to 2000 multiracial people had the option of checking a box labeled “Other” and writing in their racial identity. For statistical purposes, the USCB used whichever racial category was listed first in the “Other” category, assigning multiracial individuals to one race.

In 2010 the USCB had the following options for racial groups and allowed people to mark more than one box: “white”, “black, African American, or Negro”, “American Indian or Alaskan Native”, “Asian Indian”, “Japanese”, “Native Hawaiian”, “Chinese”, “Korean”, “Guamanian or Chamorro”, “Filipino”, “Vietnamese”, “Samoan”, “Other Asian”, “Other Pacific Islander”, and “Some other race”. The increase in options did not lead to less people identifying as Multiracial. In fact, the estimated multiracial population in the United States grew from 6.8 million to 9.0 million between the 2000 and 2010 U.S. Census (Jones & Bullock, 2012).

Similarly, students have had to self-identify according to racial groups for the purposes of educational testing. Educational testing within this review is defined as any kind of standardized

measurement of student behavior, knowledge, or ability (American Educational Research Association (AERA) et al., 2014), that is used within the field of education. Examples of educational tests are state accountability tests, nationally administered tests such as The Iowa Assessments (Riverside Insights, 2021), and college admissions exams such as The SAT (College Board, 2021) and the GRE (Graduate Record Examinations) (ETS, 2021). Assessments that are not standardized, such as classroom tests, are excluded from this definition of educational testing.

While the US Census is not a “test”, often whatever racial/ethnic categories are used by the USCB are frequently used by researchers, testing companies, government agencies, etc. Initially USCB data, and the lack of allowance for multiple races on USCB data, may seem to have no connection to the use of racial/ethnic identity on tests. However, due to the historical background of race and slavery in America, there are significant connections. One of the first connections can be made with the outcome of *Plessy vs. Ferguson* (1896), which essentially stated that racial segregation was constitutional if services provided were “equal”. Plessy, the plaintiff in the case, was Multiracial and was seven-eighths Caucasian and one-eighth Black. Based upon the court’s decision, someone who was Multiracial was in the end, assigned the minority race. While Plessy’s case had to do with which rail car he could sit in, based upon race, the outcome of this case led to continued segregation throughout the United States in many institutions, including education, which led to comparison of educational outcomes by race/ethnicity to demonstrate that the services offered were “equal”. Even after desegregation, educational outcomes continue to be assessed along racial/ethnic categories to assess concerns with fairness and equity.

Prior to allowing students to choose multiple racial/ethnic options or to selecting an option such as Two or More Races, students were forced to select one option. The practice of forcing multiracial students to choose one racial/ethnic category puts into question the validity of results that compare outcomes by race/ethnicity used in evaluating racial/ethnic group differences. Test scores are considered valid when there is sufficient evidence and theory to support the interpretation of scores for the intended purposes (AERA et al., 2014). Valid test scores are measuring what they are intended to measure, not other factors. Additionally, the practice of forcing multiracial students to choose one race raises questions about the reliability of test scores, in terms of consistent results across multiple test administrations (AERA et al., 2014), when interpreting racial/ethnic group mean scores. By forcing multiracial students to choose one race their test scores were included in race/ethnic group means that do not necessarily represent how they identify racially/ethnically. The guideline given in the *Standards for Educational and Psychological Testing* (AERA et al., 2014) states that when reporting scores for subgroups, such as by racial/ethnic categories, the subgroups are to be clearly defined and the construct measured needs to be comparable across groups. The practice of forced choice option for racial/ethnic identification seems to be in direct contradiction to the guideline.

Multiracial Identification

Research on the theory of racial identity first posited that multiracial individuals developed their identity through a “marginal man” model (Shih & Sanchez, 2005). The “marginal man” model states that those who are biracial exist on the margins of society as they live within two different worlds (Park, 1928). However, with time the “marginal man” model was seen as a negative approach to racial identity and new theories were posited in which multiracial identity was the same, or “equivalent” to monoracial identity development (Shih &

Sanchez, 2005). Following the “marginal man” model was the variant model which established that racial identification for multiracial individuals is uniquely different than that of monoracial individuals (Poston, 1990). In the variant approach multiracial individuals are thought to arrive at an integrated multiracial identity. Further research has countered the variant approach and has led to an ecological model (Rockquemore et al., 2009). The major assumptions of the ecological model are: 1) multiracial individuals are fluid in their racial identity based upon contextual factors, 2) the process of racial identity for multiracial individuals is not linear and is a lifelong process, 3) forcing multiracial individuals to identify as Multiracial or Monoracial perpetuates the same issues in terms of analysis of outcomes by racial/ethnic categories (Rockquemore et al., 2009).

The fluidity of racial identification among multiracial individuals based upon social context is supported by a review of multiracial identity literature conducted by Shih & Sanchez (2005). Factors such as the parental identification of race/ethnicity for their multiracial children (Campbell & Eggerling-Boeck, 2006), geographic location (Brunsma, 2006), changes in social groups over a lifespan (Hitlin et al., 2006), and peers (Renn, 2003), all affect how multiracial individuals identify racially/ethnically at any one time.

The complexity of the socialization of multiracial individuals can be seen by the mixed and somewhat contradictory results of a study conducted by Pearce-Morris and King (2012) that analyzed the well-being of multiethnic children (multiethnic was used within this study for people who identify as two or more races or ethnicities) compared to monoethnic children (referred to as same-ethnic in the study). When controlling for background and family characteristics, relationship stressors, and quality of parenting, multiethnic children were found to have a statistically significant negative affect compared to their monoethnic peers, yet there

were no differences found between the two groups in overall well-being, positive affect, or behavior problems. These results are particularly important as monoethnic children were used as a control group, which means that the higher negative affect for multiethnic children is significant beyond what is expected in the population. The authors were not able to explain these contradictory findings but noted that these results could be due to sample characteristics.

When analyzing young adults who identified as Asian/White in America and Canada it was found that those with an integrated view of their racial identity (equally positive feelings towards being Asian and White) had lower distress levels than those who identified primarily as Asian or White (Chong, 2012). Villegas-Gold and Tran (2018) also found significant positive relationships between an integrated view of racial identity, well-being, and self-esteem. Multiracial individuals who lived in families in which socialization celebrated the different racial/ethnic identities within the family demonstrated greater well-being than those whose families were less egalitarian in their socialization. But the relationship between an integrated racial identity and self-esteem was moderated by the physical appearance of multiracial individuals. Specifically multiracial individuals with more phenotypic racial ambiguity (physical features that make it harder for others to racially categorize someone) were found to have lower levels of self-esteem.

Differences between multiracial students and monoracial students in terms of academic achievement have seen mixed results in the research. Research by Herman (2009) found no difference in academic achievement between multiracial and monoracial students when accounting for the type of educational program within the school (college preparatory versus general education, vocational, English as a second language or business track) but Herman (2009) found that a student's multiracial racial identity has a statistically significant effect on

student's academic performance. Students who are Black/White or Hispanic/White and identify themselves as White had statistically significantly higher grades than those who identified themselves as Black or Hispanic. There was no statistically significant difference in achievement for Asian/White students who identified as Asian. These results were found while controlling for demographic, contextual, and prior achievement factors. The results indicate that there is an interaction between racial identity for multiracial youth who identify with the minority portion of their race and that their achievement is different from their peers. Phinney and Alipuria (1996) found no difference between multiracial youth and their monoracial peers in academic performance while in high school based on GPA. Harris and Thomas (2001) found similar results with the exception that Asian/White students had statistically significantly worse GPAs than White or Asian students, when controlling for school racial composition.

While research involving multiracial individuals is complex and there exists some contradictory findings, based upon the above research it is argued that multiracial individuals are different enough in their socialization compared to their monoracial peers that they should be treated differently when looking at group differences. Both increased distress levels and negative affect have been shown to decrease student scores on tests (Jones, 1994; Pretz, Totz, & Kaufman, 2010; Shang et al., 2013) and there are clearly contextual factors involving racial/ethnic identity that influence achievement of multiracial individuals. Additionally, multiracial youth who self-identify with a minority race may be impacted by stereotype threat (Herman, 2009). Stereotype threat occurs when a person of minority status feels anxious about confirming a negative stereotype and this feeling causes the test-taker to score lower than predicted (Steele & Aronson, 1995; Wicherts et al., 2005). The student is aware of the negative

stereotype which exists for her/his racial identity and the fear of reinforcing this stereotype by scoring low on the test leads to a lower test score than would be expected for that student.

Search Method

For this literature review an initial web search of the APAPsycInfo database was conducted using the following phrases “educational testing biracial and multiracial”, “testing multicultural students”, “student self-report of race ethnicity”, “educational testing ethnic minorities”, and “analysis by student race ethnicity”. Articles were then chosen to provide a general picture of what quantitative research has shown regarding multiracial individuals and educational testing. The focus of this article was on quantitative methods of measurement and therefore qualitative and mixed methods studies were not included in the review.

Based upon feedback provided by peer review further systematic searches were done of the literature. A search of the terms “achievement test*” OR “standardized test*” OR “educational test*” AND “biracial” OR “multiracial” OR “two or more races” resulted in 11 articles between 1968 – 2020 in the APA PsycInfo database and in six articles between 1975 – 2020 in the ERIC ProQuest database. None of these were relevant to this review. A further search of the terms "ACT" OR "SAT" OR "PSAT*" OR "PreACT" OR "Stanford Achievement" OR "Stanford 10" OR "Iowa Assessment*" OR "Iowa Test of Basic Skills" OR "GRE" OR "TOEFL" OR "Educational Records Bureau" OR "ERB" OR "CTP" OR "Terra Nova" OR "Comprehensive Testing Program" OR "California Assessment" OR "AP" AND "biracial" OR "multiracial" OR "two or more races" OR "multiple races" in ERIC ProQuest resulted in 19 articles and in APAPsycInfo in 39 articles. From the above only five articles were applicable to this review. Secondary searches through the websites of educational testing companies were then used to find specifics to individual tests. With the notable exceptions of the College Board, ACT

and ETS, most educational testing companies do not share any information regarding their tests. One can clearly see between the literature review and educational testing websites that there is limited research regarding educational testing and multiracial individuals.

While a few articles that individually address data collection, data coding, and/or data analysis of multiracial individuals could be found, none were found that address these topics collectively within educational testing. This review fills that gap in the literature. Often race and ethnicity are used interchangeably, which can lead to confusion. For brevity in writing, this review focuses on multiracial individuals versus multiethnic individuals. But, as noted above, the lines between the two are not clear. At times it may seem that the language used within the manuscript is contradictory in terms of race versus ethnicity and some of the terms used are offensive by today's standards. An effort was made to remain true to the content of the summarized research by using the language used within the cited research. When quoting racial/ethnic categories from an instrument, such as the U.S. Census, the terms are written within this manuscript as they are on the instrument. Otherwise, all race/ethnicity categories that are used as noun within the manuscript are capitalized, including Multiracial, to be uniform.

Review Results

This section is divided into five categories: data collection, data coding, data analysis,, implications, and recommendations. The discussion section includes potential impacts on policies, limitations of the review, as well as suggestions for further research.

Data Collection

Educational testing manuals reveal the lack of accounting for multiracial students when asking students to self-report their racial/ethnic identity. Only recently, in 2015-2016, did the SAT Suite of Assessments (the SAT, PSAT/NMSQT, PSAT 10, and PSAT 8/9 tests) add two

new options for Race/Ethnicity categories: Native Hawaiian or Pacific Islander and Two or More Races (Marini et al., 2019). A review of the *ACT® Technical Manual* (2020a) does not include multiracial individuals when reporting the demographics of test-takers until 2015-2016. The *SAT Suite of Assessments Technical Manual* (2017) does not include students of Two or More Races as the data came from the 2011 – 2014 cohort, which was prior to the inclusion of Two or More Races as an option. The Iowa Assessments (Riverside, 2012), a K-12 standardized achievement test, uses the 2010 U.S. Census ethnic and racial groupings by first having students indicate ethnicity as Hispanic or Non-Hispanic and then by selecting among racial categories which includes the option of Two or More Races. A test administration manual from 2010 for the Stanford Achievement Tests Series, Tenth Edition (a set of standardized, nationally normed achievement tests for students in grades K – 12) for the state of Arkansas gives directions for students to choose all racial/ethnic categories that apply (Arkansas Department of Education & NCS Pearson, 2010). It could not be determined when tests such as the Iowa Assessments and the Stanford achievement Tests started to have these kinds of options for students to select more than one racial/ethnic identity. As of the 2018 administration of the GRE, a standardized test used by many colleges as part of graduate admissions, there was not an option for GRE test takers to identify as two or more races/ethnicities (ETS, 2019). It is important to note that it could not be determined from the literature if GRE test takers are able to select more than one race/ethnicity.

Data Coding

Allowing test takers to select multiple racial/ethnic categories or to select an option which indicates they identify as multiple races raises several challenges for researchers if the goal is to compare outcomes across racial/ethnic categories. The practical issue in coding the

data of multiracial individuals is simply that individuals can only be counted once in the same study. This is one of the reasons why the USCB and previous studies assigned people one race. There is no easy solution to this dilemma. Options are to treat all multiracial individuals as their own group, or to further break the multiracial group down into respective parts so that someone who identifies as Black/White is in a different category than someone who identifies as Asian/Hispanic. A question this author has had is whether it is possible to use fractions to represent people, such as someone who identifies as Asian/White coded as .5 Asian and .5 White.

Support for having multiracial individuals be their own racial category is the commonality that they share, the factor of belonging to more than one racial social category and most likely being influenced by multiple cultural values. Research has shown that multiracial youth are more aware of the concept of race at an earlier age than monoracial youth, including monoracial youth of minority status (Kich, 1992, Root, 1992). Additionally, the fluidity of racial/ethnic identity among multiracial individuals (Doyle & Kao, 2007; Harris & Sim, 2001) lends itself to multiracial individuals being their own category. How a multiracial student identifies racially/ethnically depends upon a variety of factors such as the options available on the demographic question (Harris & Sim, 2001), the racial/ethnic background of the multiracial student (students who are partially Black tend to identify as Black) (Doyle & Kao, 2007; Phinney & Alipuria, 1996), and social context (Phinney & Alipuria, 1996). Multiracial students attending a predominately white college campus were found to identify as White more often than multiracial students attending a minority white college campus. This fluidity does not exist for monoracial students (Phinney & Alipuria, 1996).

The argument for breaking multiracial group data down into more granular data is based on the idea that there are still differences within multiracial people that are important. Research has shown there are different outcomes for members of different multiracial groups (Campbell, 2009; Charmaraman, Woo, Quach, & Erkut, 2014). Because race is a social construct the racial identity of multiracial students is influenced by their socialization experiences, which are affected by geography. For example, a multiracial student in Hawaii is going to have a different socialization experience than a multiracial student in Nebraska and those differences need to be taken into consideration (Herman, 2009). These differences exist for many reasons, one of which is simply that 24.0% of people in Hawaii identify as Multiracial while 2.3% of people in Nebraska identify as Multiracial (Jones & Smith, 2001). Lumping all multiracial individuals into one category does not account for these differences. These differences speak to the greater variability that exists within racial/ethnic groups than between them (Suzuki & Valencia, 1997). Due to the large variability within multiracial individuals and since they are a minority, studies that include multiracial identity as its own racial category need to be very careful in thinking through the approach to sampling (Root, 1992). Small subgroup population numbers tend to not properly measure group characteristics which means the results are not generalizable to the larger multicultural population (Goldstein & Morning, 2000; Herman, 2009).

The use of fractions to code for racial/ethnic categories will lead to needing to use more careful contrasts in coding to look at differences and this solution suffers from the same issue as all studies that compare differences by race, it assumes that race/ethnicity are mutually exclusive. Literature analyzing this approach was not found except for one study that looked at fractional assignment to analyze trends in race across the United States from 1990 to 2000 (Allen & Turner, 2001). Prior to 2000 it is impossible to know the number of people who identified as

Multiracial when using data from the USCB. Allen and Turner (2001) did find fractional assignments that could be used for some biracial individuals. But there are many limitations to their study. For example, they did not include Hispanics in the study, which constituted approximately 8% of the U.S. population at the time. Due to the struggle for many to distinguish between race and ethnicity, they could not find fractional assignments for biracial individuals who are part American Indian or Alaska Natives, and the study only applied to those who identify as Biracial. Another option is to use imputation methods that make use of a “primary race” assignment and covariates (Schenker & Parker, 2003). But this option is only possible if past research has some form of multiple race measurement, which most educational tests lack.

Data Analysis

In testing, between-group variability (differences in two or more groups) is compared to within-group variability (differences among members of the same group) to determine if the differences between two groups is significant. Concerns of fairness or equity arise if patterns of differences exist between test scores and variables for different groups (Camilli, 2006). No analyses of racial/ethnic comparisons could be found pertaining to K-12 achievement tests such as the Iowa Assessments or the Stanford Achievement Tests. The review revealed that often even if data on multiracial students was available, that data was not necessarily used in analyses of outcomes. When using data from 2015-2016 to compare student performance on the ACT by race/ethnicity the multiracial group is not included (as well as other racial groups such as American Indian/Native Alaskan, Native Hawaiian/Other Pacific Islander, etc.) (ACT, 2020a), yet these groups were included in the collection of demographic information within the same manual. The *ACT® Technical Manual* (2020a) does not state why these groups were not included in the analysis. ETS has a collection of studies which analyze GRE performance by

ethnic groups, not by racial groups (Wendler & Bridgeman, 2014). A search of the term ethnic throughout the collection reveals that none of the included studies had a category for multiple ethnicities.

More students identified as Two or More Races (N = 464) from a pilot test of the PreACT in May 2016 than as Asian (N = 364) and yet, students identifying as Two or More Races were not included in the racial/ethnic subgroup analyses of PreACT scores, while the smaller sample of students who identify as Asian was included. Obviously students who identify as Asian should be included in racial/ethnic subgroup analyses of tests scores. But it seems that it is also obvious that the data exists with a large enough sample size for multiracial students to be included in the subgroup analysis and yet, they were not.

Psychometrically a test is considered “fair” when individuals from different groups, of the same ability, have similar average item scores (Camilli et al., 2013). Psychometrics is the study of tests used to measure psychological constructs such as intelligence or achievement (Price, 2017). Tests would be considered “fair” if multiracial individuals of certain abilities had comparable scores to equal ability members from other racial groups. But to whom should we compare multiracial individuals? To the majority racial group, as we do minority groups? Or, should we compare them to the multiple racial groups to which they self-identify? And, what are the consequences of choosing one comparison over the other?

One can measure this concept of “fairness” by measuring what is known as differential item functioning (DIF). There are two kinds of DIF, content relevant DIF and content irrelevant DIF. Content relevant DIF occurs when the items on the test accurately distinguish between test takers on the construct of measure, such as differentiating between those with low and high math achievement. Content irrelevant DIF occurs when items on the test inappropriately distinguish

between test takers based on something other than the construct of measure, such as students of different race/ethnicities of similar ability having statistically different scores on an item (AERA et al., 2014). Presence of content relevant DIF and lack of content irrelevant DIF provide evidence towards test score validity. It is standard procedure for test makers of standardized tests to assess DIF across racial/ethnic categories in test preparation so that items that demonstrate content irrelevant DIF can be eliminated from the test. When measuring DIF the practice has been to use the majority group, White, as the reference group (Camilli, 2006) and to compare the minority group(s), the focal group(s), to the reference group. The *Fairness Report for the ACT Tests 2015-2016* (2017) did not include multiracial individuals as a category when measuring DIF.

Tests such as the ACT and the SAT are often analyzed for predictive validity of college success as they may be used in determining college admittance or scholarships. The ideal would be that scores on these tests predict college success equally for members of all racial/ethnic groups. In the recent manuscript in which ACT evaluates the validity of ACT scores in predicting college first-term grades multiracial individuals are included in a category labeled “multiple/unknown” (Radunzel & Mattern, 2020). A study on the predictive validity of the SAT found that students who identify as Multiracial perform better on the SAT (along with students who identify as White and Asian), than students from other racial/ethnic categories (Marini et al., 2019). SAT scores of students who identified as Two or More Races on the SAT in 2017 was found to “overpredict” those students first year college grade point averages (Marini et al., 2019), meaning that these students had slightly lower first year college grade point averages than expected given their scores on the SAT. An analysis of the effects of superscoring (students reporting the highest ACT subscores on college applications versus reporting scores from their

most recent ACT test) on subgroup differences in 2018 ACT scores included multiracial students as their own group (Mattern & Radunzel, 2019). The study found that use of superscores versus reporting the most recent ACT scores did not affect subgroup differences.

Helms (2008) analyzed the effects of a variety of different cut scores on eligibility, such as might be used for college acceptance, between Black and White test takers when taking within-group variability into account. Data from 31 different studies in which cognitive ability test scores were grouped by race were analyzed. It was found that the use of the same cut scores across race led to statistically significantly harder standards for Black than White test takers, when within-group variability was *not* taken into consideration. The results varied depending upon the cut scores used, but regardless of which cut scores were used, the results indicate that the use of cut scores, which do *not* take into consideration within-group variability, favors White test takers. While this study only looks at Black and White test takers, it speaks to the need to include within-group variability when categorizing by race. The inclusion of within-group variability is even more important for those who are Multiracial as their within-group variability is even larger (Suzuki & Valencia, 1997).

Implications

To have valid test interpretations, educational tests are given using standardized methods and procedures. When taking standardized educational tests students are asked at the beginning stages of the test to answer demographic questions such as birth date, gender, and race/ethnicity. While the demographic questions are standardized and test administrators may use the same phrases to guide students in answering the demographic questions, the actual act of having to answer questions about race/ethnicity may impact a multiracial test taker's score on the test due to anxiety about having to choose one race (Chong, 2012; Shih & Sanchez, 2005). The actual

wording of the race/ethnicity identification question may impact the score on the test for a multiracial individual as research has shown that multiracial individuals take issue with having to identify as “other” (Johnson et al., 1997; Panter et al., 2009). In other words, multiracial students may feel as if they are being treated differently than monoracial students solely due to the demographic questions. Townsend et al. (2009) found that multiracial students forced to select one race had statistically significantly lower scores on performance self-esteem, feelings of agency, feelings of self-efficacy and motivation than those who could check all possible racial categorizations that applied. Aside from the negative psychological impact of having to select one race, lower values for agency, self-efficacy, and motivation have been shown to be related to lower achievement test scores (Bruning et al., 2011).

The most important implication is simply that any comparison along racial/ethnic categories should be questioned. When conducting research, one assumes that the groups being compared are mutually exclusive. Race and ethnicity are not mutually exclusive. And yet, it should also be noted that research which makes comparisons between racial/ethnic categories is valuable to assess important concerns such as fairness and equity (AERA et al., 2014).

Marini et al., 2019 states that due to the inclusion of Two or More Races on the suite of SAT assessments in 2016, the composition of all racial groups has changed which makes any comparisons to earlier SAT research inaccurate. The concern raised by Marini applies to all educational research that includes data in which multiracial students were forced to choose one racial/ethnic category including longitudinal comparisons by race/ethnicity that include data from earlier dates. The sheer number of educational studies for which this applies is simply astounding. The sample sizes for each racial/ethnic category are not accurate as they do not

reflect students who would have selected more than one category or the option of Two or More Races, had those options been available.

Some may argue that the number of multiracial students is so small that the changes in sample sizes are negligible. That may have been true at one time point, however recent figures indicate that the argument of being a negligible amount is no longer warranted. The PreACT, a nationally standardized test that is intended for students in the 10th grade to predict how students would perform on the ACT exam, does not include multiracial individuals in its listing of fall demographics (ACT, 2020b). But 5.6% of the Fall testing sample are listed as Other and 9.3% are listed as Missing. Surely, these two categories contain a portion of multiracial individuals. PreACT 2018-2019 Operational Test Data reveal that more students identified as Two or More Races (N = 36,466) than Asian (N = 26,417) (ACT, 2020b). Approximately 4.4% of all ACT test-takers in 2018 identified as Multiracial (Mattern & Radunzel, 2019). A study of test validity for the SAT that was conducted in 2017 lists 4% of the sample identifying as Two or More Races and lists 3% of SAT test-takers who were graduating seniors as Two or More Races (Marini et al., 2019). One cannot pick a date to determine when the number of multiracial individuals was large enough to make a difference in sample sizes. In addition, the effects of changes in subgroup sample sizes on study outcomes depends upon the overall studies initial sample size as well as subgroup sample sizes and the portions of each subsample to the overall sample size.

Recommendations

A study that compared the responses of multiracial people to different types of questions regarding their racial makeup found that multiracial people identified differently depending upon the type of question used, even when all question types allowed for the possibility of sharing that they were multiracial (Johnson et al., 1997). Multiracial individuals preferred questions about

race/ethnicity in this order: questions that allowed them to state that they are Multiracial, a forced-choice option in which they could choose “not classifiable”, followed by a forced-choice option with “other” as an option. Respondents noted that the term “other” implies that you do not belong to a group.

More recently, a study conducted with the purpose of informing changes to the demographic questions on the GRE found that multiracial respondents most desired the opportunity to select more than one race (Klieger et al., 2013). Somewhat similarly to the finding by Johnson et al., (1997), some respondents in the more recent study expressed discomfort with an option of “Some Other Race” as “Some” was seen as pejorative. Respondents indicated that alphabetizing the racial/ethnic categories was seen as more neutral compared to what has primarily done in practice, which is to list White as the first option. The Urban Institute recently put out some guidelines for collecting racial/ethnic data (Schwabish & Feng, 2021) that advise in place of the word “Other” one could use “Another Race”, which is more inclusive language

In terms of data coding, literature regarding other coding alternatives to distinguish between racial/ethnic groups could not be found. But some have suggested using additional measures to inform how to categorize students. Helms (2006) proposes the use of psychological constructs that deal with race/ethnicity in place of current racial grouping practices to account for the interaction between racial identity and test scores. If this approach is used then fairness is not test dependent, but rather is sample dependent. According to the joint American Educational Research Association (AERA), American Psychological Association (APA) and National Council on Measurement in Education (NCME) 2014 Standards, fairness is interpreted as “responsiveness to individual characteristics and testing contexts so that test scores will yield valid interpretations for intended uses.” (p. 50). Helms (2006) defines test fairness as “the

removal from test scores of systematic variance, attributable to the test takers' psychological characteristics, developed in response to socialization practices or environmental conditions, that are "irrelevant to measurement of the construct of interest to the test user or assessor" (p. 847).

The use of racial/ethnic psychological constructs has merit and should be further explored.

However, it also means that test takers would have to take more than one assessment as they would need to be tested on achievement and on the psychological constructs of race/ethnicity.

Somewhat similarly, Sternberg, Grigorenko, and Kidd (2005) propose using multiple measures or factors that account for culture, in addition to test scores.

Of course, the challenge then becomes how to measure the experiences of multiracial individuals on a psychological measure. Salahuddin and O'Brien (2011) created the Multiracial Challenge and Resilience Scale (MCRS). The scale was developed using a sample of 317 self-identified multiracial individuals ranging in age from 18 – 53 years who lived in large metropolitan areas. Individuals in the sample were given accepted validated scales that measure self-esteem, depression, social connectedness, ethnic identity, and racial encounters in addition to the developed MCRS. The initial MCRS scale showed convergent validity with these validated measures. This first study resulted in a factor structure comprised of 1) Others' Surprise and Disbelief Regarding Racial Heritage, 2) Lack of Family Acceptance, 3) Multiracial Discrimination, 4) Challenges with Racial Identity, 5) Appreciation of Human Differences, and 6) Multiracial Pride. A measure of test-retest reliability found all measures reliable over time except Lack of Family Acceptance scores. Researchers could not account for the lack of reliability over time for Lack of Family Acceptance but thought that it might be due to participants feelings changing over time based upon recency, or lack of recency, of family interactions.

A second study conducted to assess the internal validity and reliability of the MCRS scale and the stability of the factor structure found that the MCRS demonstrated good internal consistency reliability for all factors except for Challenges with Racial Identity (Salahuddin & O'Brien, 2011). Further study of this scale is warranted as the researchers note that the majority of those who participated in both studies were women who had higher education levels and income than the general population. But the MCRS is certainly a step in the right direction in terms of being able to measure the socialization experience of multiracial individuals.

There really are few analyses that include multiracial individuals, and it seems that the biggest recommendation is for more research involving multiracial individuals. Because of the large within-group variability that exists within the multiracial group one of the recommendations made is to use other methods such as qualitative. When measuring test fairness Camilli (2006) states the value in using qualitative data through processes such as sensitivity reviews to assess cultural impact as well as DIF analysis. Sensitivity reviews are procedures involved in test development that ensure that test materials reflect the diversity within society and that the wording of test items do not offend different subgroups within society. The author recommends using mixed methods research, which integrates quantitative and qualitative data (Plano Clark & Ivankova, 2016), as it seems it would be a particularly useful way to include test performance information of racial/ethnic groups that may be too small in sample size to include within the larger analyses of group results.

Okazaki and Sue (1995) created a list of recommendations for any analysis that includes minorities, including Multiracial individuals. They recommend researchers 1) make explicit the definitions of ethnicity and race being used within the study, 2) provide more detailed descriptions of the sampling method and sample obtained, 3) include enough sample

information, such as comparisons of ethnic differences on measured items, so that comparisons can be made between studies, 4) use multiple measures and multiple methods, 5) use expert cultural or ethnic consultants both prior to conducting the study and to assist in interpretation of the study, and 6) use caution in assuming the findings are valid, but instead have the findings lead to more hypotheses for testing. Charmaraman et al. (2014) make similar suggestions and include the guideline of breaking the multiracial group into more granular data if sufficient sample size is available and if it serves the purpose of the study.

In 2003 the American Psychological Association (APA) came out with guidelines pertaining to the research of multiracial people and implications for the field of psychology. The first guideline is that psychologists need to be aware of their own biases. This guideline applies to all research and should be taken into consideration when testing multiracial individuals. Secondly, psychologists are encouraged to be aware of the importance of recognizing multicultural differences. Ignoring the uniqueness of multiracial individuals when testing and when assessing the validity of test scores does not recognize this uniqueness. The APA also recommends that psychological research should recognize the importance of research being culture-centered.

DIF analysis has tended to use dummy codes (0s and 1s) to indicate membership in the reference group or the focal group. Mayhew and Simonoff (2015) recommend using effect code (group average responses) as opposed to dummy codes to address issues of test fairness regarding interpretation of test scores. The language used when interpreting results from a study that uses dummy coding for racial categorization tends to have the dominate, White, race as the norm. Effect coding would remove the language of a “normed” race and instead replace the interpretation with language that would speak to the expected outcome of each racial group

compared to an “overall level”. This has the added benefit of being able to include the expected outcome of the dominant race compared to an overall level, something that is missing when researchers use dummy codes. Another option, particularly when evaluating multiracial youth, is to use monoracial minority youth as the control group (Cauce et al., 1992), as opposed to monoracial white youth. Meaning, a student who is Black and White would be compared to Blacks as the control group, not White. The question this leads to is what to do with students who identify with more than one racial/ethnic minority category?

Discussion

Based upon the review of literature it seems that while race is hard to define, using a social construct approach makes the most sense. After all, it is in the socialization process that children learn about the construct of race. In the U.S., multiracial individuals face unique socialization practices compared to their monoracial peers. In particular, the act of having to select one race prior to the 2000 census, instead of being able to select all races that apply, is just one example of an institutional practice that affects the socialization of multiracial individuals. Research is clear that formation of identity is different for multiracial individuals than monoracial individuals and these differences have to do with socialization practices.

The use of contextual information when interpreting standardized test scores by subgroups is recommended (AERA et al., 2014). Interactions between a test taker’s internalized racial or cultural experience, their environmental experience, and the test or test taking process (Herman, 2009; Sternberg et al., 2005) affect test scores. Racial identification for multiracial children leads to a unique interaction between their racial identity and scores on tests as they most likely have more than one set of cultural values influencing them (Herman, 2009; Pearce-Morris & King, 2012). Rockquemore et al., (2009) provides a very thoughtful review of the

challenges in creating a multiracial identity theory due to the multiplicity of cultural influences on racial identification for multiracial individuals.

For purposes that many would deem good or bad, group differences continue to be analyzed in educational testing. Some would deem analysis of group differences good as they aid in ensuring that groups are being treated fairly during test administration and in assessing fair impact in use of test scores. Comparing group differences helps in assessing the validity and reliability of results. However, the consequences of comparing group differences can also be interpreted as bad when group membership is somewhat vague in its construct and when individuals do not neatly fit into the provided categories. The latter examples apply to the testing of multiracial individuals and comparing people based upon racial/ethnic categories.

One option is to simply do away with racial categorization (Mays et al., 2003). However, to do this ignores that there are racial group differences in scores which impact opportunities differently for members of different races. Therefore, this option does not currently seem tenable. It seems the best option is to continue to use racial grouping when appropriate in comparing group scores. But changes should be made that more accurately reflect the uniqueness of multiracial individuals, including, but not limited to: allowing individuals to check more than one race when applicable on all demographic information, provide as much clarity as possible regarding how the construct of race and ethnicity is defined and used within a study, analyze the data using effect coding with multiracial people as their own group, if possible break the multiracial group into more granular data, and use measures of the psychological construct of racial identity, such as the MCRS, along with other measures in order to capture the interaction between racial identity and the measures. Guidelines for related topics such as language when writing analysis results and data visualization exist (Schwabish & Feng, 2021).

While the use of standardized test scores such as the ACT and SAT for purposes of college enrollment are currently declining, they are still used by many colleges and if not used for college admittance they are often used for scholarships. Because of the educational benefits a diverse student population provides, race may be taken into consideration by colleges when making admittance decisions (*Grutter vs Bollinger*, 2003). There are no guidelines on how multiracial students should be operationalized within this context (Sanchez et al., 2020). Should they be identified as a minority? Should multiracial students who are part White be categorized with monoracial students of minority status in scholarship opportunities? Would this be seen as an act of taking away a scholarship from someone who is a monoracial minority? Sanchez et al., (2020) argues for established policies in this area. However, the researchers also express concern based upon how multiracial individuals are categorized within the legal system (they tend to be categorized according to their minority status, regardless of how they identify).

The varied results of the included research can be attributed to many things. Charmaraman et al. (2014) analyzed the research methods used in 133 studies of multiracial individuals from 1990 – 2009. They found that multiracial study samples were more likely to contain people from the West, those who identify as Black/White, individuals from college-educated households, women, and young adults or adolescents. Additionally, it was found that most studies forced multiracial participants into one racial category. All of which may lead to misleading conclusions.

There is still quite a bit of research that can and should be done regarding testing and multiracial individuals. We are just beginning to have access to data in which we can measure trends that include multiracial identity. It is understandable that the major testing organizations such as the ACT and the College Board, need to wait for enough data which accurately measured

multiracial individuals before multiracial individuals could be included in their analyses. It is hoped and expected that since both the ACT and the SAT started collecting data on multiracial individuals in 2015-2016 more recent research will be forthcoming. Research regarding how to statistically account for multiracial individuals is needed. There are many studies that compare monoracial individuals to one another that should now be replicated with multiracial individuals. Lastly, measuring the impacts of certain statistical practices on the use of test scores for multiracial individual is lacking.

One of the limitations of this review is that it does not address both qualitative and mixed methods analyses that may be important in measuring multiracial individuals. The limitation was one that the author deliberately chose as analyses which involve group comparisons are quantitative. Therefore, the focus was on quantitative research. While there were recommendations to include more qualitative data when interpreting test scores of multiracial individuals, it was surprising to see that there were no recommendations to use more mixed methods within the literature as the earlier recommendation came from the author. Mixed methods has the distinct advantage of integrating both quantitative and qualitative data, which allows for greater representation of members from both large and small subgroups within the testing population.

This review was written from a primarily psychometric view and therefore focuses on measurement. However, the real importance and value of a review on educational testing and multiracial individuals is only realized when one connects the measures to actual people and understands that how one measures psychological constructs and how one uses those measures affect people. Multiracial individuals should not have to identify as other, or lesser, or fractional.

They are whole people, and measurement practices need to better reflect their lived and experienced reality as measurement practices affect those lived experiences.

References

ACT. (2017). *Fairness Report for the ACT Tests*.

https://www.act.org/content/dam/act/unsecured/documents/2015-2016_Fairness_Report_for_the_ACT_Tests_WebSecured.pdf

ACT. (2020a). *ACT® Technical Manual*.

https://www.act.org/content/dam/act/unsecured/documents/ACT_Technical_Manual.pdf

ACT. (2020b.) *PreACT® Technical Manual*.

https://www.act.org/content/dam/act/unsecured/documents/PreACT_Technical_Manual_2020.pdf

Allen, J. P., & Turner, E. (2001). Bridging 1990 and 2000 census race data: Fractional assignment of multiracial populations. *Population Research and Policy Review*, 20, 513-533. <https://link-springer-com.libproxy.unl.edu/article/10.1023%2FA%3A1015666321798>

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). Joint Committee on Standards for Educational and Psychological Testing. *Standards for Educational and Psychological Testing*. Washington, D. C.: AERA

American Psychological Association (APA). (2003). *Guidelines on multicultural education, training, research, practice, and organizational change for Psychologists*.

<http://doi.apa.org/getdoi.cfm?doi=10.1037/0003-066X.58.5.377>

AP News (1997, April 22). Tiger Woods describes himself as ‘Cablinasian’. *AP News*.

<https://www.apnews.com/458b7710858579281e0f1b73be0da618>

- Arkansas Department of Education, NCS Pearson. (2010). *Test Administration Manual. Stanford Achievement Test Series, Tenth Edition*. <https://pdf4pro.com/view/test-administration-manual-state-board-of-education-34b652.html>
- Berry, C. M., Clark, M. A., & McClure, T. K. (2011). Racial/ethnic differences in the criterion-related validity of cognitive ability tests: A qualitative and quantitative review. *Journal of Applied Psychology, 96*(5), 881-906. <https://doi.org/10.1037/a0023222>
- Bruning, R. H., Schraw, G. J., & Norby, M. M. (2011). *Cognitive psychology and instruction*. Boston, MA: Allyn & Bacon.
- Brunsma, D. L. (2006). Public categories, private identities: Exploring regional differences in the biracial experience. *Social Science Research, 35*(3), 555–576. <https://doi.org/10.1016/j.ssresearch.2004.10.002>
- Camilli, G. (2006). Test fairness. In R. Brennan (Ed.), *Educational measurement* (4th ed., pp. 221-256). Westport, CT: American Council on Education and Praeger.
- Camilli, G., Briggs, D. C., Sloane, F. C., Chiu, T. (2013). Psychometric perspectives on test fairness: Shrinkage estimation. In K. F. Geisinger (Ed.-in-Chief), B. A. Bracken, J. F. Carlson, J. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Assoc. Eds.), *APA handbooks in psychology: APA handbook of testing and assessment in psychology: Volume 3. Testing and Assessment in school psychology and education*. Washington, D.C.: American Psychological Association
- Campbell, M. E. (2009). Multiracial groups and educational inequality: A rainbow or a divide? *Social Problems, 56*(3), 425-446. <https://doi.org/10.1525/sp.2009.56.3.425>

- Campbell, M. E., & Eggerling-Boeck, J. (2006). “What about the children?” The psychological and social well-being of multiracial adolescents. *The Sociological Quarterly*, 47(1), 147–173. <https://doi.org/10.1111/j.1533-8525.2006.00041.x>
- Cauce, A. M., Yumi, H., Mason, C., Aquilar, T., Ordonez, N., & Gonzales, N. (1992). Between a rock and a hard place: Social adjustment of biracial youth. In M. P. P. Root (Ed.), *Racially mixed people in America* (pp. 207-222). Newbury Park, CA: SAGE Publications, Inc.
- Charmaraman, L., Woo, M., Quach, A., & Erkut, S. (2014). How have researchers studied multiracial populations? A content and methodological review of 20 years of research. *Cultural Diversity and Ethnic Minority Psychology*, 20(3), 336–352. <https://doi.org/10.1037/a0035437>
- Chong, V. (2012). *Racial identity, family, and psychological adjustment in Asian-White biracial young adults* [Doctoral dissertation, University of Windsor]. Electronic Theses and Dissertations, 4797. <https://scholar.uwindsor.ca/etd/4797>
- College Board. (2021). *The SAT*. <https://collegereadiness.collegeboard.org/sat>
- Doyle, J. M., & Kao, G. (2007). Are racial identities of multiracials stable? Changing self-identification among single and multiple race individuals. *Social Psychology Quarterly*, 70(4), 405-423. <https://www.jstor.org/stable/20141804>
- ETS. (2019). *A snapshot of the individuals who took the GRE® general test*. https://www.ets.org/s/gre/pdf/snapshot_test_taker_data_2019.pdf
- ETS. (2021). *GRE® Guide to the Use of Scores*. https://www.ets.org/s/gre/pdf/gre_guide.pdf

- Goldstein, J.R. & Morning, A. J. (2000). The multiple-race population of the United States: Issues and estimates. *Proceedings of the National Academy of Sciences*, 97(11), 6230-6235. <https://doi.org/10.1073/pnas.100086897>
- Grutter v. Bollinger. (2003). *Oyez*. <https://www.oyez.org/cases/2002/02-241>
- Harris, D. R., & Sim, J. J. (2001). *An empirical look at the social construction of race: The case of multiracial adolescents*. [Research Report 39]. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.22.6942&rep=rep1&type=pdf>
- Harris, D. R. & Thomas, J. L. (2001, March). *The educational costs of being multiracial: Evidence from a national survey of adolescents*. [Presentation] Population Association of America, Washington, D.C.
- Helms, J. E. (2006). Fairness is not validity or cultural bias in racial-group assessment: A quantitative perspective. *American Psychologist*, 61(8), 845-859. <https://doi.org.10.1037/0003-066X.61.8.845>
- Helms, J. E. (2008). Implications for social policy of variability in racial groups' test scores: How cut scores on tests of cognitive abilities, knowledge, or skills matter. *American Psychologist*, 63(8), 721-739. <http://web.b.ebscohost.com.libproxy.unl.edu/ehost/pdfviewer/pdfviewer?vid=1&sid=3f98e657-2a95-44ce-9568-3974023d7155%40pdc-v-sessmgr01>
- Herman, M. R. (2009). The Black-White-other achievement gap: Testing theories of academic performance. *Sociology of Education*, 82(1), 20-46. <https://www.jstor.org/stable/40376036>

- Hitlin, S., Scott Brown, J., & Elder, G. H. (2006). Racial self-categorization in adolescence: Multiracial development and social pathways. *Child Development, 77*(5), 1298–1308. <https://doi.org/10.1111/j.1467-8624.2006.00935.x>
- Jones, M. C. (1994). *Life stress and reading comprehension test scores in the middle school student*. [Master's Thesis]. Kean College. <https://files.eric.ed.gov/fulltext/ED369057.pdf>
- Jones, N. A., & Smith, A. S. (2001). *The two or more races population: 2000*. U.S. Census Bureau. <https://www.census.gov/prod/2001pubs/c2kbr01-6.pdf>
- Jones, N. A., & Bullock, J. (2012). *The two or more races population: 2010*. U.S. Census Bureau. <https://www.census.gov/prod/cen2010/briefs/c2010br-13.pdf>
- Kich, G. K. (1992). The developmental process of asserting a biracial, bicultural identity. In M. P. P. Root (Ed.) *Racially mixed people in America* (pp. 304-317). Newbury Park, CA: SAGE Publications, Inc.
- Mattern, K., & Radunzel, J. (2019). *Does superscoring increase subgroup differences*. <https://www.act.org/content/dam/act/unsecured/documents/R1774-superscoring-subgroup-2019-07.pdf>
- Mayhew, M. J., & Simonoff, J. S. (2015). Non-White, no more: Effect coding as an alternative to dummy coding with implications for higher education researchers. *Journal of College Student Development, 56*(2), 170–175. <https://doi.org/10.1353/csd.2015.0019>
- Mays, V. M., Ponce, N. A., Washington, D. L., & Cochran, S. D. (2003). Classification of race and ethnicity: Implications for public health. *Annual Review of Public Health, 24*(1), 83–110. <https://doi.org/10.1146/annurev.publhealth.24.100901.140927>

- Okazaki, S., & Sue, S. (1995). Methodological issues in assessment research with ethnic minorities. *Psychological Assessment*, 7(3), 367-375.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.905.4003&rep=rep1&type=pdf>
- Panter, A. T., Daye, C. E., Allen, W. R., Wightman, L. F., & Deo, M. E. (2009). It matters how and when you ask: Self-reported race/ethnicity of incoming law students. *Cultural Diversity and Ethnic Minority Psychology*, 15(1), 51-66.
<https://doi.org/10.1037/a0013377>
- Park, R. E. (1928). Human migration and the marginal man. *American Journal of Sociology*, 33(6), 881-893. <https://www.jstor.org/stable/2765982>
- Pearce-Morris, J. & King, V. (2012). The well-Being of children living with interethnic parents: Are they at a disadvantage? *Journal of Family Issues*, 33(7), 898-919.
<https://doi.org.10.1177/0192513X11420938>
- Phinney, J. S. & Alipuria, L. L. (1996). At the interface of cultures: Multiethnic/multiracial high school and college students. *The Journal of Social Psychology*, 136(2), 139-158.
<http://web.a.ebscohost.com.libproxy.unl.edu/ehost/pdfviewer/pdfviewer?vid=1&sid=4842795e-deca-4cec-895f-0ff3ec19232b%40sdc-v-sessmgr01>
- Plano Clark, V., & Ivankova, N. V. (2016). *Mixed methods research: A guide to the field*. SAGE Publications, Inc.
- Plessy vs. Ferguson, Judgement, Decided May 18, 1896; Records of the Supreme Court of the United States; Record Group 267; *Plessy v. Ferguson*, 163, #15248, National Archives.
<https://www.ourdocuments.gov/doc.php?flash=false&doc=52>
- Poston, W. S. C. (1990). The biracial identity development model: A needed addition. *Journal of Counseling & Development*, 69, 152-155.

<http://web.a.ebscohost.com.libproxy.unl.edu/ehost/pdfviewer/pdfviewer?vid=1&sid=4b429052-7a54-44ce-aba7-c8c2cbf17cba%40sessionmgr4007>

Pretz, J. E., Tetz, K. S., & Kaufman, S. B. (2009). The effects of mood, cognitive style, and cognitive ability on implicit learning. *Learning and Individual Differences, 20*, 215-219.

[https://www-sciencedirect-](https://www-sciencedirect-com.libproxy.unl.edu/science/article/pii/S1041608009001095?via%3Dihub)

[com.libproxy.unl.edu/science/article/pii/S1041608009001095?via%3Dihub](https://www-sciencedirect-com.libproxy.unl.edu/science/article/pii/S1041608009001095?via%3Dihub)

Price, L. R. (2017). *Psychometric methods: Theory into practice*. The Guilford Press.

Radunzel, J., & Mattern, K. (2020). *A case study: ACT section retest scores and superscores are predictive of first-term grades*. ACT.

<https://www.act.org/content/dam/act/unsecured/documents/R1817-he-section-retest-2020-05.pdf>

Riverside Insights. (2021). *Iowa assessments*. <https://info.riversideinsights.com/iowa-assessments>

Riverside. (2012). *Iowa assessments Form E technical summary*. https://central-rivers-juiceboxinteract.netdna-ssl.com/wp-content/uploads/2017/03/H5_Iowa-Assessments-Form-E-Technical-Summary.pdf

Root, M. P. P. (Ed.) (1992). *Racially mixed people in America*. Newbury Park, CA: SAGE Publications.

Johnson, T. P., Jobe, J. B., O'Rourke, D., Sudman, S., Warnecke, R. B., Chávez, N., Chapa-Resendez, G., & Golden, P. (1997). Dimensions of self-identification among multiracial and multiethnic respondents in survey interviews. *Evaluation Review, 21*(6), 671-687.

<https://doi.org/10.1177/0193841X9702100602>

- Klieger, D. M., Adler, R., & Ezzo, C. (2013). Asking differently About race and ethnicity: New needs for a changing population. *ETS Research Report Series*, 2013(2), i–59.
<https://doi.org/10.1002/j.2333-8504.2013.tb02344.x>
- Marini, J., Westrick, P. A., Young, L., Ng, H., Shmueli, D., & Shaw, E. J. (2019). *Differential validity and prediction of the SAT®*. <https://files.eric.ed.gov/fulltext/ED597325.pdf>
- Renn, K. A. (2003). Understanding the identities of mixed-race college students through a developmental ecology lens. *Journal of College Student Development*, 44(3), 383–403.
<https://doi.org/10.1353/csd.2003.0032>
- Rockquemore, K. A., Brunnsma, D. L., & Delgado, D. J. (2009). Racing to theory or retheorizing race? Understanding the struggle to build a multiracial identity theory. *Journal of Social Issues*, 65(1), 13–34. <https://doi.org/10.1111/j.1540-4560.2008.01585.x>
- Salahuddin, N. M., & O'Brien, K. M. (2011). Challenges and resilience in the lives of urban, multiracial adults: An instrument development study. *Journal of Counseling Psychology*, 58(4), 494–507. <https://doi.org/10.1037/a0024633>
- Sanchez, D. T., Gaither, S. E., Albuja, A. F., & Eddy, Z. (2020). How policies can address multiracial stigma. *Policy Insights from the Behavioral and Brain Sciences*, 7(2), 115–122. <https://doi.org/10.1177/2372732220943906>
- Schenker, N., & Parker, J. D. (2003). From single-race reporting to multiple-race reporting: Using imputation methods to bridge the transition. *Statistics in Medicine*, 22(9), 1571–1587. <https://doi.org/10.1002/sim.1512>
- Shang, J., Fu, Q., Dienes, Z., Shao, C., & Fu, X. (2013). Negative affect reduces performance in implicit sequence learning. *PLoS ONE*, 8(1), e54693.
<https://doi.org/10.1371/journal.pone.0054693>

- Shih, M., & Sanchez, D. T. (2005). Perspectives and research on the positive and negative implications of having multiple racial identities. *Psychological Bulletin*, 131(4), 569–591. <https://doi.org/10.1037/0033-2909.131.4.569>
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811. <https://doi.org/10.1037//0022-3514.69.5.797>
- Sternberg, R. J., Grigorenko, E. L., & Kidd, K. K. (2005). Intelligence, race, and genetics. *American Psychologist*, 60(1), 46–59. <https://doi.org/10.1037/0003-066X.60.1.46>
- Suzuki, L. A., & Valencia, R. R. (1997). Race-ethnicity and measured intelligence. *American Psychologist*, 52(10), 1103-1114. <http://web.a.ebscohost.com.libproxy.unl.edu/ehost/pdfviewer/pdfviewer?vid=1&sid=6f845a07-bfb7-48a7-aec0-08fa3702f130%40sessionmgr4008>
- Townsend, S. S. M., Markus, H. R., & Bergsieker, H. B. (2009). My choice, your categories: The Denial of multiracial identities. *Journal of Social Issues*, 65(1), 185-204. <Rhttps://doi.org/10.1111/j.1540-4560.2008.01594.x>
- U.S. Census Bureau. (2017). *Race & ethnicity*. <https://www.census.gov/mso/www/training/pdf/race-ethnicity-onepager.pdf>
- Villegas-Gold, R., Tran, A. G. T. T. (2018). Socialization and well-being in multiracial individuals: A moderated mediation model of racial ambiguity and identity. *Journal of Counseling Psychology*, 65(4), 413-422. <https://doi.org/10.1037/cou0000277>
- Wendler, C., & Bridgeman, B. (2021). *The research foundation for the GRE® revised general test: A compendium of studies*. ETS. https://www.ets.org/s/research/pdf/gre_compendium.pdf

Wicherts, J. M., Dolan, C. V., & Hessen, D. J. (2005). Stereotype threat and group differences in test performance: A question of measurement invariance. *Journal of Personality and Social Psychology*, 89(5), 696-716.

<http://web.b.ebscohost.com.libproxy.unl.edu/ehost/pdfviewer/pdfviewer?vid=1&sid=74d9f7fd-f339-4b74-ba0c-312164224560%40pdc-v-sessmgr03>