

---

# The Meaning of Race in Psychology and How to Change It

---

## *A Methodological Perspective*

---

Janet E. Helms, Maryam Jernigan, and Jackquelyn Mascher  
*Boston College*

---

*The primary purpose of this article was to offer a methodological critique in support of arguments that racial categories should be replaced as explanatory constructs in psychological research and theory. To accomplish this goal, the authors (a) summarized arguments for why racial categories should be replaced; (b) used principles of the scientific method to show that racial categories lack conceptual meaning; (c) identified common errors in researchers' measurement, statistical analyses, and interpretation of racial categories as independent variables; and (d) used hierarchical regression analysis to illustrate a strategy for replacing racial categories in research designs with conceptual variables. Implications for changing the study of race in psychology are discussed.*

**R**ace has no consensual theoretical or scientific meaning in psychology, although it is frequently used in psychological theory, research, and practice as if it has obvious meaning (cf. Yee, Fairchild, Weizmann, & Wyatt, 1993; Zuberi, 2001). Some psychologists contend that race refers to biological characteristics of individuals as reflected in their physical appearance (e.g., Rowe, 2002), some argue that it is a pseudonym for impoverished backgrounds (Eisenman, 1995), whereas others assert that race is a social construction that maintains a sociopolitical hierarchy in U.S. American society (e.g., American Psychological Association, 2003; Helms, 1994).

In fact, because race lacks precise meaning, various psychologists have long challenged the scientific merit of studying or using race as an explanatory construct in psychological theory, research, and, by implication, practice (Phinney, 1996; Yee et al., 1993). In fact, according to Yee (1983), the Council of Representatives of the American Psychological Association passed a resolution more than two decades ago opposing the use of the concept of race to explain human behavior. This resolution was deemed necessary because of society's continued reification of folk definitions of race in spite of considerable substantive advice to the contrary offered by nonpsychological professional and scientific organizations. The rationale for the original resolution was elaborated on in the "Guidelines on Multicultural Education, Training, Research, Practice, and Organizational Change for Psychologists" (American Psy-

chological Association, 2003). Yet neither the original resolution nor the recently adopted "Guidelines" appear to have had much effect on the study or conceptualization of race in psychology.

Instead, psychologists have ignored the fact that race itself has no shared conceptual definition by tacitly agreeing to use factitious racial categories (e.g., Black and White) as independent or predictor variables in their theories and research designs as if the categories convey whatever conceptual meaning of race the researcher intends. Equating race with racial categories gives scientific legitimacy to the conceptually meaningless construct of race, thereby perpetuating racial stereotypes and associated problems in society. Moreover, it permits the discipline of psychology to function as an "objective" science even though it has granted a conceptually meaningless concept (i.e., race) so central a role in its theory, research, and practice (Fairchild, 1991; Zuberi, 2001).

A common theme, albeit implicit, in the previously cited critiques of the manner in which race has been investigated in psychology has been that racial categories ought to be replaced as independent variables in psychological theory and research. The following four strategies for accomplishing the recommended replacement can be summarized from the race-focused critiques of current methodological practices in psychology:

1. Substitute the concepts of ethnicity, ethnic group, or ethnic identity for race or racial group (Betancourt & López, 1993; Phinney, 1996; Yee, 1983). By concepts, advocates of this approach mean specification of factors such as values, customs, or traditions rather than merely substituting alternative labels for race or racial group. Rationales offered in support of this recommendation include (a) improving psychological research designs and theories by shifting psychologists' focus away from group-level factors to "individualistic traits" (Yee, 1983, p. 21), (b)

---

Janet E. Helms, Maryam Jernigan, and Jackquelyn Mascher, Institute for the Study and Promotion of Race and Culture and Department of Counseling, Developmental, and Educational Psychology, Boston College.

Correspondence concerning this article should be addressed to Janet E. Helms, Department of Counseling, Developmental, and Educational Psychology, Boston College, 318 Campion Hall, Chestnut Hill, MA 02467. E-mail: helmsja@bc.edu



**Janet E. Helms**

encouraging researchers to study race as a social construction rather than a biological entity (American Psychological Association, 2003), and (c) illuminating cultural phenomena underlying human behavior (Betancourt & López, 1993; Phinney, 1996).

2. Avoid using racial categories in research designs without a clear conceptual reason for doing so (Dole, 1995; Pedhazur & Schmelkin, 1991). The rationale for this recommendation is that because racial categories encompass such a wide array of unspecified attributes, it is too tempting to “fall into the trap of ‘explaining’ [racial category] differences [on the dependent variable]” by means of racial categories instead of identifying the variables associated with racial categories (e.g., exposure to discrimination, in-group bias) that relate to or affect the dependent variables in research designs (Pedhazur & Schmelkin, 1991, pp. 175–176; Phinney, 1996).

3. Replace racial categories as independent variables with independent variables derived from racial categorization (RC) theories. Unlike users of racial categories as independent variables, RC theorists do not contend that they are studying race per se. RC theorists define constructs based on people’s experiences of categorizing or being categorized into one mutually exclusive group rather than another. Thus, this perspective advocates substituting conceptually meaningful RC constructs for racial categories. Many theories exist whose constructs could be used to realize this strategy. These include, but are not limited to, social category theory (e.g., Allport, 1954), stereotype threat theory (Steele, 1997), racism theories (Clark, Anderson, Clark, & Williams, 1999), and racial identity theories (Cross, 1978; Helms, 1995).

4. Use statistical analyses of theory-derived variables to determine whether some of the myriad of constructs

potentially encompassed by racial categories or derived from RC theories can be used to replace racial categories. Although this strategy is not well known, it potentially subsumes the others. Also, perhaps because it has been used primarily to test specific hypotheses in a small number of studies, the broader applications of the strategy have not yet been recognized (Ellis & Ryan, 2003; Manly et al., 1998; Ong & Phinney, 2002; Osborne, 2001).

A logical inference from these critiques is that use of racial categories as if they are precise measures of some genuine psychological theoretical construct accords scientific legitimacy to what are essentially racial stereotypes that psychologists share with the larger society and the professional environments in which the psychologists function. Many psychologists (e.g., Helms, 1992; Hilliard, 1984; Yee, 1983), focusing primarily on the domain of intellectual testing, have attempted to alert psychologists to the illogic and potential harm to society of using racial stereotypes as if they are legitimate scientific constructs (i.e., independent variables).

It is not clear why psychology as a discipline has resisted the multiple calls to stop reifying racial categories in its theory and research, and, therefore, its implicit racial theories. Perhaps the resistance has occurred because although the rationales for the four replacement strategies have been well reasoned, their focus has been on persuading psychologists of the logic of replacing racial categories rather than on elucidating the methodological limitations of using racial categories as independent variables that make their replacement necessary.

Therefore, our primary goal in the present article is to offer a methodological critique of the use of racial categories as independent variables in psychological theory and research. We intend our perspective to be supportive of the calls for replacement previously described. To accomplish our goal, we (a) use principles of the scientific method to compare use of psychological constructs as independent variables with use of racial categories as independent variables, (b) identify common errors of interpretation in measuring and analyzing racial categories, and (c) illustrate a procedure for replacing racial categories with conceptual constructs (i.e., Replacement Strategy 4).

To identify common interpretational errors, we examined studies cited in the PsycINFO database (<http://www.apa.org/psycinfo/>) over a short period of time. For the most part, we use scores on intellectual tests as our dependent/criterion variables in our discussion because (a) racial categories are frequently used as independent variables in this line of research and (b) the methodological principles developed in research on intellectual testing often become the standards for other lines of psychological scientific inquiry (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999; Zuberi, 2001). Demonstrating the fallacy of using racial categories as independent variables should have broader implications for changing the meaning of race in psychology.



**Maryam  
Jernigan**

## Racial Categories as Independent Variables

It does not seem to be obvious why it is desirable to replace racial categories as independent variables in research designs. Therefore, we use principles from the scientific method to show that racial categories lack conceptual meaning relative to other psychological constructs. Race reportedly is as emotionally laden a topic among psychologists as it is among the lay public more generally (Scarr, 1988; Yee et al., 1993; Zuckerman, 1990), which makes it difficult to have an objective discussion about needed changes in psychology's study and conceptualization of race.

In an attempt to avoid stimulating resistance to the idea of changing the manner in which racial categories are conceptualized in research designs, we (a) summarize the general principles for defining an independent variable according to the scientific method as it is practiced in psychology, (b) illustrate the principles using a neutral but familiar construct (i.e., depression), (c) illustrate the principles using a RC construct (i.e., stereotype threat), and (d) explain why racial categories do not conform to these principles, whereas depression, stereotype threat, and, by implication, analogous constructs do.

### General Principles

Table 1 illustrates the steps in defining an *independent* variable as outlined by Cacioppo, Semin, and Berntson (2004). In Step 1, psychologists observe psychological phenomena (e.g., attitudes, behaviors) whose existence or functioning they would like to explain. In Step 2, they develop or locate a theory whose hypothetical constructs potentially explain the psychological phenomena of inter-

est. The theories and related constructs might come from any of several philosophical orientations represented in psychological literature including biological, environmental, or intrapsychic.

At Step 3, the researchers' selected construct is conceptualized as an independent variable, which is operationally defined (Step 4) by means of procedures consistent with the chosen theory, such as manipulation (i.e., experimental conditions) or measurement (e.g., test scores) of some phenomena. Once the independent variable has been so defined, hypotheses can be formulated that relate it to the phenomena that served as the catalysts for theoretical explication (i.e., dependent or outcome variables). Specification of dependent/outcome variables occurs by means of an analogous theory-driven developmental process (not shown).

### Depression

Table 1 illustrates the principles with depression as the psychological phenomenon. At Step 1, the researcher observes a number of symptoms (e.g., sadness, loss of appetite), which are verifiable by the individuals themselves as well as outside observers. Depending on the researcher's philosophical assumptions about the causes of psychological phenomena, he or she may choose theories focused on different aspects of these symptoms. For example, theorists who believe in biological bases of behavior may use theories that attribute the symptoms to biological mechanisms such as deficits in serotonin, whereas intrapsychic theorists might attribute them to characteristics of individuals such as negative cognitions, and environmental theorists might view social support as relevant. On the basis of these hypothetical constructs, the researchers develop measures or manipulations (i.e., independent variables) that reflect their conceptualization of the mechanisms that cause depression. These independent variables, in turn, are used to test hypotheses about whether depression, as conceptualized, is related to dependent variables, as conceptualized by the relevant theory of depression.

### Racial Categories

Table 1 uses the steps of the scientific process for defining independent variables to show the breakdown in the process when racial categories are used as independent variables. The breakdown occurs because racial categories function as the all-encompassing theoretical explanation of observed behavior (Step 2) as well as the independent variable (Step 3). No explicit conceptual framework guides the procedures by which research participants are assigned to one category rather than another (Step 4). Instead, the researchers' implicit beliefs about the meaning of race (e.g., physical appearance, self-designation) serve as the operational definitions of racial categories and, consequently, the independent variable is also amorphous. Thus, theorists who endorse biological bases of behavior infer biology from racial categories, environmentalists infer context from the same racial categories, as do intrapsychic theorists with respect to individual processes. A variable that means everything means nothing.





**Jackquelyn  
Mascher**

Racial categories, used as independent variables, are not based on theory and do not involve purposeful manipulation or measurement of research participants' behaviors or attributes. As a consequence, researchers infer causation from between-groups differences on the dependent variable according to their preferred implicit philosophical orientation. Thus, if between-groups differences on intellectual test scores are found in a study, biologically focused theorists interpret them as evidence of genetic causation, environmentally focused theorists interpret them as evidence of contextual causation, intrapsychic theorists interpret them as evidence of the quality of functioning of individual processes, and so forth. In other words, when between-groups differences in the phenomena that the researcher chooses to study occur, it is because the groups were somehow different. This type of circular reasoning is the converse of how explanatory constructs typically are developed according to the scientific method as it is supposed to be used in psychology (Cacioppo et al., 2004).

### **RC Constructs**

Table 1 summarizes the process of developing RC constructs as independent variables (i.e., Replacement Strategy 3). The general premise of this approach is that phenomena of interest to psychologists occur because of differences in conditions of racial socialization or experiences. Explanatory theories, constructs, and independent variables in RC theories evolve out of observable psychological phenomena related to RC (e.g., racism, group cohesion). In this perspective, RC-focused independent variables may be used to explain dependent variables or phenomena (e.g., test scores) that have no obvious racial attributes.

For example, Steele (1997) developed the theory of stereotype threat to explain the lower performance of Black

test takers relative to White test takers on tests of intellectual abilities. Steele's original theoretical formulation was that lower performance was attributable to research participants' attempts to avoid performing on the tests in a manner that would confirm society's negative stereotypes about the intellectual abilities of people who society categorizes as Black. He used both threat versus nonthreat experimental conditions (i.e., categorical variables) and Black and White college students as his independent variables. Analog test scores were his dependent variables. Subsequent researchers have shown that stereotype threat may also be measured, rather than or in addition to being manipulated, using measures developed for that purpose (e.g., Ployhart, Ziegert, & McFarland, 2003).

Nevertheless, Steele (1997) (a) observed a phenomenon in need of explanation, test scores (Step 1); (b) developed a theory-based explanation of it, fear of confirming negative stereotypes (Step 2); (c) defined an independent variable consistent with the conceptual rationale, stereotype threat (Step 3); and (d) specified procedures for studying threat, experimental manipulations of participants' perceptions of tests (Step 4). Therefore, when differences between or within groups occurred under different conditions of threat, it was reasonable for the researchers to infer that stereotype threat caused the differences.

### **How Conceptual Independent Variables Differ From Racial Categories**

In sum, racial categories should not be used to explain psychological phenomena because the categories have no conceptual meaning. Assignment of research participants to a racial category reveals something about the researchers' beliefs about race but nothing about the behaviors or attributes of the research participants. Nor does such assignment mean that persons in one category have more or less race than those in another category. Therefore, when racial categories are used as independent variables, racial attributes are being inferred from the researcher's underlying beliefs about the nature of race even though racial categories cannot accurately reflect such beliefs.

In contrast, conceptually based psychological constructs may be measured or manipulated and, in some cases, both operations are possible. Depression, for example, may be operationally defined as a continuous variable (e.g., scores on an inventory assessing depressed thinking) or a categorical variable based on raters' judgments (e.g., depressed thoughts vs. not depressed thoughts). Membership in the depressed category means that the person manifests more of the attributes that typify depressed people according to some conceptual model. In a similar manner, membership in the stereotype-threat condition (i.e., category), for example, rather than the nonthreat condition means that the person has been exposed to manipulations intended to arouse fear of confirming negative racial stereotypes.

Consequently, if psychologists' intentions are to uncover the psychological aspects of individuals that racial categories mask, then they should make greater use of RC constructs and variables in their research designs because

**Table 1**

*Summary of General Steps in the Scientific Method for Developing Independent Variables as Applied to Depression, Racial Categories, and Racial Categorization Constructs*

Step 1: Catalyst psychological phenomenon	Step 2: Formulation of theory	Step 3: Independent variable defined	Step 4: Operational definition
Defining psychological phenomena			
Behaviors Attitudes Processes	Biological Environmental Intrapsychic		Experimental manipulations Measurement
Depression as the psychological phenomenon			
Symptoms (e.g., sadness, loss of appetite)	Biological	Serotonin deficits	Measured serotonin
	Environmental Intrapsychic	Negative cognition Social support	Frequency of negative thoughts
Racial categories as the psychological phenomenon			
Nonspecific	Categories	Categories (e.g., Black, White, Asian)	Self-description Researcher observation Archival data Combination procedures
Racial categorization constructs as the psychological phenomenon			
Low test scores In-group bias Resilience	Stereotype threat Racial identity Racism theory	Stereotype threat conditions Racial identity attitudes Level of racism	Experimental manipulations Scores on a scale

they are amenable to manipulation, measurement, and interpretation. If their intentions are merely to describe racial-category differences on a myriad of dependent variables, then they should be advised that they are engaging in an enterprise with little scientific value.

## Identifying Common Misinterpretations of Racial Categories

The perspective that psychologists wrongly infer racial causation from racial categories, as they are used in their research designs, is not new, as it has been frequently argued by proponents of the four replacement strategies previously described. So, we attempted to discover whether there were common misconceptions about the measurement and statistical analysis of racial categories that fueled contemporary researchers' ostensible resistance to replacing them as independent variables. Toward this goal, we conducted a relatively informal abbreviated search of the PsycINFO database, which covered the time span from January 1, 2004 to June 2, 2004. We used the search terms *race*, *study*, and *variables* because we wanted to identify empirical studies quickly. Hence, the results of our search are not comprehensive but hopefully provide a reasonable sampling of the manner(s) in which contemporary psychol-

ogy researchers use racial categories in their research designs.

Our search returned a total of 112 citations. One was in a language that none of us comprehended and so it was excluded from further consideration. Another 11 could not be located during the time period available. Thus, we manually reviewed 100 documents and identified 38 that were usable empirical studies of racial variables. The smallness of our sample should not be problematic because 73% of them used racial categories as independent variables, predictors, covariates, or as controls for race. The remainder used them to describe sample composition. We discuss general issues derived from our review to avoid the appearance of blaming any single researcher for what historically have been standard research practices where the study of race is concerned.

## Measurement and Analysis of Racial Categories

Our noncomprehensive examination of measurement and research design texts suggests that their authors do not give proper attention to the unique aspects of measuring and analyzing racial categories as independent variables, predictors, covariates, and so forth. Most of the texts treated racial categories (or pseudonyms for them) as if they were conceptually meaningful measures of independent vari-

ables but did not consider the unique interpretational and analytic issues that use of racial categories as independent variables evokes. Therefore, we discuss some improprieties in measurement of racial categories and interpretation of results of statistical analyses involving them that we discovered in our database search.

It is convenient when discussing these issues to use regression and correlation terminology primarily because doing so emphasizes the points that (a) use of racial categories as independent variables always implies a between-groups comparison, (b) racial categories can be no more than conceptually vacuous correlates of other variables no matter what statistics the researcher uses to analyze them, and (c) "all statistical analyses (e.g., analysis of variance, *t* test, *R*, *R*<sub>c</sub>) are correlational." This latter point means, in part, that effect sizes "analogous to *r*<sup>2</sup> [that is, percentage of variability in participants' dependent variable scores that is predicted by racial categories] can be computed in all parametric analyses" (Thompson, 2002, p. 68; also see Thompson, 2000, for a discussion of multivariate analyses as correlational).

### **Racial Categories as Measures**

Measurement is the quantification of attributes (Nunnally, 1967). Perhaps we disposed of the misperception that racial categories represent attributes in our earlier discussion of the process of defining independent variables. However, researchers assume that because they can assign numbers to racial categories and these numbers appear to relate to or affect other variables, they have measured some underlying racial attribute of research participants. Here we attempt to show why this assumption is not true by discussing some common misinterpretations of racial categories as quantified independent variables. We discuss measurement and interpretational issues together because it was difficult to separate them.

### **Nonreporting**

Virtually none of the researchers described the procedures used (i.e., operational definitions) to create racial categories in their Methods sections. Moreover, except for occasional table notes in results sections, researchers either did not report or wrongly reported how they quantified racial categories. Without information about the researchers' coding system, independent researchers are unable to interpret their results. Thus, the omissions meant that the researchers' implicit racial theories could not be challenged.

### **Racial Attributions**

Inferring from table notes, one sees that researchers sometimes used dummy coding to quantify racial categories used as independent variables. However, many researchers interpreted their dummy-coded variables to be measures of racial characteristics of participants rather than the regression equivalent of comparing the mean scores of racial categories using analyses of variance (ANOVAs) or *t* tests. Typically, when dummy coding is used, the researcher develops a dichotomous variable or set of variables to represent racial categories in the study by labeling mem-

bers of the focal group (e.g., Blacks) with ones and members of the standard or comparison group (e.g., Whites) with zeros. However, either category may be assigned either number without changing the nature of the interpretation of the results. In other words, the numbers themselves convey no conceptual meaning.

Yet researchers often indicated that they had dummy coded only one racial category (e.g., "We dummy coded Latinos") even though there were several racial categories in their study and it would have been impossible for them to analyze only one group using dummy codes because the phrase implies a comparison between means. Even though the mean scores of the identified racial group must have been contrasted against the average scores of other groups in the study, the researchers often used the dummy-coded variable to make implicit racial-category inferences about only the group of interest to the researcher. For example, the researcher might report, "Latinos had low self-esteem," thereby, attributing a negative trait to them rather than "Latinos had lower average levels of self-esteem than Whites."

In a correlation/regression analysis, a positive sign for the regression coefficient indicates that the group coded with ones (e.g., Blacks) had a higher mean score on the dependent measure (e.g., intellectual test scores), and a negative sign means that the comparison group had a higher mean score on the dependent measure. Sometimes researchers interpreted the dummy-coded variables as continuous variables reflecting the racial or ethnic traits of the focal group as opposed to the groups' mean differences on the dependent measure. This error was reflected in statements, such as "Being Latino was related to lower self-esteem" as opposed to "The group classified as Latino had lower average self-esteem scores than the group classified as Whites."

### **Attributing Racial Meaning**

Perhaps because racial categories were often the only independent variable in their studies, researchers typically went beyond merely using them to describe differences. Instead, on the basis of significant between-groups differences, they often drew connections between variables that were not even measured in their studies (e.g., "Previous studies have found that Blacks are academically disengaged"), thereby, seemingly reinforcing preexisting stereotypes.

In sum, because racial categories are so imprecisely defined but are so easily quantified, researchers attribute more meaning to them than is merited (Pedhazur & Schmelkin, 1991). Equating the quantification of racial categories with measurement of an intended underlying racial construct leads to perpetuating the myth that race or ethnicity has been measured and, therefore, may be used interpretatively.

## **Research Design and Statistical Analyses of Racial Categories**

Many methodologists and researchers appear to conflate racial categories, as used in research designs, with race as

nonscientists use it. For both groups, confusion about the etiology of racial categories in society appears to underlie their misinterpretation of racial categories, although the confusion is manifested differently.

### **Methodologists**

Many methodologists appear to believe that racial categories, used as independent variables, reflect innate attributes of research participants. Consequently, they provide nonsensical advice for analyzing them in which the methodologists' racial beliefs are obliquely embedded. Two examples may illustrate this point.

**Example 1.** In their discussion of best practices for analyzing moderators and mediators, Frazier, Tix, and Barron (2004) advised that a "given variable may function as a mediator or moderator, depending on the theory being tested" (p. 116), but they consider race to be a "group variable [that] is *naturally occurring* [italics added]" (p. 117), meaning that researchers who say that they are studying race do not have to conform to the same standards of rigor in their research designs as are required for interpretation of other categorical or categorized variables.

It is difficult to imagine a psychological theory that would account for the natural occurrence of racial categories, but it is the case that racial categories, as used by psychologists, are not naturally occurring. If they were, researchers would not have been able to assign and reassign research participants to racial categories according to the researchers' sample size needs. Note that Asian Americans, for example, may be concomitantly "honorary Whites" and "not Whites" in the same study (Liang, Li, & Kim, 2004, p. 112). Each of the not-Whites racial groups (i.e., categories) is susceptible to being collapsed into a single amorphous "minority group" or buried in a White group to compensate for their small numbers in any given study.

**Example 2.** In his discussion of use of hierarchical linear models to analyze longitudinal data, Weinfurt (2000) described race (i.e., racial categories) as "intact groups" (p. 339). He defined *intact groups* as "[those that] exist outside of the research context, over which the researcher has no control" (p. 356). Typically, in research designs, *intact* implies that group members share experiences, attributes, or conditions (e.g., intact classrooms). In intellectual testing, the typical large-scale racial-group comparison study involves thousands of participants. For example, Pennock-Román (1993, Table 4, p. 22) reported the following sample sizes for citizens who took the SAT and GRE: Asian Americans ( $N = 1,983$ ), Blacks ( $N = 2,614$ ), Mexican Americans ( $N = 511$ ), and non-Hispanic Whites ( $N = 47,756$ ). It is not plausible that the test takers within each of the categories shared enough experiences or attributes to render them intact.

When groups are intact, the researchers' analytic procedures are supposed to compensate for attributes of the preexisting groups in some manner (e.g., analysis of covariance [ANCOVA]). Use of ANCOVA, for example, is intended to statistically match individuals in the two groups on some preexisting or concurrent attribute (e.g., English proficiency) other than the independent variable (e.g., ste-

reotype-threat intervention) that might bias the study's outcome through its effect on the dependent variable (e.g., intellectual test scores).

Yet because racial categories are not intact groups, researchers should not use analytic strategies intended to compensate for their intactness. Also, because racial categories do not measure shared (as opposed to researcher-inferred) attributes of research participants, they should not be analyzed as covariates themselves because it is not clear what constructs the researcher is compensating for. In our overview of current practices, it appeared that researchers often engaged in both practices.

### **Researchers' Confusion**

Researchers' implicit theories about the meanings of racial categories are also evident in their statistical analyses. When racial categories function as independent variables in analyses, they have the property that they always signify between-groups comparisons. This aspect may be obscured because of (a) researcher's racial beliefs, (b) procedures conducted on measures or samples during preliminary analyses, (c) attempts to control for race, and (d) use of multivariate analyses.

**Researcher beliefs.** Researchers often hypothesize that a correlate of racial categories (e.g., poverty) explains Behavior X (e.g., minorities and poor people do X). But to test their hypotheses, they study the main effects of racial categories created by the researcher (e.g., minority group vs. White group) or socioeconomic status on a dependent measure without regard to whether poverty levels and racial categories are related in their designs. Significant findings related to either variable are then interpreted as an effect due to race or ethnicity.

**Preliminary operations.** Sometimes researchers conduct preliminary analyses on their independent measures to determine whether the scores of racial categories differ on them and either discard or alter their measures if differences are found. By doing so, the researchers are actually treating racial categories as the primary measure of the independent variable in their design, when instead the discarded or unaltered variable should serve in this role because it can potentially replace conceptually meaningless racial categories in subsequent analyses (cf. Replacement Strategy 2).

**Controlling for race.** A fairly common strategy for validating test scores in research on intellectual testing is to analyze relationships between scores on intellectual tests, used as independent variables, and some criterion variable (e.g., academic performance), while disregarding previously discovered or known racial-group differences on the test score independent variable. Researchers often refer to this strategy as *controlling* or *adjusting* for race or, alternatively, testing *moderators* (cf. Frazier et al., 2004).

However, when between-groups racial differences on test scores used as dependent variables exist, then using the test scores as independent variables in other studies or with other variables, without investigating conceptually based alternative explanations for such differences, begs the question of what the racial categories mean with respect to



test scores. Not providing an empirically supported reason for why the racial categories differ merely allows researchers to infer proof of whatever stereotypic interpretation of the test score differences that they desire.

**Multivariate analyses.** A full discussion of the role of racial categories in the various types of multivariate analyses (e.g., canonical correlation, hierarchical linear models) is beyond the scope of this article. However, the critical factor here is that even in these models, racial categories still represent between-groups fixed effects just as they do when simpler analytic strategies (e.g., *t* tests, ANOVAs) are used. *Fixed* refers to the fact that their numerical values are set by the researcher rather than defined on the basis of participants' behaviors or attributes. *Between-groups* signifies that the means of dummy coded racial categories (i.e., aggregated data) on the dependent variables are compared just as they are in simpler analyses.

When researchers use longitudinal designs, for example, they collect multiple measures of the dependent measure(s) over occasions, but the coded nature of racial categories (e.g., Blacks = 1, Whites = 0) does not change. Also, their status as nonconceptual nonexperimental variables does not change (Weinfurt, 2000). Therefore, when racial categories are used as a group-level variable to predict individuals' patterns of change over time (e.g., growth curves), a positive regression coefficient indicates that the mean of the dependent measures, averaged across occasions, of the group coded with ones is higher than the comparison group.

In sum, regardless of how racial categories are quantified and analyzed, they cannot be more than they are, place holders for conceptual constructs. All of the psychological research, in which racial categories have been compared on some dependent measures, provides evidence of a variety of variables associated or not associated with racial categories. Yet these associations reveal nothing about the racial characteristics of research participants because racial categories are conceptually void.

## Implementing Replacement Strategy 4

Replacement Strategy 4 offers a methodology for implementing the other three strategies, thereby potentially replacing racial categories in psychological theory and research with conceptually meaningful constructs. The strategy permits researchers to address the questions of whether conceptually based independent variables representing ethnicity or cultural factors (Strategy 1), associated with racial categories (Strategy 2), or derived from RC theories (Strategy 3) replace racial categories as predictors of dependent measures (Strategy 4). Another label for *replace* is *mediate*.

### Description and Example

Use of the strategy requires four components: (a) an independent variable, intended to replace racial categories, developed according to principles of the scientific method (see Table 1); (b) a dependent variable of interest (e.g., test scores); (c) hypotheses, derived from the same theory as

the independent variable, describing how the independent variable is expected to relate to or affect the dependent variable; and (d) racial categories to be replaced.

We use hierarchical multiple regression (HMR) analysis to describe the test of conceptual independent variables as mediators of ostensible relationships between racial categories and dependent variables as did Osborne (2001), Ong and Phinney (2002), and Ellis and Ryan (2003). Manly et al. (1998) used ANCOVA, but multiple regression analysis would have yielded identical results assuming its assumptions were met (Cohen, Cohen, West, & Aiken, 2003). Researchers, intending to use more complex statistical analyses (e.g., multilevel regression analyses), may think of this procedure as a preliminary analysis to determine whether scores on the conceptual independent variable permit them to eliminate racial categories as a group-level variable in their analyses. If not, then they must still be cautious about attributing conceptual meaning to the group effect as opposed to merely describing it.

The proposed HMR analysis consists of three steps: (a) entry in the first step of the conceptual variable(s) that is intended to represent the explanatory construct in the research design, (b) entry of the racial-categories variable(s) in the second step to determine the amount of variance it predicts or describes beyond what is explained by the conceptual variable, and (c) determination of whether the effect size and significance level of the racial-category variable(s) suggests that its relation to the dependent variable appears to be meaningful, that is, adds importantly to the prediction of scores on the dependent variable beyond what is explained by the conceptual independent variable(s).

Of the three studies of intellectual testing that fortuitously used the strategy, Ellis and Ryan's (2003) report of their descriptive data (e.g., means, standard deviations) was most complete (Table 1, p. 2615), which permitted us to replicate their Phase 2 findings (Table 3, p. 2617) within rounding error. For their conceptual variable in Step 1 of their hierarchical regression analysis, they entered the variable of ineffective test-taking strategies (ITT). It explained 20% of the variability among test scores, their dependent variable. In Step 2, the researchers entered the dummy coded variable for racial categories (Black = 1, White = 0). It described an additional 11% of the variance in test scores.

Using the principle that all parametric analyses are correlational, which was previously articulated (Thompson, 2002), researchers can convert percentages of variance explained (e.g.,  $r^2$ ) to  $d$  (i.e., the number of standard deviations that separate the mean scores of two groups). Thus, Ellis and Ryan's (2003) conceptual variable (ITT) explained the equivalent of one standard deviation (i.e.,  $d = 1.00$ ) of difference between the mean scores of their Black and White participants, which is the prototypical standardized mean difference between test scores of Black and White test takers (Sackett, Hardison, & Cullen, 2001). Yet racial categories continued to describe almost one standard deviation ( $d = 0.70$ ) of separation after ITT was statistically removed.



## Interpretation

The results of the replacement analysis in this example suggest that the conceptual independent variable (i.e., ITT) explained as much variance in intellectual test scores as is typically described by racial categories when Black and White test takers are compared (Sackett et al., 2001). Yet the results also suggest that the researchers should continue to conceptualize and search for additional conceptual factors, perhaps more explicit RC variables, to account for the additional sizable variance that was still described by comparing racial categories.

Nevertheless, Ellis and Ryan's (2003) findings with respect to ITT, their conceptual variable, may be reasonably interpreted as a "cause" of participants' behaviors (i.e., test scores), whereas the portion due to racial categories should not be interpreted as explanatory. Development of interventions intended to change participants' ITT might be a reasonable outcome of their findings, but no intervention logically follows from racial category differences.

## Conclusion and Implications

Psychologists have expressed a variety of beliefs about what race is or is not, which we chose not to reiterate in any great detail here because these suppositions are essentially irrelevant to what psychologists do with respect to research practices. Psychologists' beliefs about race, for the most part, have not been subjected to empirical investigation using the scientific method as it is supposedly practiced in psychology. Instead, regardless of their theoretical orientations (e.g., biological, environmental, intrapsychic), psychologists have knowingly continued to use and interpret essentially the same single, flawed, atheoretical operational definitions of race (i.e., factitious racial categories) as if racial categories constitute both theoretical constructs and measures of immutable racial characteristics of research participants, when, in fact, they do neither.

In lay society, imputing behavior to a person solely on the basis of the person's ascribed racial category is called *racial profiling*. It is a practice that increasingly is being recognized as unjust. Yet in psychology, entire content areas (e.g., high-stakes testing) rest on the belief that it is acceptable to use membership in arbitrary racial categories to explain individuals' behaviors (e.g., test performance). A harmful consequence to society of this practice is that scores on intellectual tests, for example, are used to make decisions about selection and placement even though it is known that the test scores differentially favor or disfavor test takers assigned to one racial category rather than another. Such usage strengthens racial stereotyping. We proposed the basics of a strategy by which psychologists can begin to redress the wrongs caused by racial profiling as it has been condoned in psychology.

Someday, some bright litigant might pose the question, "Why do racial groups (i.e., categories) differ on X behavior?" More criterion validity or test bias studies, as they have been conducted historically, will not provide answers to the question because such studies seek causation in the racial properties of dependent measures (e.g., tests)

rather than in the attributes of researchers or research participants that result from the phenomena of RC. Unless more psychologists join RC theorists in conceptualizing and measuring or manipulating theory-based independent variables, derived from individuals' experiences of RC, rather than continuing to infer traits and behaviors from racial category ascriptions, the consequences to psychology as an objective science could be dire.

## REFERENCES

- Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Addison Wesley.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Psychological Association. (2003). Guidelines on multicultural education, training, research, practice, and organizational change for psychologists. *American Psychologist, 58*, 377–402.
- Betancourt, H., & López, S. R. (1993). The study of culture, ethnicity, and race in American psychology. *American Psychologist, 48*, 629–637.
- Cacioppo, J. T., Semin, G. R., & Berntson, G. G. (2004). Realism, instrumentalism, and scientific symbiosis: Psychological theory as a search for truth. *American Psychologist, 59*, 214–223.
- Clark, R., Anderson, N. B., Clark, V. R., & Williams, D. R. (1999). Racism as a stressor for African Americans: A biopsychosocial model. *American Psychologist, 54*, 805–816.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- Cross, W. E., Jr. (1978). The Thomas and Cross models of psychological nigrescence: A review. *Journal of Black Psychology, 5*, 15–31.
- Dole, A. (1995). Why not drop race as a term? *American Psychologist, 50*, 40.
- Eisenman, R. (1995). Why psychologists should study race. *American Psychologist, 50*, 42–43.
- Ellis, A. P. J., & Ryan, A. M. (2003). Race and cognitive-ability test performance: The mediating effects of test preparation, test-taking strategy use and self-efficacy. *Journal of Applied Social Psychology, 33*, 2607–2629.
- Fairchild, H. H. (1991). Scientific racism: The cloak of objectivity. *Journal of Social Issues, 47*, 101–115.
- Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology, 51*, 115–134.
- Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing. *American Psychologist, 47*, 1083–1101.
- Helms, J. E. (1994). The conceptualization of racial identity. In E. Trickett, R. Watts, & D. Birman (Eds.), *Human perspectives on people in context* (pp. 285–311). San Francisco: Jossey-Bass.
- Helms, J. E. (1995). An update of Helms's White and People of Color racial identity models. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 181–198). Thousand Oaks, CA: Sage.
- Hilliard, A. G., III. (1984). IQ testing as the emperor's new clothes: A critique of Jensen's *Bias in Mental Testing*. In C. R. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing* (pp. 139–169). New York: Plenum Press.
- Liang, C. T. H., Li, L. C., & Kim, B. S. K. (2004). The Asian American racism-related stress inventory: Development, factor analysis, reliability, and validity. *Journal of Counseling Psychology, 51*, 103–114.
- Manly, J. J., Miller, S. W., Heaton, R. K., Byrd, D., Reilly, J., Velasquez, R. J., et al. (1998). The effect of African American acculturation on neuropsychological test performance in normal and HIV-positive individuals. *Journal of the International Neuropsychological Society, 4*, 291–302.
- Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.
- Ong, A. D., & Phinney, J. S. (2002). Personal goals and depression among

- Vietnamese American and European American young adults: A mediational analysis. *Journal of Social Psychology, 142*, 97–108.
- Osborne, J. W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology, 26*, 291–310.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.
- Pennock-Román, M. (1993). *Differences among racial and ethnic groups in mean scores on the GRE and SAT: Longitudinal comparisons*. Princeton, NJ: Educational Testing Service.
- Phinney, J. S. (1996). When we talk about American ethnic groups, what do we mean? *American Psychologist, 51*, 918–927.
- Ployhart, R. E., Ziegert, J. C., & McFarland, L. A. (2003). Understanding racial differences on cognitive ability tests in selection contexts: An integration of stereotype threat and applicant reactions research. *Human Performance, 16*, 231–259.
- Rowe, D. C. (2002). IQ, birth weight, and number of sexual partners in White, African American, and mixed-race adolescents. *Population, 23*, 513–524.
- Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist, 57*, 305–306.
- Scarr, S. (1988). Race and gender as psychological variables. *American Psychologist, 43*, 56–59.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629.
- Thompson, B. (2000). Canonical correlation analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 285–316). Washington, DC: American Psychological Association.
- Thompson, B. (2002). “Statistical,” “practical,” and “clinical”: How many kinds of significance do counselors need to consider? *Journal of Counseling & Development, 80*, 64–71.
- Weinfurt, K. P. (2000). Repeated measures analyses: ANOVA, MANOVA, and HLM. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 317–361). Washington, DC: American Psychological Association.
- Yee, A. H. (1983). Ethnicity and race: Psychological perspectives. *Educational Psychologist, 18*(1), 14–24.
- Yee, A. H., Fairchild, H. H., Weizmann, F., & Wyatt, G. E. (1993). Addressing psychology’s problem with race. *American Psychologist, 48*, 1132–1140.
- Zuberi, T. (2001). *Thicker than blood: How racial statistics lie?* Minneapolis: University of Minnesota Press.
- Zuckerman, M. (1990). Some dubious premises in research and theory on racial differences. *American Psychologist, 45*, 1297–1303.