

© Health Research and Educational Trust
DOI: 10.1111/1475-6773.12132
RESEARCH ARTICLE

“Which Box Should I Check?”: Examining Standard Check Box Approaches to Measuring Race and Ethnicity

Abbey Eisenhower, Karen Suyemoto, Fernanda Lucchese, and Katia Canenguez

Objective. This study examined methodological concerns with standard approaches to measuring race and ethnicity using the federally defined race and ethnicity categories, as utilized in National Institutes of Health (NIH) funded research.

Data Sources/Study Setting. Surveys were administered to 219 economically disadvantaged, racially and ethnically diverse participants at Boston Women Infants and Children (WIC) clinics during 2010.

Study Design. We examined missingness and misclassification in responses to the closed-ended NIH measure of race and ethnicity compared with open-ended measures of self-identified race and ethnicity.

Principal Findings. Rates of missingness were 26 and 43 percent for NIH race and ethnicity items, respectively, compared with 11 and 18 percent for open-ended responses. NIH race responses matched racial self-identification in only 44 percent of cases. Missingness and misclassification were disproportionately higher for self-identified Latina(o)s, African-Americans, and Cape Verdeans. Race, but not ethnicity, was more often missing for immigrant versus mainland U.S.-born respondents. Results also indicated that ethnicity for Hispanic/Latina(o)s is more complex than captured in this measure.

Conclusions. The NIH's current race and ethnicity measure demonstrated poor differentiation of race and ethnicity, restricted response options, and lack of an inclusive ethnicity question. Separating race and ethnicity and providing respondents with adequate flexibility to identify themselves both racially and ethnically may improve valid operationalization.

Key Words. Measurement of race and ethnicity, health disparities research, National Institutes of Health (NIH) race and ethnicity reporting, racial and ethnic self-identification

Racial and ethnic health disparities continue to exist today in risk, prevalence, access to care, and quality of services (Betancourt et al. 2003; Brondolo, Gallo, and Myers 2009; U.S. Department of Health and Human Services 2013). While there is increasing research to identify these inequalities and their

causes, and to develop prevention and intervention programs, there are also heightened methodological concerns about the accuracy and utility of the current approach to measuring race and ethnicity in health disparities research (McKenzie and Crowcroft 1994; Jones 2001; Lee, Mountain, and Koenig 2001; Rivara and Finberg 2001; Laws and Heckscher 2002; Ford and Kelly 2005; Ramírez et al. 2005). Particularly, questions have been raised about the commonly-used approach outlined by revised Directive Number 15 from the U.S. Office of Management and Budget (OMB), with five race categories (American Indian/Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Black or African-American, White) and two ethnic categories [Hispanic/Latina(o) and non-Hispanic/Latina(o)]. Although the OMB's approach to measuring race and ethnicity is used and required for all NIH-funded research (this measure will be referred to hence as the NIH measure), the NIH has itself identified a need to improve this and other measures of race and ethnicity (National Institutes of Health 2001). This is not surprising given that these federal surveillance categories were established for legal purposes and not intended for use in scientific research (Hahn and Stroup 1994). Foremost among potential problems with current measures such as the NIH measure are that (1) it is unclear whether current measures reflect participants' actual experiences, are meaningful to participants, and are used by participants in ways that researchers intend and (2) measures may not reflect current scientific and public health understandings of the actual meanings and differentiation of race and ethnicity. Such categorical approaches and associated problems may contribute to missing data (Hahn and Stroup 1994) and misclassification across race groups (Landale and Oropesa 2002; Campbell and Rogalin 2006), both of which undermine the validity and reliability of findings from health disparities research.

Defining Race and Ethnicity: The Incongruence of Conceptual Understandings and Standard Measures

Race is a social categorization imposed on people related to physical appearance for the purpose of making hierarchical power-based distinc-

Address correspondence to Abbey Eisenhower, Ph.D., Assistant Professor, Department of Psychology, University of Massachusetts Boston, 100 William T. Morrissey Blvd., Boston, MA 02125; e-mail: abbey.eisenhower@umb.edu. Karen Suyemoto, Ph.D., is with the Departments of Asian American Studies and Psychology, University of Massachusetts Boston, Boston, MA. Fernanda Lucchese, M.A., and Katia Canenguez, M.A., M.Ed., are also with the Department of Psychology, University of Massachusetts Boston, Boston, MA.

tions in social relations (Smedley and Smedley 2005; Markus 2008). Ethnicity is a social categorization based on shared cultural values and meanings such as relational styles, values, language, and customs that is more usually self-claimed or developed in relation to feelings of belonging to a chosen community (Smedley and Smedley 2005; Markus 2008). Research supports this conceptual distinction and indicates differential relations of race and ethnicity to health (Cokley 2005; Deaux et al. 2007; Nazroo et al. 2007).

Research has indicated that *race* relates to health disparities in physiological arousal, psychological distress generally, and specific psychiatric symptoms, including depression, anxiety, somatization, obsessive-compulsive symptoms, and interpersonal sensitivity (Clark et al. 1999; Klonoff, Landrine, and Ullman 1999; Jones 2001; Harrell, Hall, and Taliaferro 2003). These effects are racially based in that they are related to (1) institutional racism, which affects access to goods, services, and opportunities such as differential access to education or health care; (2) interpersonal racism, which directly affects psychological experience or may affect access due to service providers' biases; or (3) internalized racism, which may affect health risks directly such as through increased risk-taking due to self-devaluation, or indirectly as when experiences of helplessness related to internalized racism affect health compliance (Jones 2001; Ford and Kelly 2005).

Research has also indicated that *ethnicity* contributes to health disparities. Cultural practices of ethnic groups, such as health beliefs (Castro and Hernández-Alarcón 2002), dietary practices, health-related values, attitudes about diet and exercise (Castro, Shaibi, and Boehm-Smith 2009), and community-level differences in risks and resource access (Horowitz et al. 2004) result in ethnic health disparities that cross racial categories. Differences within NIH racial groups underscore the need to attend to variability that could be related to greater exposure to racialization and racism in the United States or to ethnic variation. For example, Caribbean-Americans and African-Americans, although both may fall within the "Black or African-American" NIH race category, differ in self-reported general health, as well as in specific health outcomes (e.g., hypertension) and health-related inequalities such as income and education (Nazroo et al. 2007).

Despite the clear conceptual distinction and the research demonstrating differential effects of race and ethnicity, current measures continue to confound race and ethnicity (Afshari and Bhopal 2002; Kaplan and Bennett 2003). For example, the response option "Black or African-American" confounds

ethnicity (African-American) with race. The option “Hispanic/Latina(o)” confounds race as ethnicity, at least for some Hispanic/Latina(o)s experiencing inequities related to racialization. The confounding of Hispanic/Latina(o) race as ethnicity is particularly problematic given the substantial evidence that Latina(o)s experience racism and racialized identity, demonstrating an effect of race, not just ethnicity, on their experience (Landale and Oropesa 2002; Vaquera and Kao 2006; McDonnell and de Lourenço 2009; Frank, Akresh, and Lu 2010). The general confounding makes it difficult to accurately assess whether ethnic or racial aspects are underlying identified health disparities.

Don't We All Have Both? Incongruence between Participants' Understandings and Standard Measures of Race and Ethnicity

The brief review above illustrates that race and ethnicity are conceptually distinct. Implied in this distinction is that all individuals have both racial and ethnic experiences. However, the standard measure of race and ethnicity may not capture participants' lived experiences in both categories, potentially contributing to missing or inaccurate data. For example, Caribbean immigrants may check “Black or African-American” as the closest available race option, or they may skip the question because the specific inclusion of “African-American” in the Black category does not reflect their understanding of themselves as Black and the ethnicity question does not allow for an ethnic identity specification that makes this distinction clearer. Waters, for example, found that second-generation Caribbean Black immigrants vary in their identification as Black (racial identity), immigrant, or ethnic-specific (e.g., Jamaican or Haitian) and that these identities related to different attitudes about opportunities and responses to racism (Waters 1994). Such variability within second-generation Black immigrants would not be captured by looking solely at either race or ethnicity, even if immigration generation were accounted for. Moreover, as the Waters (1994) study suggests, the lived experiences of race and ethnicity also vary in relation to immigrant status. We have limited understanding of the ways in which recent immigrants internalize concepts of race and ethnicity as they adapt to U.S. society, yet these findings suggest that the lived experiences of many immigrants may not be captured by the NIH measure.

Furthermore, the fact that many Latina(o)s self-identify *racially* as Latina(o) is not captured by the NIH measure (Landale and Oropesa

2002; Vaquera and Kao 2006; Frank, Akresh, and Lu 2010). Indeed, Campbell and Rogalin found that 78 percent of participants who previously identified as White chose the Latina(o) category when given the option to identify as Latina(o) racially (Campbell and Rogalin 2006). Other research indicates that at least some Latino(a)s in the United States, such as Brazilians, reject ethnic identification as Latina(o), while recognizing a racial identity or identification from others as Latina(o) (Afshari and Bhopal 2002). The NIH measure captures neither race nor ethnicity as these participants experience it.

Our goal is to evaluate the NIH's standard measure of race and ethnicity relative to the above-mentioned concerns about the measure's ability to accurately reflect individuals' self-identified race and ethnicity. We approach this goal by examining (1) patterns of missing data to evaluate whether the measure's operationalization of race and ethnicity contributes to participants' likelihood of skipping NIH questions and (2) patterns of agreement between responses on the NIH measure and participants' self-identified race and self-identified ethnicity.

METHODS

Participants

Data were gathered through a larger study assessing child functioning, needs, and service preferences among economically disadvantaged caregivers of young children. Participants were 236 caregivers raising young children, recruited from urban Women Infants and Children (WIC) clinics in the Boston metropolitan area; in line with WIC criteria, all families earned ≤ 185 percent of the federal poverty level. On the NIH race measure, responses other than "Black or African-American" or "White" were endorsed by very few respondents, including five "Native Hawaiian or Other Pacific Islander," seven "American Indian/Alaska Native," and five "Asian" endorsements. Thus, these categories were not included in subsequent analyses, resulting in 219 included respondents [mean age = 28.7 years (SD = 6.8); 93 percent biological mothers].

Thirty percent of respondents were immigrants from outside the United States and its territories, 5 percent were migrants from Puerto Rico, and 46 percent were born in the mainland United States; 20 percent did not provide place of birth.¹ On average, immigrant respondents had been in the United States for 10 years (SD = 7, range: 0.4–34 years). As U.S.

citizens, those who migrated from Puerto Rico may experience the structural supports and access to federal benefits available to other U.S. citizens, yet they may experience acculturation and racialization challenges similar to those migrating from outside the United States. In analyses below, we include Puerto Rican-born migrants in the “immigrant” group because we are primarily interested in whether developmental experience and familiarity with the U.S. system of racialization is related to differences in survey responses.

Fifty-five percent of those responding had a high school education or less; 37 percent had attended at least some college. The majority (51 percent) reported a native language other than English, including Spanish (21 percent of total), Cape Verdean Creole (15 percent), and 16 others. Eighty percent reported speaking English very well (66 percent) or pretty well (14 percent); 10 percent reported speaking English very little or not at all. Regarding gross household income, 45 percent of those responding reported earning \$15,000/year or less, and 72 percent reported \$25,000/year or less. Only 10 percent earned over \$35,000/year.

Procedures

Data collection occurred during May–October of 2010. Bilingual (Spanish and English) research assistants approached caregivers in WIC clinic waiting areas to request their completion of a 20-minute survey assessing child functioning and family needs. This paper-and-pencil survey was written to reflect no more than a sixth-grade reading level, and research assistants were available for questions about any items. Participants responded in English (95 percent) or Spanish. Eligible participants were those raising one or more children aged 11 months to 5 years 11 months. Parents received a \$10 honorarium.

Measures

This study focused on the items measuring race and ethnicity, which were preceded by questions about child functioning and family demographics. Race and ethnicity items were administered in the following order:

1. Open-ended measures of self-identified race and ethnicity, with the below prompts followed by a blank line for participants to write responses:

- “Race is based on how you look (often skin tone or facial features) and how you think of yourself. In your own words, what race(s) or racial group(s) do you belong to?”
- “Ethnicity typically emphasizes the common history, nationality, geography, language, food, or dress of groups of people (such as Haitian, African-American, European-American, Dominican, Irish, Cantonese, etc.). In your own words, to which ethnic group(s) do you belong?”

2. The standard NIH measure as follows:

- “Funding agencies require us to ask about your race and ethnicity in the following format:
- Which group below most accurately describes your *race*?
 - i American Indian/Alaska Native
 - ii Asian
 - iii Native Hawaiian or Other Pacific Islander
 - iv Black or African-American
 - v White
- Which group below most accurately describes your *ethnicity*?
 - i Hispanic or Latina(o)
 - ii Not Hispanic or Latina(o)”

The open-ended measure was positioned first, so that these responses were not primed by the categories provided in closed-ended items.

For ease of analysis, responses to open-ended items were subsequently aggregated into categories; see notes under Figures 1 and 3 for responses included in each aggregate category.

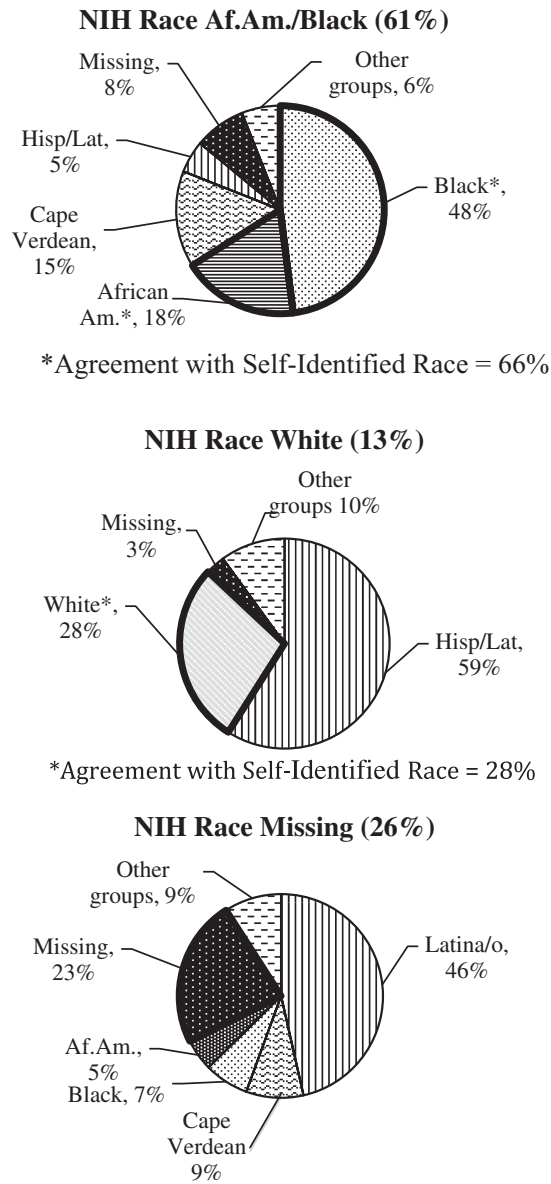
RESULTS

We evaluated the NIH race and ethnicity items by examining:

1. Rates and patterns of completion versus missingness of the NIH items relative to self-identification items. High rates of missingness on the NIH measure raise concern about whether participant characteristics used in comparisons are being fully captured. We also examined the degree to which NIH items yielded disproportionately

missing data for specific groups, testing specific hypotheses of comparative groups on missingness versus completion with chi-square analyses or Pairwise (2 × 2) Fisher’s exact tests with a Bonferroni correction for multiple tests. Finally, we compared rates of completion in immigrant versus nonimmigrant participants with chi-square analyses.

Figure 1: Aggregated Self-Identified Race Responses within Each NIH Race Category



Note. Self-identified race responses were aggregated into the following categories: “Black” includes those who self-identified as Black ($n = 46$), Black American ($n = 1$), or African/Black and Black/African ($n = 5$), Hispanic Black ($n = 4$), and Cape Verdean-Black ($n = 2$) but excludes those who identify solely as African; “African American” includes AfAm ($n = 25$), AfAm/West Indian ($n = 1$), Caribbean/AfAm ($n = 1$), and AfAm/Black ($n = 10$) but excludes those who identified solely as African; “Cape Verdean” includes Cape Verdean ($n = 23$), American/Cape Verdean ($n = 2$); “Hispanic or Latino” includes Hispanic ($n = 32$), Hispanic/Latino/a (1), Latino/a ($n = 7$), Hispanic/Puerto Rican ($n = 2$), Puerto Rican ($n = 5$), Spanish/Puerto Rican ($n = 1$), Hispanic (light skin, light + dark hair, green eyes) ($n = 1$), and Brown/Hispanic ($n = 1$); “White” includes White ($n = 6$), Caucasian ($n = 1$), and Caucasian/Eastern European ($n = 1$). “Other groups:” Aggregate categories that had fewer than 8 participants included multiracial (e.g., Mestizo, mixed, two or more, Black/White, Triguena, Black/German/Italian, Portuguese/Irish) ($n = 8$), African ($n = 1$), and Spanish or Spanish/American ($n = 6$). Pie graph sections outlined in bold represent agreement between NIH race and self-identified race.

2. Rates and patterns of agreement between NIH items and self-identified race and ethnicity responses. Low match rates raise questions about what the NIH items are measuring if they are not measuring self-identified race and ethnicity. We also examined the extent to which NIH items matched self-identification within specific subgroups. Finally, we compared agreement rates in immigrant versus nonimmigrant participants.

Preliminary analyses indicated that, within the immigrant group, rates of missingness and agreement were not related to number of years spent in the United States (t s ranged from 0.13 to 1.06, all nonsignificant). Time in the United States was therefore not controlled for in the analyses reported below.

Race

Completion versus Missingness. More than twice as many participants left the NIH race item blank than left the self-identified race item blank [26 percent ($n = 57$) vs. 11 percent ($n = 24$)]. The majority of those who skipped the NIH race item but responded to the self-identified race item self-identified as Hispanic/Latina(o) (59 percent, $n = 26$). Indeed, as shown in Figure 1, self-identified Hispanic/Latina(o)s made up 46 percent of those who skipped the NIH race item. We hypothesized that missingness on the NIH race item would be significantly higher for self-identified Hispanic/Latina(o)s than for any other group. Pairwise (2×2) Fisher’s exact tests compared missingness versus completeness between self-identified race groups, with a Bonferroni correction applied to adjust for multiple tests (acceptable $p \leq .0125$ for this number of groups). Our hypothesis was supported: self-identified Hispanic/Latina(o)s

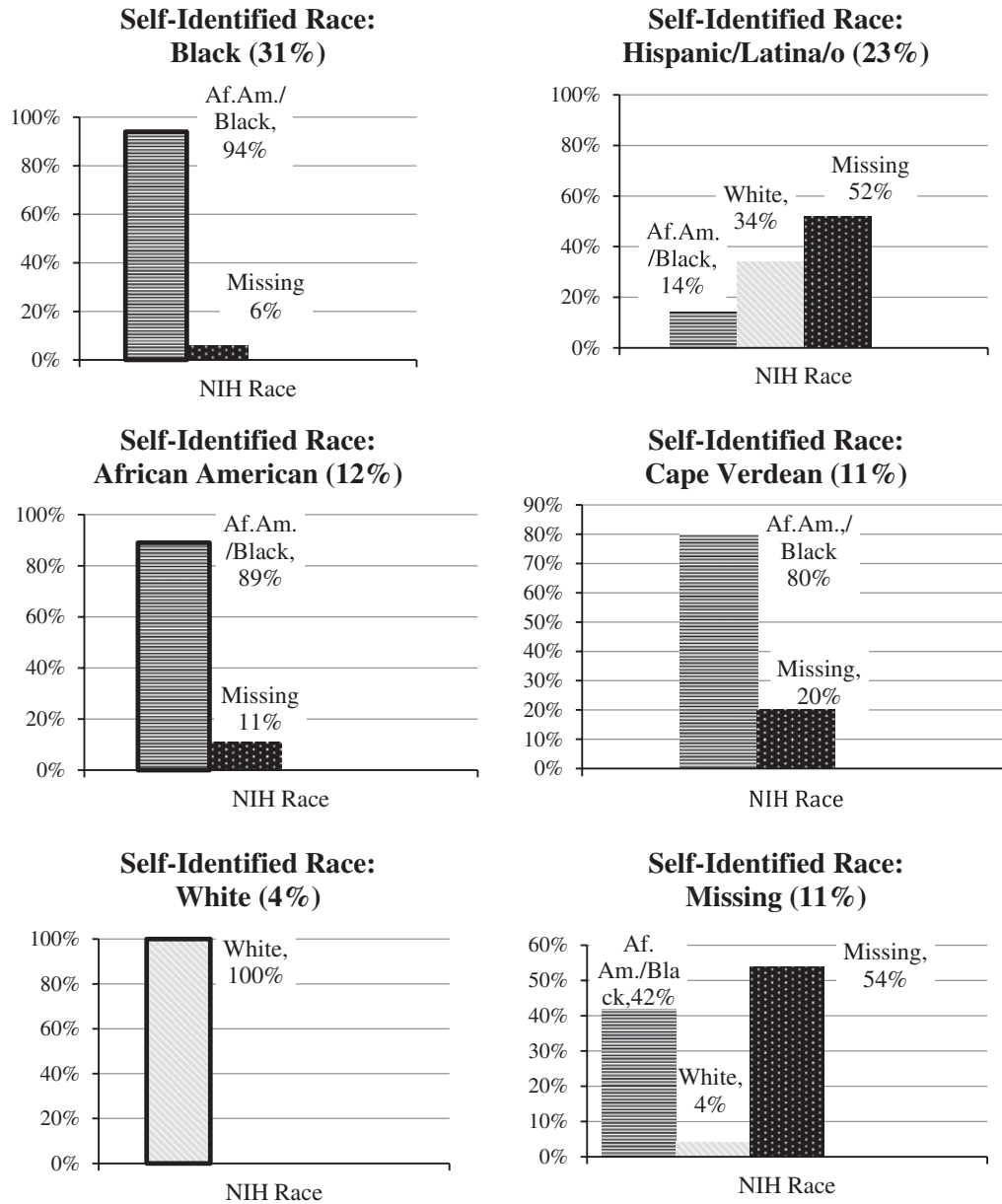
were significantly less likely to complete the NIH race item as compared to self-identified African-Americans ($p < .001$), Blacks ($p < .001$), Cape Verdeans ($p = .008$), and Whites ($p = .006$). In comparing immigrants and nonimmigrants, we included Puerto Ricans in the immigrant group in all analyses, as explained above; the NIH race item was more often missing for immigrant respondents (37 percent) than for mainland U.S.-born respondents (14 percent), $\chi^2 = 13.05$, $p < .001$.

Agreement between Measures. Responses to the NIH race item matched individuals' self-identified race for only 44 percent of participants ($n = 96$), indicating that the NIH race item did not reflect self-identified race(s) for more than half of respondents (56 percent, $n = 123$). Forty-six percent of mismatches ($n = 57$) were due to missing data on the NIH race item. However, even when including only participants who completed both the NIH and self-identified race items, match was only 64 percent. Figure 1 shows the self-identified race categories observed within each NIH race category. As shown in Figure 1, among participants who endorsed the Black or African-American NIH race category, 66 percent of these respondents self-identified as racially Black or African-American. Racially self-identified Cape Verdeans (15 percent) and racially self-identified Hispanic/Latina(o)s (5 percent) were the largest categories of mismatch. Among participants who endorsed the White NIH category, only 28 percent of these respondents self-identified racially as White; racially self-identified Hispanic/Latina(o)s were the largest category of mismatch (59 percent).

We also examined response patterns for immigrants versus mainland U.S.-born respondents. Agreement between NIH race responses and self-identified race was significantly lower among immigrant respondents (29 percent) than among mainland U.S.-born respondents (56 percent), $\chi^2 = 12.78$, $p < .001$. However, this difference was fully accounted for by the markedly lower rates of completion (higher missingness) among immigrants versus mainland U.S.-born respondents; when looking only at respondents who completed both NIH race and self-identified race items, agreement rates did not differ across immigrants (54 percent) and mainland U.S.-born respondents (66 percent), $\chi^2 = 1.88$, $p = .17$.

Patterns of Agreement and Completion by Self-Identified Race. When comparing across self-identified race groups, as shown in Figure 2, great variability is

Figure 2: NIH Race Responses within Each Aggregate Self-Identified Race Category



Note. Bars outlined in bold represent agreement between NIH race and self-identified race. Included above are any response categories that had more than five respondents.

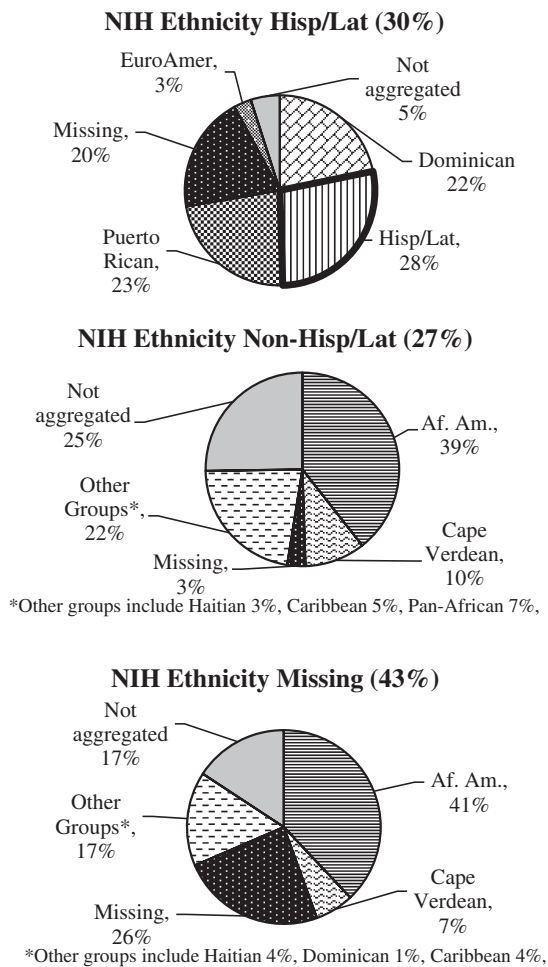
evident in the extent to which self-identified race is captured by the NIH measure. Self-Identified Black, African-American, and White respondents are reflected as such on the NIH measure at rates of 94, 89, and 100 percent

respectively. Meanwhile, racially self-identified Hispanic/Latina(o)s and Cape Verdeans, due to the absence of these groups as racial categories on the NIH measure, have match rates of 0 percent.

Ethnicity

Completion versus Missingness. More than twice as many participants skipped the NIH ethnicity item (43 percent, $n = 95$) as skipped the self-identified ethnicity item (18 percent, $n = 40$). As shown in Figure 3, of those who skipped the NIH ethnicity item, the largest group (41 percent) self-identified ethnically as African-American. Pairwise comparisons using Fisher’s exact test and

Figure 3: Aggregated Self-Identified Ethnicity Responses within Each NIH Ethnicity Category



Note. Self-identified ethnicity responses were aggregated into the following categories: “African American” includes those who self-identified as African American ($n = 61$) and Black/African American ($n = 1$); “Hispanic or Latino/a” includes those who self-identified as Catholic Hispanic (1), Hispanic (14), Latin (1), Latin-American (1), and Latino/a (2); “Puerto Rican” includes those who specifically identified as Puerto Rican ($n = 17$); “Dominican” includes those who self-identified as Dominican ($n = 13$), Hispanic/ Dominican ($n = 1$), and American-Dominican ($n = 1$); “Cape Verdean” includes those who self-identified as Cape Verdean ($n = 10$), Cape Verdean/Portuguese ($n = 2$), and American/Cape Verdean ($n = 1$). “Caribbean” includes Caribbean ($n = 2$), Jamaican ($n = 1$), parents Trinidad and Tobago ($n = 1$) and West Indian ($n = 1$). “Pan-African” includes those who self-identified as African ($n = 6$), African/Black ($n = 1$), Yoruba/African ($n = 1$); “European American” includes those who self-identified as Caucasian ($n = 1$), English ($n = 1$), English/Irish ($n = 1$), Irish ($n = 2$), White/Spanish ($n = 1$), Spanish ($n = 1$); “Haitian” includes Haitian ($n = 5$) and African Haitian American ($n = 1$). “Missing” includes all those who left the item blank ($n = 40$). “Not aggregated:” Responses not included in aggregate categories include ($n = 1$ unless otherwise indicated): food, N/A, other, both, Catholic, all of them, none ($n = 2$), American ($n = 3$). Responses of Black ($n = 7$) and White ($n = 2$) were also not included in an aggregated category as it was unclear to which aggregated category they should be assigned. Mixed ethnicity ($n = 5$) such as African American/European American or Cape Verdean/Middle Eastern were not included in any category. Pie graph sections outlined in bold represent agreement between NIH ethnicity and self-identified ethnicity.

applying a Bonferroni-corrected p -value of .00625 indicated that self-identified African-Americans were significantly more likely to skip this item than self-identified ethnic Dominicans ($p < .001$), Hispanic/Latina(o)s ($p < .001$), and Puerto Ricans ($p < .001$) but not significantly more likely than self-identified ethnic Cape Verdeans ($p = .54$), Caribbeans ($p = .77$), Euro-Americans ($p = .014$), Haitians ($p = .86$), or pan-Africans ($p = .45$).

We hypothesized that Hispanic/Latina(o) pan-ethnic individuals would skip the NIH ethnicity item at lower rates than other groups, given that the question is solely focused on Hispanic/Latina(o) ethnicity and is the only place that Hispanic/Latina(o) ethnic *or* racial experience can be indicated on standard NIH forms. Pairwise comparisons using Fisher’s exact test and applying a Bonferroni correction for multiple tests ($p \leq .00833$) indicated that those with Hispanic/Latina(o) pan-ethnic self-identifications [including Dominican, Puerto Rican, and Hispanic/Latina(o)] were significantly more likely to complete the NIH ethnicity item than those self-identifying ethnically as pan-African ($p = .008$), African-American ($p = .001$), Cape Verdean ($p = .001$), Caribbean ($p = .005$), or Haitian ($p = .002$) but did not differ from ethnically self-identified European-Americans ($p = .49$). There were no differences in completion rates between the three pan-ethnic Hispanic/Latina(o) subgroups [Dominican, Hispanic/Latina(o), and Puerto Rican]. With regard to immigrant status, the NIH ethnicity item was missing at similar rates for

immigrant and mainland U.S.-born respondents at 42 and 39 percent, respectively ($\chi^2 = 0.08, p = .77$).

Agreement between Measures. Figure 3 compares responses to the NIH ethnicity and self-identified ethnicity items. Match for ethnicity is difficult to assess, as the NIH item is actually not a general ethnicity item, but is instead a Hispanic/Latina(o)-specific ethnicity item. Among those who responded to both items, fewer than one third (28 percent) of those who endorsed the NIH response “Hispanic or Latina(o)” also self-identified ethnically as generally Hispanic/Latina(o), and an additional 45 percent self-identified as Dominican (22 percent) or Puerto Rican (23 percent). Meanwhile, among those endorsing the NIH “not Hispanic or Latino/a” category, none self-identified specifically as “not Hispanic or Latino/a.”

We examined agreement separately for immigrants and mainland U.S.-born respondents. Agreement between NIH ethnicity items and self-identified ethnicity items was similar across immigrant and mainland U.S.-born respondents. Among immigrant respondents who completed both items, 38 percent of those who endorsed the NIH response “Hispanic or Latina(o)” also self-identified ethnically as Hispanic/Latina(o), and an additional 57 percent self-identified as Dominican (43 percent) or Puerto Rican (14 percent). Meanwhile, among those endorsing the NIH response “not Hispanic or Latino/a,” none self-identified specifically as “not Hispanic or Latino/a.” Similarly, among mainland U.S.-born respondents who completed both items, 36 percent of those who endorsed the NIH response “Hispanic or Latina(o)” self-identified ethnically as Hispanic/Latina(o), and 50 percent self-identified as Dominican (11 percent) or Puerto Rican (39 percent). Among those endorsing the NIH response “not Hispanic or Latino/a,” none self-identified specifically as “not Hispanic or Latino/a.” In all, among respondents completing both items, rates of agreement between NIH ethnicity and self-identified ethnicity were not statistically different across immigrants (18 percent) and mainland U.S.-born respondents (16 percent), $\chi^2 = .06, p = .81$.

DISCUSSION

This study examines methodological concerns regarding the standard approach to measuring race and ethnicity using federally defined categories. This standard approach, which is mandated by the U.S. Office of Management

and Budget's Directive Number 15 and employed in all NIH-funded research, is widespread in its usage in spite of recognized methodological concerns. We evaluated some of these concerns by comparing responses to this NIH measure with responses to open-ended questions of racial and ethnic self-identification in a community sample.

The size and unique nature of our sample—with high rates of Cape Verdean participants, few Asian, Pacific Islander, and Native American respondents, and other community-specific patterns—limit generalizability to other communities. However, the use of a community sample, like the samples studied in much health disparities research, is also a strength; by examining this measure in a community sample we apply a more rigorous, realistic test of the measure's ability to capture participants' racial and ethnic identity even in light of the particularities and specific patterns of a community. The comparison of NIH measures to self-identified race and ethnicity on our open-ended measure is also a strength.

With regard to study limitations, our open-ended measure may be affected by participants' learned responses to race and ethnicity measures from past exposure to surveys with predefined race and ethnicity categories. Moreover, the definition of race provided in the open-ended measure focuses primarily on phenotype (e.g., skin color, facial features) and thus our data are inevitably limited to this particular construction of race. Although researchers agree that race is related to phenotype (e.g., Smedley and Smedley 2005), the social construction of race is complex. It may be impossible to operationalize well the full complexity of the social construction of race, but the approximation we have attempted here allows us to make interpretations about race and racialization.

The results support concerns about the standard approaches to assessing race and ethnicity widely used by health researchers. Questions about race and ethnicity were skipped by 6 and 11 percent of respondents (respectively) regardless of the question format, but the NIH measure resulted in more than twice as much missing data as the self-identification items for both race (26 percent vs. 11 percent) and ethnicity (43 percent vs. 18 percent). Match between NIH item responses and self-identified race and ethnicity was also low. Thus, although this article does not directly test validity, these results raise serious concerns about the NIH measure's construct validity.

Beyond the race and ethnicity items, certain other demographic variables were also missing at high rates, namely income (17 percent missing) and place of birth (20 percent missing). Although neither were missing at the level of the NIH items, these are nonetheless high rates which may reflect concerns

about financial privacy and immigration status; meanwhile, other demographic variables examined (parent education, native language, English proficiency, gender, relationship to child) had lower missingness rates of 8.7 percent on average (range: 7–11 percent). Overall, the elevated missingness rates observed for the NIH race and ethnicity items are unparalleled by other demographic items.

Results also indicate that rates of missingness and misclassification on the NIH measure vary substantially between groups: the NIH measure performs disproportionately poorly for specific groups, including Latina(o)s, Cape Verdeans, and African-Americans. While missingness could reflect an order effect—in which respondents skipped the seemingly repetitive, closed-ended measure because they had already completed the open-ended measure—such an effect would have produced similar rates of missingness across groups. The fact that missingness was disproportionately present in specific groups suggests that it is not due to a mere order effect.

Immigrant respondents were disproportionately likely to have missing race data on the NIH measure relative to mainland U.S.-born respondents. This finding further indicates that national, closed-ended measures may fail to capture the diversity of lived experiences of racism within specific community subgroups—in this case, immigrant groups—and the potential implications of these experiences. Given the profound, multi-generational social adaptation process that accompanies immigration, we would expect differences in concepts of race and ethnicity between immigrants and mainland U.S.-born individuals; research to elucidate the means by which race and ethnicity are incorporated into the immigrant experience would inform how we measure and study racism in the United States.

A full 10 percent of our respondents reported speaking English “very little” or “not at all.” While most of these individuals were Spanish speakers and chose to complete the survey in Spanish, seven of these non-English-proficient individuals (3 percent of the sample) completed the survey in English. For these individuals, whose native languages were Cape Verdean Creole (6) and Haitian Creole (1), the validity of such survey data is limited; this limitation, unfortunately, represents a pattern that occurs frequently in research with linguistically diverse communities.

The lack of a Latina(o) race option in the NIH measure may undermine the measure’s validity by increasing both missingness and misclassification. Missing responses on the NIH race item, and respondents’ decision to instead represent their identity through their responses to the NIH ethnicity item,

may reflect these individuals' identities as ethnically Latina(o)s, but it does not capture experiences of racialization. Misclassification or missing data for these respondents not only affects our ability to understand health disparities related to Latina(o) experiences but also clouds our understanding of the White and Black/African-American groups. If Latina(o)s' racialization and experiences with racism do indeed relate to increased risk or prevalence, then their inclusion in the White group will raise that group's risk or prevalence, thus obscuring important differences related to racism not only between Whites and Hispanic/Latina(o)s but also in comparisons between Whites and Blacks or other racial minorities. This is particularly problematic in samples such as ours, where the majority of those endorsing White are actually racially self-identified Latina(o)s. From a statistical perspective, this pattern demonstrates how such misclassification of race and ethnicity creates measurement error that may bias findings regarding health disparities toward the null.

Our findings also suggest that measuring ethnicity as only Hispanic/Latina(o) or non-Hispanic/Latina(o) is problematic. The fact that nearly half of self-identified ethnic African-Americans skipped the NIH ethnicity question signifies that the measure is not capturing their ethnicity-related experiences. Operationalizing ethnicity in this dichotomous way also prevents researchers from understanding ethnic differences within racialized groups; for example, our findings suggest that the NIH ethnicity item misses more specific or nuanced ethnic identifications provided by many Latina(o)s, including self-identified Puerto Ricans and Dominicans.

The poor differentiation of race and ethnicity and the lack of an inclusive ethnicity question that captures more than pan-ethnic Hispanic/Latina(o) ethnicity also combine to undermine the NIH measure's utility. Immigrants from Brazil and other parts of South or Central America may self-identify or be identified by as others as Latina(o) *racially*, but they may not *ethnically* identify as such (Zubaran 2008; McDonnell and de Lourenço 2009). These respondents would be misrepresented by the current measure; they may skip the NIH race item because there is no appropriate response but also skip the ethnicity item because their ethnicity as Brazilian cannot be captured. Similarly, because the "Black or African-American" NIH race option confounds race (Black) and ethnicity (African-American), individuals who identify as Black but are not ethnically African-American may be poorly captured. Twenty percent of the Cape Verdeans in our sample skipped the NIH race item, and 60 percent skipped the NIH ethnicity item. Their completion might be higher if African-American ethnicity was not confounded with Black race in the response options, and if there were an

ethnicity option that enabled Cape Verdeans to differentiate themselves from African-Americans. While Cape Verdeans represent a small proportion of the population nationwide, there are over 37,000 Cape Verdeans in the community we sampled (US Census Bureau 2010); the measure's poor performance with this group indicates an inability to capture experiences of regionally specific groups of interest to public health researchers. The ability to capture the identities of such ethnic groups is essential given the established differences in diet, health behaviors, and health beliefs among ethnic groups and, in turn, disparities in health that are attributable to ethnic, not racial, differences (Castro, Shaibi, and Boehm-Smith 2009).

CONCLUSIONS

Our results inform current policies and practices regarding the collection of race and ethnicity data by health researchers, health plans, and health care providers. These findings indicate that utilizing a measure that conceptually separates race and ethnicity, and that provides respondents with adequate flexibility to identify themselves both racially and ethnically, decreases missing data and misclassification and, as such, may increase validity. A more detailed, granular measure of race and ethnicity would enhance validity. Employing an open-ended response approach requires time-consuming coding and is impractical for large-scale public health research; nevertheless, operationalization of a closed-ended, multiple-choice measure can be improved. Feasible changes might include the following: (1) adding a Hispanic/Latina(o) race category; (2) differentiating Black racialization from African-American ethnicity by rewording the response to only "Black"; and (3) designing an ethnicity variable that captures greater variability in responses, beyond the Hispanic/Latina(o) versus non-Hispanic/Latina(o) dichotomy, such as offering various, fine-grained ethnicity options that are tailored to the community of interest or an extensive menu of options organized by geographical region.²

Finally, even self-identified race is not necessarily the best indicator of one's racialization and experiences of racism; although the impact of ascribed race is not well understood, how others ascribe one's race also affects individuals' racialized experiences (Landale and Oropesa 2002; McDonnell and de Lourenço 2009). Broadly, future research using improved measurement tools may better elucidate race- and ethnicity-related patterns in health and may inform the development and evaluation of health prevention and intervention efforts.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: We wish to thank the research assistants who worked on data collection and processing, including Diana Cortes, Stephanie Moronta, and Gavin O'Brien, graduate student Nicholas Mian and Professor Alice Carter for their oversight of data collection and Professor Carter's initial contribution to conceptualization, and graduate student Hillary Hurst Bush for her initial work in data aggregation. Partial funding for this study came from the Horizon Center at the University of Massachusetts Boston, funded by the National Institute on Minority Health and Health Disparities (NIMHHD) of the NIH under Award Number P20 MD002290-05. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. An early version of these findings was presented at the 2012 Annual Convention of the American Psychological Association. The authors do not declare any potential personal conflicts of interest.

Disclosures: None.

Disclaimers: None.

NOTES

1. Percents do not sum to 100 percent due to rounding.
2. We have developed such a measure (Suyemoto et al. 2012) and are currently using it in a variety of research studies. Contact karen.suyemoto@umb.edu.

REFERENCES

- Afshari, R., and R. S. Bhopal. 2002. "Changing Pattern of Use of 'Ethnicity' and 'Race' in Scientific Literature." *International Journal of Epidemiology* 31 (5): 1074.
- Betancourt, J. R., A. R. Green, J. E. Carrillo, and O. Ananeh-Firempong. 2003. "Defining Cultural Competence: A Practical Framework for Addressing Racial/Ethnic Disparities in Health and Health Care." *Public Health Reports* 62 (1): 293–302.
- Brondolo, E., L. C. Gallo, and H. F. Myers. 2009. "Race, Racism and Health: Disparities, Mechanisms, and Interventions." *Journal of Behavioral Medicine* 32 (1): 1–8.
- Campbell, M. E., and C. L. Rogalin. 2006. "Categorical Imperatives: The Interaction of Latino and Racial Identification." *Social Science Quarterly. Special Issue: Ethnicity and Social Change* 87 (5): 1030–52.

- Castro, F. G., and E. Hernández-Alarcón. 2002. "Integrating Cultural Factors into Drug Abuse Prevention and Treatment with Racial/Ethnic Minorities." *Journal of Drug Issues* 32: 783–810.
- Castro, F. G., G. Q. Shaibi, and E. Boehm-Smith. 2009. "Ecodevelopmental Contexts for Preventing Type 2 Diabetes in Latino and Other Racial/Ethnic Minority Populations." *Journal of Behavioral Medicine* 32 (1): 89–105.
- Clark, R., N. B. Anderson, V. R. Clark, and D. R. Williams. 1999. "Racism as a Stressor for African Americans: A Biopsychosocial Model." *American Psychologist* 54 (10): 805–16.
- Cokley, K. O. 2005. "Racial(Ized) Identity, Ethnic Identity, and Afrocentric Values: Conceptual and Methodological Challenges in Understanding African American Identity." *Journal of Counseling Psychology* 52 (4): 517–26.
- Deaux, K., N. Bikmen, A. Gilkes, A. Ventuneac, Y. Joseph, Y. A. Payne, and C. M. Steele. 2007. "Becoming American: Stereotype Threat Effects in Afro-Caribbean Immigrant Groups." *Social Psychology Quarterly* 70 (4): 384–404.
- Ford, M. E., and P. A. Kelly. 2005. "Conceptualizing and Categorizing Race and Ethnicity in Health Services Research." *Health Services Research* 40 (5 Pt 2): 1658–75.
- Frank, R., I. R. Akresh, and B. Lu. 2010. "Latino Immigrants and the U.S. Racial Order: How and Where Do They Fit In?" *American Sociological Review* 75 (3): 378–401.
- Hahn, R. A., and D. F. Stroup. 1994. "Race and Ethnicity in Public Health Surveillance: Criteria for the Scientific use of Social Categories." *Public Health Reports* 109 (1): 7–15.
- Harrell, J. P., S. Hall, and J. Taliaferro. 2003. "Physiological Responses to Racism and Discrimination: An Assessment of the Evidence." *American Journal of Public Health* 93 (2): 243–8.
- Horowitz, C. R., K. A. Colson, P. L. Hebert, and K. Lancaster. 2004. "Barriers to Buying Healthy Foods for People with Diabetes: Evidence of Environmental Disparities." *American Journal of Public Health* 94 (9): 1549–54.
- Jones, C. P. 2001. "Invited Commentary: 'Race', Racism, and the Practice of Epidemiology." *American Journal of Epidemiology* 154 (4): 299.
- Kaplan, J. B., and T. Bennett. 2003. "Use of Race and Ethnicity in Biomedical Publication." *Journal of American Medical Association* 289 (20): 2709–16.
- Klonoff, E. A., H. Landrine, and J. B. Ullman. 1999. "Racial Discrimination and Psychiatric Symptoms among Blacks." *Cultural Diversity and Ethnic Minority Psychology* 5 (4): 329–39.
- Landale, N. S., and R. S. Oropesa. 2002. "White, Black or Puerto Rican? Racial Self-Identification among Mainland and Island Puerto Ricans." *Social Forces* 81 (1): 231–54.
- Laws, M. B., and R. A. Heckscher. 2002. "Racial and Ethnic Identification Practices in Public Health Data Systems in New England." *Public Health Reports* 117 (1): 50–61.
- Lee, S. S., J. Mountain, and B. A. Koenig. 2001. "The Meanings of 'Race' in the New Genomics: Implications for Health Disparities Research." *Yale Journal of Health Policy, Law, and Ethics* 1: 33–75.

- Markus, H. 2008. "Pride, Prejudice, and Ambivalence: Toward a Unified Theory of Race and Ethnicity." *American Psychologist* 63: 651–70.
- McDonnell, J., and C. de Lourenço. 2009. "You're Brazilian, Right? What Kind of Brazilian Are You? The Racialization of Brazilian Immigrant Women." *Ethnic and Racial Studies* 32 (2): 239–56.
- McKenzie, K., and N. S. Crowcroft. 1994. "Race, Ethnicity, Culture, and Science." *British Medical Journal* 309: 286–7.
- National Institutes of Health. 2001. "Toward Higher Levels of Analysis: Progress and Promise in Research on Social and Cultural Dimensions of Health: NIH Publication No. 21-5020." [accessed on May 22, 2012]. Available at http://obsr.od.nih.gov/pdf/HigherLevels_Final.PDF
- Nazroo, J., J. Jackson, S. Karlsen, and M. Torres. 2007. "The Black Diaspora and Health Inequalities in the US and England: Does Where You Go and How You Get There Make a Difference?" *Sociology of Health and Illness* 29 (6): 811–30.
- Ramírez, M., M. E. Ford, A. L. Stewart, and J. A. Teresi. 2005. "Measurement Issues in Health Disparities Research." *Health Services Research* 40 (5 Pt 2): 1640–57.
- Rivara, F. P., and L. Finberg. 2001. "Use of the Terms Race and Ethnicity." *Archives of Pediatric and Adolescent Medicine* 155 (2): 119.
- Smedley, A., and B. D. Smedley. 2005. "Race as Biology is Fiction, Racism as a Social Problem Is Real: Anthropological and Historical Perspectives on the Social Construction of Race." *American Psychologist* 60: 16–26.
- Suyemoto, K. L., L. Roemer, S. M. Erisman, D. W. Holowka, C. Fuchs, and H. Barrett-Model. 2012. *UMass Boston Comprehensive Race and Ethnicity Questionnaire, Revised*. Unpublished measure.
- US Census Bureau. 2010. "United States Census 2010" [accessed on May 1, 2012]. Available at http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_10_1YR_B04006&prodType=table
- U.S. Department of Health and Human Services. 2013. "2012 National Healthcare Disparities Report AHRQ Publication No. 13-0003." Rockville, MD: Agency for Healthcare Research and Quality [accessed on August 23, 2013]. Available at <http://www.ahrq.gov/research/findings/nhqdr/nhdr12/2012nhdr.pdf>
- Vaquera, E., and G. Kao. 2006. "The Implications of Choosing 'No Race' on the Salience of Hispanic Identity: How Racial and Ethnic Backgrounds Intersect among Hispanic Adolescents." *The Sociological Quarterly* 47: 375–96.
- Waters, M. C. 1994. "Ethnic and Racial Identities of Second-Generation Black Immigrants in New York City." *International Migration Review* 28 (4): 795–820.
- Zubaran, C. 2008. "The Quest for Recognition: Brazilian Immigrants in the United States." *Transcultural Psychiatry* 45 (4): 590–610.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.