

Sources of Nonreplicability in Aging Ethnoracial Health Disparities
Research

Ian M. McDonough – University of Alabama

Erin R. Harrell – University of Alabama

Sheila R. Black – University of Alabama

Rebecca S. Allen – University of Alabama

Patricia A. Parmelee – University of Alabama

Deposited 03/08/2021

Citation of published version:

McDonough, I., Harrell, E., Black, S., Allen, R., Parmelee, P. (2021): Sources of Nonreplicability in Aging Ethnoracial Health Disparities Research. *Nonreplicability in Health Disparities Research*.

This is an accepted manuscript. The fully published version can be found at:

DOI: <https://doi.org/10.1037/pag0000606>

Sources of Nonreplicability in Aging Ethnoracial Health Disparities Research

Ian M. McDonough^{1,2}

Erin R. Harrell^{1,2}

Sheila R. Black^{1,2}

Rebecca S. Allen^{1,2}

Patricia A. Parmelee^{1,2}

¹The University of Alabama, Department of Psychology

²Alabama Research Institute on Aging

Author Note

Correspondence should be addressed to Ian M. McDonough, The University of Alabama, Department of Psychology, 505 Hackberry Lane, BOX 870348, Tuscaloosa, AL 35487 USA

Email: immcdonough@ua.edu, Phone: 205-348-2808; Fax: 205-348-8648. All individuals have agreed to be an author. The authors declare no conflicts of interest.

Abstract

The older adult population in the U.S. is becoming increasingly diverse across a constellation of factors including ethnoracial group, socioeconomic status, and immigration status. However, our understanding of the consequences of this diversity for cognitive and mental health is masked by the lack of inclusion of diverse sample characteristics, the use of assessments that might hold a different meaning for different groups of people, and analytical choices that do not probe the impact of diverse characteristics or assume an unwarranted degree of homogeneity within groups. Each of these factors not only hinders our ability to understand various psychological mechanisms that differ as a function of age, but also threatens the likelihood of replicability across aging research studies. This paper provides our perspective on three key sources of nonreplicability in ethnoracial health disparities research among older adults: 1) what is lost in creating monolithic groups rather than identifying subgroups of minorities, 2) understanding aging from the perspective of intersecting identities, and 3) biases of research materials. We also provide recommendations to increase replicability in aging research with respect to the challenges outlined. Approaching questions on aging from a health disparities lens, can both increase the generalizability of research outcomes and improve initiatives of social justice that are long overdue.

Keywords: cognition, mental health, intersectionality, mixed methods, latent class analysis

Sources of Nonreplicability in Aging Ethnoracial Health Disparities Research

As the U.S. population ages, ethnoracial minorities are projected to become the fastest growing group of older adults. Differences in health outcomes between ethnoracial minorities and non-Hispanic Whites (the majority) have been defined broadly as health disparities. Some of these disparities in the U.S. can be traced to factors including inequitable structural policies (e.g., in housing, policing, education; Salter, Adams, & Perez, 2018). For example, housing policies such as redlining and exclusionary zoning has downstream negative effects of neighborhood funding, access to fresh and high-quality food, and access to healthcare—all of which can affect mental and cognitive health (Swope & Hernández, 2019). Other sources of disparities stem from historical distrust of the medical and research communities following events like the Tuskegee Syphilis Study (Freimuth et al., 2001), and lack of cultural competence of the health care workforce such as pain management (Betancourt et al., 2003). These historical roots set the stage both for negative cognitive and mental health among older ethnoracial minorities and for a frank lack of scientific understanding of those potential negative outcomes. Some researchers have suggested that ethnoracial diversity may be a key contributing factor to nonreplicability of various psychology findings (Yip et al., 2018). When adding the heterogeneity of the aging process to existing differences among ethnoracial groups, one can begin to appreciate the complicated challenge of approaching age-related questions in health disparities research. Here, we provide perspective on sources of nonreplicability in ethnoracial health disparities research among older adults by focusing on three key issues: 1) what is lost in creating combined subgroups of minorities into monolithic groups, 2) understanding aging from the perspective of intersecting identities (including age/cohort characteristics), and 3) biases of research materials. Following explication of these issues, we provide recommendations to increase replicability in health disparities research and more general aging research.

Identifying Subgroups of Ethnoracial Minorities: What is Lost in Creating Monolithic Groups?

A large body of aging research, especially in experimental psychology, has not historically considered ethnoracial category in conceptualization, recruitment, or analysis (Lewis & van Dyke, 2018). Of those that have, dichotomous categories are often used that compare large ethnoracial groups (e.g., “Asian,” “Black”) to non-Hispanic Whites. Although well intentioned, the grouping of ethnoracial groups discounts differences within these groups that can obscure meaningful patterns. Although this method supports statistical power and simple explanations, this assumed homogeneity might be one major source of nonreplicability. Treating diverse minority groups monolithically assumes shared ethnic heritage that may not exist. For example, as compared to those who immigrated to the U.S., persons of African descent born as U.S. citizens often diverge in culture and experiences such as socialization, education, diet, and spiritual beliefs. Yet studies rarely distinguish between smaller, meaningful subgroups such as people of African descent who immigrated to the U.S, people of African descent born in America, and Caribbean Blacks who may have been born in the U.S. or in the Caribbean and are of West Indian or Caribbean descent. Such monolithic grouping is not limited to Blacks. Asian Americans represent over 50 different ethnic groups with more than 100 different languages and a diverse set of characteristics such as country of origin, religion, generational status, duration of residence in the U.S., and socioeconomic indicators (Islam et al., 2010).

Creating monolithic groups not only reduces generalizability but risks excluding findings that can provide insight into factors that may be important for health and well-being among subgroups within a population. One finding that has long gone unnoticed due to the treatment of Hispanics as a single group is the “happiness paradox” among foreign-born Hispanic individuals (Calvo et al., 2017). Despite the fact that older Hispanic immigrants often face more inequalities than other ethnoracial groups, many report greater life satisfaction than native-born Hispanics and non-Hispanics in the U.S. Similarly, a study exploring life satisfaction among older Turkish immigrants observed that first-generation immigrants reported higher levels of happiness than counterparts from the same region of origin who never left home (Baykara-Krumme & Platt, 2018), again emphasizing the importance of considering

immigration status. Researchers have suggested that unique cultural factors that emphasize social integration and collectivism (the needs, goals, views, and beliefs of the group are emphasized over those of the individual) may be contributing to these paradoxical outcomes (Schwartz et al., 2010). Social integration factors such as connection to religious groups also have been posited as an explanation for “the race paradox” in which non-Hispanic Black Americans are as happy as or even happier than Whites—an unexpected outcome given that Black Americans are more socio-economically disadvantaged and in poorer physical health than Whites (Tang et al., 2019).

In accordance with socio-emotional selective theory (SST; Carstensen & Löckenhoff, 2003), these groups may be choosing to immerse themselves in religious or volunteer groups even as the larger social network contracts. The positive affect that results from engaging in valued social exchanges may increase social support that older individuals may need to carry out activities of daily living (Baltes, 1997). Although SST has been shown to apply to all older adults, older ethnic minorities may exhibit greater socio-emotional selectivity because many may be at higher risk of cognitive and functional decline. For example, dementia risk is reported to be twice as high for Black Americans compared to non-Hispanic Whites (Plassman et al., 2007; Shadlen et al., 2006; Yaffe et al., 2013). In fact, dementia risk is a good example of the problems of generalizing across minority subgroups. When dementia incidence was investigated among Asian American subgroups, dementia incidence was 20% higher among Filipino Americans than other Asians such as Chinese Americans (Mayeda et al., 2017). The increase risk was attributed in part to distinctive cultural traditions, years of exposure to U.S. culture, and varying degrees of social integration (Mayeda et al., 2017).

In a similar vein, research has shown that when subgroup membership and socioeconomic predictors are included, a different picture emerges in regard to ethnoracial differences in mental disorders and self-reported mental health (Kim et al., 2012; Jang et al., 2014). Kim et al. (2012) highlighted misdiagnosed psychiatric disorders as a result of the common practice of lumping all Asian Americans together. Across Filipino, Vietnamese, and Chinese samples, Vietnamese Americans rated

their mental health as poorest, which was attributed to the uniquely stressful circumstances that led to their moving to the United States (Kim et al., 2012). Research has demonstrated similar variability in mental health within subgroups of Black Americans. Black Americans with Generalized Anxiety Disorder (but not Major Depressive Disorder) self-reported poorer overall mental health whereas Caribbean Blacks with Major Depressive Disorder (but not Generalized Anxiety Disorder) reported poorer mental health (Assari, Dejman, & Neighbors, 2016). These findings suggest that the increasing diversity of the population demands appreciation of subtle cultural differences among subsets of individuals who have previously been treated as a homogenous group.

By creating monolithic groups, researchers also risk displacing or ignoring individuals who identify with more than one ethnoracial group because they do not neatly fit into one category. Identifying with more than one race presents sociocultural challenges in regard to how people are accepted, how they are treated in healthcare settings, and how they might attempt to avoid racism and discrimination (Albuja, Sanchez, & Gaither, 2018; Davenport, 2016). Creating monolithic groups in data collection methods can inadvertently promote distrust of our research methods and, in some instances, motivate individuals to engage in “racial passing” such that they identify with a single ethnoracial group—the choice of which might be made out of fear for how they will be treated based on how they identify (Albuja et al., 2018). In some instances, this fear may dissuade participants from diverse groups from participating in research, once again limiting the data we are able to collect and the conclusions that can be drawn.

Understanding Aging from the Perspective of Intersecting Identities

In addition to the importance of identifying subgroups, understanding how intersecting identities may influence research outcomes might inform sources of nonreplicability. One’s current age or cohort represents one of many potential identities that might intersect to affect cognition and mental health. We refer to age/cohort as the influence of historic time and place on attitudes and behaviors in older adults (Elder, 1998). Many factors already mentioned earlier (e.g., ethnoracial group, immigration status,

neighborhood, age group) have been recognized as a source of stigma (Cary & Chasteen, 2015). These compounding identities can serve as multiple sources of structural inequality and psychological distress, which in turn can be detrimental to various domains of function (cognition, health, vitality).

Although health disparities research has recently recognized the importance of intersecting identities that lead to compound inequalities (e.g., women of color who identify as transgender), the addition of age often has seldom been included in these notions of intersectionality. Intersecting identities interact simultaneously and dynamically to influence one's experiences within society (Bauer, 2014; Bowleg, 2012; Earnshaw et al., 2018). Bowleg (2012) identified three primary tenets of an intersectional perspective relevant to public health, each of which may be a source of nonreplicability: 1) social identities are not independent, 2) people from multiple historically oppressed and marginalized groups must be considered prior to perspectives of the majority, and 3) multiple social identities at the micro level intersect with macro level structural factors to produce disparate health outcomes. Although not all intersecting identities might be of sufficient value to merit study (Bauer, 2014), more research on the impact of various identities in older adults is needed to tease apart the combinations of most value from those that have little impact.

Even though older adults across ethnoracial groups may share cohort similarities (e.g., having shared health concerns regarding cognitive functioning and independence), health and socioeconomic differences among minority groups give rise to unique barriers and approaches to ameliorating these concerns. Some key intersecting identities that can affect barriers are levels of income, education, and access to healthcare. Compared with non-Hispanic Whites, older Hispanics and Black Americans often (but not always) have lower incomes and more cognitive and functional limitations (Jackson et al., 2010). These two groups also are more likely to report not having health care coverage, healthcare savings, or advanced care planning (Choi et al., 2019, 2020). While only 15% of non-Hispanic Whites reported cost as a barrier to accessing healthcare, sizeable proportions of non-Hispanic Blacks (24.5%), Hispanics (23.1%), and American Indian/Alaskan Native (19.1%) said that they could not afford to see a

physician when needed (James et al., 2017). As discussed above, these group averages do not apply evenly to all group members. Instead, such barriers are more likely when for ethnoracial minorities that have intersecting identities with other factors associated with structural inequality (e.g., access to housing, education, and employment).

The intersectionality of being an ethnic and older adult sexual minority places one at greater risk of stigmatization, discrimination, and victimization (Emlet, 2016). Fredriksen-Goldson et al. (2014) showed that transgender older adults have increased risk for poverty, financial barriers to healthcare access, and greater internalized and enacted stigma compared to their LGB peers. They also had higher rates of disability and lower overall social support (Fredriksen-Goldson et al., 2013). This decreased level of social support might underlie findings of higher rates of health risk behaviors. Studies have shown excessive drinking among lesbian and bisexual women compared to their heterosexual peers (Dilley et al., 2010; Conron, Mimiaga, & Landers, 2010); differences have also been found with gay and bisexual men, who are more likely to smoke and drink excessively compared to their heterosexual peers (Wallace et al., 2011).

One health challenge that has come out of this stigmatization is the rate of HIV status among this community. Although HIV status impacts psychological distress regardless of sexual orientation or gender identity, 70 percent of Americans with HIV are expected to be 50 and older in 2020; thus, the disease, disproportionally affects older adults (APA.org). Older adults living with HIV are significantly more likely to have an Axis I mental health diagnosis, substance abuse issues, hypertension, diabetic mellitus, arthritis, and lower physical functioning (Beatie, Mackenzie, & Chou, 2015). HIV among older adults also can affect multiple domains of cognition and everyday functioning, including medication adherence, instrumental activities of daily living, and driving a vehicle (Cody & Vance, 2016).

Biases of Research Materials

The increasing older population presents an important problem with regard to psychometric assessments (Mindt et al., 2013). Most assessments for older adults were normed on predominantly non-

Hispanic White samples. As described by Helms (1992), such exclusive construction of assessments implicitly determines a) what constitutes factors of intellectual ability, b) what language standards are used in materials and test administration, c) the cultural relevance of pictures, words, and names used in tests, d) what setting is optimal including the characteristics of the tester, e) that taking tests is a valuable approach to assessing cognition, f) and even who chooses what a “correct” answer should be. These factors are important to consider when interpreting cognitive and mental health assessments because they represent a cultural bias. Cultural bias refers to a test yielding significantly different results among subgroups of the population with similar ability levels and is often illustrated via Differential Impact Functioning (DIF; Ng et al., 2018).

In fact, when intelligence testing was being standardized in the 1940s and 1950s, psychometric tests were indiscriminately administered to Slavic immigrants from non-English speaking countries; these immigrants unfortunately bore the label “feeble minded” for poor test performance due to language barriers (Samelson, 1982). Thus, accounting for sociocultural factors can impact the sensitivity and specificity of an assessment—especially assessment of cognitive functioning. Sensitivity refers to a test being able to detect group membership or outcome (e.g., dementia) and specificity refers to the extent to which the test correctly indicates the absence of group membership or outcome (e.g., accurately indicating the absence of dementia). In regard to cognitive tests related to dementia specifically, the sensitivity and specificity of most tests are poor for people of color relative to non-Hispanic Whites (Freedman & Manly, 2018). These poor sensitivity and specificity rates are problematic from both clinical and scientific perspectives. Clinically, the lack of accuracy for people of color has led to a delay in a diagnosis of dementia until the latter stages of the disease that, in turn, can impact public health decisions (Avila et al., 2020; Bertola et al., 2020). As a result, these individuals lose the opportunity to engage in financial and healthcare planning. From a scientific perspective, understanding the different experiences of cognitive decline can help identify alternative pathways or subtypes of disease states.

Evidence of Ethnoracial Differences in Cognitive Test Performance

Given the possibility that cognitive assessments are culturally biased, what evidence exists that performance on standardized cognitive assessments can be treated at face value between ethnoracial groups? Several studies have addressed this issue by testing measurement invariance of cognitive measures between groups. A test is invariant when it demonstrates that it measures the same construct, has the same relationships to the construct, and has similar residual variance scores between groups (Avila et al., 2020; Bertola et al., 2020; Jones, 2003). Although several statistical tests can assess measurement invariance, structural equation modeling (SEM) is currently the most popular. Using a large national data set, Jones (2003) tested measurement invariance between Black and White groups in the modified version of the Telephone Interview for Cognitive Status. Older Black participants performed worse across the cognitive tests than non-Hispanic White participants. Critically, while 13% of the of observed racial differences in cognitive scores were attributed to social determinants (e.g., sex, age, SES, health behaviors, and health conditions), 73% of the differences could be attributed to measurement artifact (after controlling for social determinants).

This finding suggests that the observed test scores do not measure the construct in the same way. Jones (2003) also found that some of the relationships between environmental/demographic factors and cognition differed between the two groups. For example, whereas lower education and smoking had a greater negative impact on cognition in Blacks than Whites, greater income benefitted cognition more in Whites than Blacks. The tasks that showed the greatest measurement bias largely involved language/semantic tasks (e.g., naming) and mental arithmetic (i.e., counting backward by seven).

Similar violations of measurement invariance have been found in language tasks (i.e., naming and word similarities) between non-Hispanic Whites and Hispanics (Avila et al., 2020). Differences also have been found on tasks often considered to be indicative of non-language domains like Trail Making Test B, which is used to test executive function or processing speed differences (Bertola et al., 2020).

However, even this task requires fluently processing the order of the alphabet (i.e., a language ability) to do well. Thus, cultural biases, often through language, might influence the validity of cognitive tests.

Factors to Consider when Interpreting Psychological Assessments

Increasingly, researchers have recognized the importance of considering socio-historical factors when interpreting psychometric assessments (Garcia et al., 2017). Below we outline several key factors that may impact the interpretation of cognitive assessments: health, education, cognitive reserve, acculturation, discrimination, and perceived stress.

Health

Black Americans, Native Americans, and Hispanics suffer from diabetes, hypertension, kidney disease, and heart disease at greater rates than non-Hispanic Whites. All of the aforementioned diseases are associated with a decline in cognition. For example, diabetes is associated with decrements in memory, language, processing speed, executive functioning, and visuospatial abilities (Moheet, Mangia, & Seaquist, 2015; Ruis et al., 2009). Heart failure is associated with a decline in executive functioning and processing speed (Cameron, 2016). Thus, elevated levels of chronic diseases in some ethnoracial minorities might be one factor underlying concomitant ethnoracial differences in cognition (Tang et al., 2019). Notably, deficits in cognitive functioning due to these chronic health conditions can be independent from cognitive deficits due to dementia. For example, some researchers argue that heart failure is associated with cognitive impairment, but the pattern of impairments is different from dementia (Cameron et al., 2016). Other researchers argue that many of the aforementioned chronic health conditions are risk factors for dementia (Deckers et al., 2015; Livingston et al., 2017). Thus, health might serve as a mediating factor of apparent ethnoracial group differences.

Education and Cognitive Reserve

Systemic racism has resulted in marked differences in the quality of education as well as the level of educational attainment in Black Americans (Barnes et al., 2016). Due to Jim Crow, Black Americans living in the South were required to attend segregated high school during the 1940s, 1950s,

and 1960s. Black Americans living in the South also were prohibited from attending a number of major universities (Brown & Davis, 2009). In the North, although no laws mandated segregation, schools were for the most part still segregated (Logan, 1933). In addition, schools serving children of color throughout the country have historically received less funding than schools serving White students (Brown & Davis, 2009).

Studies also have emphasized the importance of quality of education, thereby moving beyond simple metrics such as highest degree obtained or years of education. Manly and colleagues (2004) found that when Blacks and Whites were matched on reading level (a proxy for quality of education), no differences in cognition were found. Although accounting for reading level does not always explain ethnoracial differences in cognitive performance, other studies have shown that measures of educational quality at least attenuate the differences (Crowe et al., 2013; Fyffe et al., 2011). Notably, these differences in education coincide with the differences in language and domains relying on language found between ethnoracial groups summarized in the previous section.

Education does more than just confer knowledge. It is associated with the development of new cognitive skills which may ultimately lead to cognitive reserve in old age (Contador et al., 2017; Stern, 2009). Cognitive reserve is a proxy for the brain's capacity to withstand neuropathological damage, or the ability to utilize a reservoir of cognitive processing mechanisms to compensate for decline in cognition. Cognitive reserve theory assumes that education and exposure to cognitively complex tasks facilitate more efficient cognitive operations and alternative cognitive networks that allow the brain to compensate for pathology that might occur during the course of aging. Implicit in these ideas is that early life barriers in access to quality education and resources in ethnoracial minorities can prevent building cognitive reserve in other life stages (e.g., via wealth, job complexity, mental engagement; Glymour & Manly, 2008). This limitation of access has downstream consequences throughout the adult lifespan, especially in older age when cognitive reserve is called upon to compensate for the presence of vascular burden or pathology. Note that research also is beginning to show that such early life buildup of

cognitive reserve can reduce the likelihood of accumulating vascular injury or pathology in the first place (Binette et al., 2020; Wirth et al., 2014).

Acculturation

Acculturation refers to the extent to which individuals adopt the language, customs, and values of their own culture versus the values, customs, and language of the dominant culture (Helms, 1992; Freedman & Manly, 2018). Acculturation affects cognitive style, test strategy, “testwiseness,” and familiarity with test items. Thus, test takers from minority groups with higher levels of acculturation to the culture on which the test is normed perform better on verbal and nonverbal tests in comparison with individuals whose cultural experiences are less mainstream and more bound to their culture of origin. Presumably, individuals from minority groups with more exposure to the dominant culture or who have successfully learned to adapt to the dominant culture are more familiar with test items or the terms used to describe those items typically included in assessment measures and are more likely to have adopted the “cognitive style” associated with optimal performance on assessment tests (Barnes, 2016; Freedman & Manly, 2018; Helms, 1992). As described earlier, the experiences of ethnoracial groups are quite variable and this variation extends to acculturation among Black Americans, Hispanics, and other subpopulations in America (Freedman & Manly, 2018).

Acculturation also is a critical factor to consider for immigrants to the US. From this perspective, acculturation occurs when individuals incorporate the customs, worldviews, and mores of a foreign culture into their own cultural system. Different acculturation styles have been described based on the degree to which immigrants have embraced their native or host culture (Bernstein et al., 2011; Berry, 2003). The acculturation style adopted is influenced by multiple factors. In many cases, language proficiency in the host culture may be a critical factor. To the extent that the dominant language is not learned proficiently, less acculturation and greater discrimination may occur. As mentioned earlier, acculturation to the dominant language is an important factor to consider when interpreting cognitive test performance. Lack of acculturation is associated with lower test performance in part because of

assessment bias. For example, older Latino Americans with high levels of acculturation performed better on cognitive measurements than Latinos with lower levels of acculturation (Martinez-Miller et al., 2020).

The degree of integration and/or assimilation often is dependent on age. Older immigrants are more likely to resist assimilation than younger immigrants (Cheung, 1989), in part because older immigrants find it more difficult to adopt the language and practices of the host country. In many instances, they are able to find a cultural enclave in which they may speak their native language and enjoy the foods from their heritage culture. For example, many older Chinese immigrants prefer to limit their interactions to Chinatown where they can speak their native language and observe familiar customs (Cheung, 1989). Another obstacle for older immigrants is the shift in values. Many immigrants of Asian, African, and Hispanic descent have migrated from collectivist cultures in which societal norms require that the family and community take care of old people (Albertini et al., 2019; Gao et al., 2019). As younger family members who are also immigrants become acculturated, tension sometimes arises.

Perceived Stress and Discrimination

Chronic stress has been clearly documented to impact cognition and mental health (Shields & Slavich, 2017). Discrimination is one type of chronic stress that is frequently faced by ethnoracial minorities that can be further exacerbated by intersecting identities. In one of the first studies to document this relationship in cognitive aging, Barnes et al. (2012) showed that discrimination among older Black Americans was associated with lower cognitive performance, especially episodic memory. Beatty Moody et al. (2019) found a significant positive correlation between reported levels of discrimination and white matter lesion volumes among Black Americans. White matter lesions are associated with strokes, dementia and cognitive decline (Pantoni, Poggesi, & Inzitari, 2007).

Solutions to Understand Aging Subgroups and Enhance Replicability

An understanding of health disparities requires integration of multiple determinants, which can be challenging to quantify using traditional methods. The NIH Health Disparities Research Framework

laid out 62 factors that may fuel disparities among older adults (Hill et al., 2015). Given this complexity, sophisticated methods and analytic strategies are needed. Most importantly, scientists must be open to flexible approaches to these issues including altering current assessments, considering qualitative or mixed methods, and person-centered techniques to discover and address research questions that enhance replicability in an increasingly diverse population.

Strategies for Breaking Apart Monolithic Groups

Strategies for breaking apart monolithic groups, while well-intended, can pose a number of challenges. One key strategy is to understand the magnitude of heterogeneity and the importance in evaluating the validity and interpretation of results (Mayeda et al., 2017). Because many researchers may begin with clear *a priori* hypotheses regarding health disparities, they can explicitly acknowledge the reporting of exploratory hypotheses to test for generalizability across diverse groups and briefly outline the importance of such tests (cf. Dotson & Duarte, 2020). The appropriateness of sample sizes should be considered in analyses and trends reported if the numbers of participants do not reach appropriate power as required by a power analysis. We also acknowledge that defining subgroups can be a challenge; definitions of groups will change over time with new research and current sociocultural climate of the nation and, frankly, will not satisfy everyone. As a start, though, researchers should review the current distinctions as determined by the U.S. Census or the National Institutes of Health. For example, the 2020 U.S. Census breaks down Hispanics/Latinx by the following: those whose origin is Mexican, Puerto Rican, Cuban, Spanish-speaking Central or South American countries, or “other Hispanic/Latino, regardless of race” (U.S. Census Bureau, 2019). While a checkbox approach can be sufficient, also having an “other” with an open-ended option for participants to explain their identity further would provide clarification for any ambiguities or disagreements with the choice of options. It may also be helpful to include questions about parents’ race/ethnicity to help pinpoint multiracial/multiethnic group identity.

Once group categories are defined, recruitment of participants representing these categories remains another challenge. How do we avoid intentional oversampling to obtain adequate representation of any given subgroup? One way is to engage in purposeful sampling. In contrast to probabilistic or random sampling, which is used to ensure the generalizability of findings, purposeful sampling involves identifying and selecting individuals or groups of individuals that are especially knowledgeable about or experienced with a phenomenon of interest (Creswell & Plano Clark, 2011; see later section on Qualitative Inquiry and Mixed Methods). For this reason, purposeful sampling can be used to reduce mistrust and increase participation among historically marginalized groups and experiences (e.g., those with diverse immigration experiences). As researchers, we need to be able to build trust among participants that research findings will not be used to reinforce negative stereotypes surrounding ethnoracial group (Thompson et al., 2003) or expose them to unnecessary risk (Corbie-Smith, Thomas, & George, 2002; Brandon, Isaac, & LaVeist, 2005). One benefit of purposeful sampling is that sample sizes can remain modest as dictated by traditional power analyses. Should a researcher attempt to recruit many different ethnoracial or intersecting categories for a single study, practical limitations with recruitment and participant remuneration may limit the power needed to conduct appropriate comparisons.

As with purposeful sampling, strategies need to be adjusted based on the type of research study. In qualitative research, moderators of focus groups should be trained to make participants feel comfortable sharing both negative and positive perceptions, recognizing that individuals may respond negatively to disparaging information but positively to messages of hope (Nicholson et al., 2008). In addition, in research conducted using study centers (common with cancer and Alzheimer's disease research), recruitment of diverse populations may be improved by creating community advisory boards, delivering culturally targeted education programs (Scharff et al., 2010), and partnering with community-based organizations who are already serving groups in the community and have built a trusted relationship with those groups. Community-based participatory research (CBPR) models can improve

community relationships by equitably engaging stakeholders in the research process and valuing the strengths that each brings (Minkler & Wallerstein, 2011).

Additionally, the field of psychology may benefit from having more research investigators from diverse backgrounds. The Institute of Medicine and the Department of Health and Human Services explicitly recommend increasing the number of underrepresented minorities in research positions (Smedley, Stith & Nelson, 2003). Such diversity can improve not only recruitment, but also the research design in novel ways, e.g., by providing creative assessments and questionnaires that emphasize incorporating measures to best identify subgroups.

Addressing Bias in Assessment Tests

The easy solution to address bias in assessment tests is to use ethnoracial-specific norms, which have been shown to diagnose neurocognitive disorders with appropriate specificity and sensitivity (Werry, Daniel, & Bergström, 2019). Some researchers have suggested that norms adjusting for educational quality (rather than simply years of education) may help adjust scores without having to create ethnoracial-specific norms (Manly et al., 2002). Other suggestions have been to statistically adjust or match groups of participants by health conditions or common SES variables such as wealth or income. In regard to different levels of health, some investigators have recommended screeners to assess cognitive impairment secondary to cardiovascular disease. For example, the National Institute for Neurological Disorders and Stroke (NINDS) and the Canadian Stroke Network (CSN) recommend a short cognitive screening assessment to identify cardiovascular patients who are experiencing cognitive impairment but may or may not meet the threshold to be classified as demented. Research using such screeners demonstrated that measures of executive functioning, immediate, and delayed recall had adequate sensitivity and specificity to provide useful information to healthcare providers (Cameron et al., 2016). Overall, when researchers control for group differences in health, differences in performance between ethnoracial minorities and non-Hispanic Whites can be attenuated (Aiken-Morgan et al., 2018;

Jones, 2003). Another way to minimize group differences is to estimate mean performance by freely estimating the intercepts for different ethnoracial groups (Avila et al., 2020).

However, these easy solutions present several challenges that also need to be addressed. First, the criteria for adjustment (e.g., education, SES, health) may not be measurement invariant between groups (Jones, 2003). Even accounting for common SES variables assumes standards of social class that represent non-Hispanic White acculturation. Second, adjusting data with different norms or specific variables fails to address the underlying reasons for differences in cognition (Freedman & Manly, 2018). What substantial steps, then, can we take to start addressing the societal inequalities that cause differences in test assessments?

Perhaps the first step is ensure that aging researchers and clinicians be culturally and linguistically competent (Allen & Walsh, 2000). For example, researchers examining the attitudes and/or the mental health of older Chinese Americans should be aware of filial piety and how that might influence responses to questions about autonomy (Allen & Walsh, 2000; Cheung, 1989; Gao et al., 2019). Additionally, they can consider the degree of acculturation, whether among ethnoracial minorities in the U.S. or among immigrant populations (e.g., Manly et al., 2004). When assessing older adult immigrants, researchers and clinicians must be aware of language barriers. Older immigrants are less likely than younger adults to become proficient in a second language (Cheung, 1989), which can be a problem for those without familial support. While efforts have been made to translate some cognitive assessments into various languages, translators often do not fully consider linguistic, cultural, or educational differences to make sure the translation conveys the same message to individuals in multiple cultures (Ng et al., 2018).

Proponents of culturalist perspectives suggest that rather than controlling for cultural influences through culture-fair tests (e.g., Cattell, 1940), hypothesized cultural factors such as behaviors, beliefs, and values should be measured and explicitly tested to identify the sources of test differences (Helms, 1992). Better yet, those cultural factors could be implemented in the construction, administration, and

interpretation of newly designed cognitive tests. Ideally, such tests would not require language “code switching” in which minority groups would have to take on the language and interpretation set by highly educated and White individuals. As summarized by Helms (1992), cognitive tests can be created with intellectual dimensions in mind that might be culturally appropriate, but still have the potential for universal appeal. Some of these dimensions include the ability to integrate feelings with thoughts and actions, valuing the group more than the individual, acquiring and testing through oral expressions, and consideration of immaterial forces in one’s thinking (Boykin & Toms, 1985; Helms, 1992). Although more challenging to implement, construction of such tests would ensure that at broader aspects of intellectual ability are considered rather than only a non-Hispanic White mindset.

Qualitative Inquiry and Mixed Methods

Solutions to many of the aforementioned challenges can be made more manageable by supplementing traditional analyses by idiographic methods including qualitative inquiry and mixed methods. Qualitative research is used to explore and understand the meaning individuals or groups ascribe to experiences (Creswell, 2012, 2014) and is frequently used in psychology for development of theory or generation of new research questions in a relatively unexplored area of research (Howitt, 2019). Many studies using an intersectional lens have relied upon qualitative inquiry.

Sofaer (1999) describes the following applications for qualitative methods in health services research: measuring unique or unexpected events, describing the experience and interpretations of individuals with widely differing stakes and roles, providing a voice for underrepresented self-identified minorities, generating research questions and developing theories, and exploring mechanisms. Methodological standards for qualitative research include credibility, dependability, confirmability, transferability, and authenticity (Cope, 2014). A qualitative study is considered credible if the experience described is immediately recognizable to individuals that share the same experience (Sandelowski, 1986). Dependability is the constancy of data over similar experiences and is supported by investigator “triangulation” or agreement. Confirmability is established when investigators clearly

describe how conclusions and interpretations are established, showing how findings were derived directly from the data. Confirmability is frequently supported by extensive use of quotations from participants. Transferability indicates that findings can be applied to other settings or groups and have meaning and impact with individuals such as readers not involved in the study. Authenticity refers to whether the researcher expresses the feelings and emotions of participants in a faithful manner and also is supported by the frequent use of quotes.

The most commonly used qualitative approach in psychology is constructivist grounded theory (Charmaz & Henwood, 2017; Mills, Bonner, & Francis, 2006). The goal in this approach is to develop a theory based on accrued knowledge and qualitative inquiry investigating the lived experience of individuals in regard to specific phenomena (e.g., subjective age, attitudes, memory complaints, resilience). Such methods can be used to identify meaningful subgroups (thereby breaking apart monolithic groups), better understand the impact of intersecting identities, and can be used to help minimize biases in surveys or cognitive tests to match the language styles or competencies of the audience they intend to measure. For example, Windsong (2018) describes opportunities and challenges in using qualitative interview methods to describe and understand intersectionality of race and gender as related to neighborhood experiences and influences. Using similar methods to incorporate an intersectional lens including age/cohort to critical phenomena of aging may improve replicability. Rather than being unscientific or “soft”, however, incorporation of such designs requires careful forethought and planning to obtain and analyze qualitative data with methodological rigor. Collecting such qualitative data alongside quantitative data also can be quite manageable with recommended sample sizes between five to 25 individuals or until no new theoretical insights are introduced (Creswell, 2014).

Addressing monolithic groups and intersectionality: Novel applications of latent class analysis

As described earlier, many studies treat ethnoracial groups as a monolithic and homogenous. However, scientists are beginning to realize that ethnoracial group and other intersecting factors (e.g.,

SES) are not simply “additive” effects. But for many individuals, such intersections are tied to one’s identity and should not be statistically teased apart as done in regression models. However, we also acknowledge that it would be a gross generalization to conclude that some factors are so highly correlated that they cannot be separated at all (e.g., minority status and low SES). Researchers have argued that such variable-centered analyses obscure diversity, offer over-generalized conclusions, and speak more to the variables themselves than to the people that the analyses are intended to investigate (Bogat et al., 2005; Scotto Rosato & Baer, 2012).

What, then, are we to do when several intersecting factors are highly correlated with one another and yet we do not want to group people together into large homogenous groups? Whereas most aging research has relied on variable-centered methods (e.g., analysis of variance or regression), a valuable approach to investigating ethnoracial and intersecting identities is a person-centered approach. Person-centered methods can reveal subsamples based on varying distributions of inputted characteristics. One such method that is growing in popularity within the field of psychology is latent class analysis (LCA). For details on the mathematics underlying this method, we refer readers to Eid et al. (2003) and Berlin et al. (2014). LCA is a data-driven technique that identifies clusters of participants that share a common set of dichotomous or Likert-scale data. For example, a study from the UK created clusters based on SES, occupation, ethnoracial group, and immigration status (Goodwin et al., 2017). They found separate clusters for White British that either encompassed professionals who were homeowners (class 1) or “economically inactive” (i.e., unemployed, temporarily sick, retired, etc.) people that were renters (class 2). They also found separate clusters that consisted of Black Africans who were either students with high income backgrounds (class 3) or professionals with high income that migrated to the UK (class 6). Yet another class consisted of Black Caribbeans who were nonmigrant skilled workers (class 4). When comparing the clusters on mental health, the “economically inactive” Whites and Black African students were both at higher odds of reporting a common mental disorder (e.g., depression, anxiety, cognitive problems, somatic symptoms, etc.) than the White professionals. Note that other Black groups (Black

African professionals and Black Caribbean nonmigrant skilled workers) did not differ in mental health from the White professionals, highlighting the importance of identifying group heterogeneity. As shown above, the characteristics entered into an LCA could include the aforementioned “intersecting” factors (e.g., ethnoracial group, SES, rurality, sexual orientation).

The family of LCA includes Latent Profile Analyses, which can be applied when the characteristics are continuous rather than dichotomous. For example, Bogner et al. (2009) found that older adults’ cognition and mental health could be separated into three clusters that were not evenly distributed in the community: 1) nondepressed with average to high cognition (62% of the sample), 2) depressed with average cognition (26%), and 3) depressed with lower cognition (12%). They further tested whether the groups differed in risk for Alzheimer’s disease by comparing the frequency of having the APOE ϵ 4 allele in each group but found no group differences. This method could be applied to investigate clusters of individuals that vary on proposed mediating factors of social determinants of health. Note that although many clusters might be identified, each cluster can incorporate complex interactions without reducing the degrees of freedom that a model with 3-way or 4-way interactions might require, thereby preserving power.

Despite these strengths, some limitations exist when this class of methods. First, clusters can be hard to interpret. As mentioned earlier, the clusters can implicitly comprise complex interactions among the variables. For example, data from the Einstein Aging Study (Zammit et al., 2018) was used recently to combine cognitive measures and demographic variables (age, sex, education) into a latent class model yielding 9 different clusters! As with other data-driven techniques like exploratory factor analysis, simplified labels are often used to encompass the cluster, but with some difficulties. In the Einstein Aging Study, the labels became less and less descriptive and did not map neatly onto ethnoracial categories. For example, Black people were the majority group in two different clusters: a) poor executive and poor working memory and b) poor processing speed and executive function. These two clusters differed in age and proportion who also had a stroke, highlighting the heterogeneity among

different Black groups. Although useful, this potential complexity of interpretation requires that researchers think carefully beforehand about which factors make most sense to include and even predict potential outcomes. Such decisions might be based on theoretical principles, themes (e.g., demographic characteristics), or the inclusion of factors that make reasonable sense if formal interactions would normally be tested (ethnoracial group and class). Second, relatively large sample sizes ranging from at least 300-500 participants are recommended so that reliable clusters can be identified and no group has too few participants (Dziak et al., 2014). Third, because this approach is data driven, the clusters may not be replicable in separate samples. Thus, replicating cluster types in separate and independent samples is important to mitigate this concern.

Enhancing generalizability: Lessons from machine learning

Different approaches to solving statistical problems can pave the way for innovations across fields. The field of aging might benefit from common methods applied in machine learning, which has become quite popular in the neurosciences over the last decade (Kriegeskorte, & Douglas, 2018). Machine learning is a category of different statistical tools that often are able to integrate large amounts of data in nonlinear manners to solve complex problems. Although implementation of machine learning algorithms in ethnically diverse older adults might provide novel insights into the social determinants of health that lead to disparities in older adults, we introduce the method here to borrow standards from the technique, not to suggest use of the technique per se.

Like in the neurosciences, most experimental studies of aging consist of modest sample sizes, which threatens generalizability. Studies investigating differences in ethnoracial groups are even more problematic because of the aforementioned nuances of intersecting identities. Even large observational studies that have majority non-Hispanic White samples will be biased by uneven sample sizes of the minority ethnoracial groups, which could carry the main effects. Some of the popular machine learning techniques like support vector machine require one sample of individuals to form a statistical model (as one would implement in a regression or analysis of variance) and a second, independent sample to test

the generalizability of the model. As one might expect, a model that fits quite well in the original sample or explains a large amount of variance in the dependent variable often does not fit as well in an independent sample (e.g., Rieck et al., 2015). While this general notion is similar to testing group invariance in SEM, the manner in which machine learning implements this idea is quite different. Such a formal test is important because an independent sample that fails to perform better than chance is equivalent to a replication failure across studies. Such a finding might indicate a bias in the original sample or measures, or could indicate a large degree of heterogeneity across people such that one sample is simply not generalizable to another (Yip, 2018). If such poor generalization occurs within a single study, the authors can search for and test the origins of such biases or heterogeneity in the samples.

Testing the generalizability of a finding is common practice in the field of machine learning even if the sample is not large enough to evenly split the sample in two, which is a large barrier in other methods such as in SEM. Machine learning uses a leave-one-out-validation approach in which a statistical model is tested on a large proportion of the data (typically 75%-80%) but can also be tested on all but one of the participants in the sample. Then, the same parameters from the original statistical model are applied to predict the outcome measure in the left-out participant or group as in jackknifing (Thompson, 1989). The left-out data then serves as a completely independent test of generalizability. The predicted outcome can be compared to the actual data to derive model fit indices as one usually might implement in a standard regression model. Next, a new statistical model is created by rotating each participant/group through and leaving out a different participant/group. Thus, this leave-one-out-validation approach results in many assessments of the model parameters (e.g., beta-coefficients) that can serve as a measure of variability of generalizability and tested against a hypothesized effect size or simply to zero. In a study investigating how the neural correlates of face processing is impacted by Alzheimer's disease pathology (Rieck et al., 2015), the statistical model using the whole sample had a

prediction accuracy of 64.5%, but when the independent validation procedure was used, the accuracy fell to 48.4% (chance being 34.2%).

In the context of health disparities, this leave-one-out-validation technique also can test model generalizability to different ethnoracial groups. Most studies in cognitive aging (and likely many other fields on aging) have samples that are biased toward non-Hispanic Whites. One can easily test the extent that the model parameters estimated in non-Hispanic Whites generalize to other ethnoracial groups by applying the same model parameters to the independent group. Of course, such models assume that the measurements of episodic memory, education, and other covariates are treated at face value (i.e., are measurement invariant).

Conclusion

We focused on three main issues that increase the risk of nonreplicability of aging research from the lens of ethnoracial health disparities: treating ethnoracial groups as monolithic and homogenous, not considering intersecting identities, and biases of research materials. Importantly, some of our examples to support these ideas were not in older adult samples largely because these issues have not been sufficiently investigated in the context of aging. While we provided many examples in the field of cognitive aging research, all of these issues are as important (or even more important) in other fields in which one's identity or intersection of identities may play a more central role. Nevertheless, we have provided a glimpse of how many social determinants of health should be considered fundamental to aging research both in the construction of measures, the analyses conducted, and the interpretation of those measures. When we fail to recognize these issues, we are then approaching our research questions in a manner that will surely leave much of the scientific story untold or even introduce misleading conclusions. We currently have the opportunity to achieve both generalizable research outcomes and social justice by being more inclusive and culturally sensitive in our research choices. We are optimistic that such choices will have positive downstream consequences in aging research such as increasing trust in scientific research and the findings resulting from them. Such trust can, in turn, create an atmosphere

of openness to participating in future research efforts. As researchers, we have all contributed to studies that have ignored these issues, but by recognizing them now and offering guidance on how to correct for them in future research, our hope is that we will all raise our standards of conducting research. By considering more flexible analytic tools and embracing research participants' individuality, we can increase the value of our research outcomes.

References

- Aiken-Morgan, A. T., Gamaldo, A. A., Wright, R. S., Altaire, J. C., & Whitfield, K. E. (2018). Stability and change in cognitive status classification of black older adults. *Journal of the American Geriatrics Society*, *66*(1), 179-183.
- Albertini, M., Mantovani, D., & Gasperoni, G. (2019). Intergenerational relations among immigrants in Europe: the role of ethnic differences, migration and acculturation. *Journal of Ethnic and Migration Studies*, *45*, 1693-1706.
- Albuja, A. F., Sanchez, D. T., & Gaither, S. E. (2018). Fluid racial presentation: Perceptions of contextual “passing” among biracial people. *Journal of Experimental Social Psychology*, *77*, 132-142.
- Allen, J., & Walsh, J. A. (2000). A construct-based approach to equivalence: Methodologies for cross-cultural/multicultural personality assessment research. *Handbook of cross-cultural and multicultural personality assessment*, 63-85.
- Assari, S., Dejman, M. & Neighbors, H.W. (2016). Ethnic Differences in Separate and Additive Effects of Anxiety and Depression on Self-rated Mental Health Among Blacks. *Journal of Racial and Ethnic Health Disparities* *3*, 423–430 (2016). <https://doi.org/10.1007/s40615-015-0154-3>
- Avila, J. F., Rentería, M. A., Witkiewitz, K., Verney, S. P., Vonk, J. M., & Manly, J. J. (2019). Measurement invariance of neuropsychological measures of cognitive aging across race/ethnicity by sex/gender groups. *Neuropsychology*, *34*(1), 3-14.
- Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundation of developmental theory. *American psychologist*, *52*(4), 366.
- Barnes, L. L., Lewis, T. T., Begeny, C. T., Yu, L., Bennett, D. A., & Wilson, R. S. (2012). Perceived discrimination and cognition in older African Americans. *Journal of the International Neuropsychological Society: JINS*, *18*(5), 856-865.

Barnes, L. L., Yumoto, F., Capuano, A., Wilson, R. S., Bennett, D. A., & Tractenberg, R. E. (2016).

Examination of the factor structure of a global cognitive function battery across race and time.

Journal of the International Neuropsychological Society, 22(1), 66-75.

Bauer, G. R. (2014). Incorporating intersectionality theory into population health research methodology:

Challenges and the potential to advance health equity. *Social Science & Medicine*, 110, 10–17.

<http://dx.doi.org/10.1016/j.socscimed.2014.03.022>

Baykara-Krumme, H., & Platt, L. (2018). Life satisfaction of migrants, stayers and returnees: Reaping

the fruits of migration in old age?. *Ageing & Society*, 38(4), 721-745.

Beatie, B. E., Mackenzie, C. S., & Chou, K. L. (2015). Prevalence of psychiatric and medical

comorbidities in HIV-positive middle-aged and older adults: findings from a nationally representative survey. *Journal of Therapy and Management of HIV Infection*, 3(1), 7-16.

Beatty Moody, D. L., Taylor, A. D., Leibel, D. K., Al-Najjar, E., Katzel, L. I., Davatzikos, C., . . . Erus,

G. (2019). Lifetime discrimination burden, racial discrimination, and subclinical cerebrovascular disease among African Americans. *Health Psychology*, 38(1), 63-74.

Berlin, K. S., Williams, N. A., & Parra, G. R. (2014). An introduction to latent variable mixture

modeling (part 1): Overview and cross-sectional latent class and latent profile analyses. *Journal of pediatric psychology*, 39(2), 174-187.

Bernstein, K. S., Park, S., Shin, J., Cho, S., & Park, Y. (2011). Acculturation, discrimination and

depressive symptoms among Korean immigrants in New York City. *Community Mental Health Journal*, 47(1), 24-34.

Berry, J. W. (2003). *Conceptual approaches to acculturation*. In K. M. Chun, P. Balls Organista, & G.

Marín (Eds.), *Acculturation: Advances in theory, measurement, and applied research* (p. 17–37).

American Psychological Association.

- Berlin, K. S., Williams, N. A., & Parra, G. R. (2014). An introduction to latent variable mixture modeling (part 1): Overview and cross-sectional latent class and latent profile analyses. *Journal of pediatric psychology, 39*(2), 174-187.
- Bertola, L., Benseñor, I. M., Barreto, S. M., Moreno, A. B., Griep, R. H., Viana, M. C., ... & Suemoto, C. K. (2019). Measurement invariance of neuropsychological tests across different sociodemographic backgrounds in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). *Neuropsychology, 34*(2), 227-234.
- Betancourt, J. R., Green, A. R., Carrillo, J. E., & Owusu Ananeh-Firempong, I. I. (2003). Defining cultural competence: A practical framework for addressing racial/ethnic disparities in health and health care. *Public health reports, 118*, 293-302.
- Binette, A. P., Vachon-Preseau, É., Morris, J., Bateman, R., Benzinger, T., Collins, D. L., ... & PREVENT-AD Research Group. (2020). Amyloid and tau pathology associations with personality traits, neuropsychiatric symptoms and cognitive lifestyle in the preclinical phases of sporadic and autosomal dominant Alzheimer's disease. *Biological psychiatry*.
- Bogat G. A., Levendosky A .A., von Eye A.. The future of research on intimate partner violence: Person-oriented and variable-oriented perspectives, *American Journal of Community Psychology*, 2005, vol. 36 1/2(pg. 49-70)
- Bogner, H. R., Richie, M. B., de Vries, H. F., & Morales, K. H. (2009). Depression, cognition, apolipoprotein e genotype: Latent class approach to identifying subtype. *The American Journal of Geriatric Psychiatry, 17*(4), 344-352.
- Bowleg, L. (2012). The problem with the phrase *women and minorities*: Intersectionality—an important theoretical framework for public health. *American Journal of Public Health, 102*, 1267–1273.
<http://dx.doi.org/10.2105/AJPH.2012.300750>

- Boykin, A. W., & Toms, F. D. (1985). *Black child socialization: A conceptual framework*. In H. P. McAdoo & J. L. McAdoo (Eds.), *Sage focus editions, Vol. 72. Black children: Social, educational, and parental environments* (p. 33–51). Sage Publications, Inc.
- Brandon, D. T., Isaac, L. A., & LaVeist, T. A. (2005). The legacy of Tuskegee and trust in medical care: is Tuskegee responsible for race differences in mistrust of medical care?. *Journal of the National Medical Association, 97*(7), 951-956.
- Brown, M. C., & Davis, J. E. (2001). The historically Black college as social contract, social capital, and social equalizer. *Peabody journal of Education, 76*(1), 31-49.
- Calvo, R., Carr, D. C., & Matz-Costa, C. (2019). Expanding the Happiness Paradox: Ethnoracial Disparities in Life Satisfaction Among Older Immigrants in the United States. *Journal of Aging and Health, 31*(2), 231-255.
- Cameron, J. D., Gallagher, R., Pressler, S. J., McLennan, S. N., Ski, C. F., Tofler, G., & Thompson, D. R. (2016). Sensitivity and specificity of a five-minute cognitive screening test in patients with heart failure. *Journal of Cardiac Failure, 22*(2), 99-107.
- Cattell, R. B. (1940). A culture-free intelligence test. I. *Journal of Educational Psychology, 31*(3), 161-179.
- Carstensen, L. L. and Löckenhoff, C. E. (2003). Ageing, emotion, and evolution. *Annals of the New York Academy of Sciences, 1000*, 1, 152–179.
- Cary, L. A., & Chasteen, A. L. (2015). Age stereotypes and age stigma: Connections to research on subjective aging. *Annual review of gerontology and geriatrics, 35*(1), 99-119.
- Charmaz, K., & Henwood, K. (2017). Grounded theory methods for qualitative psychology. *The SAGE handbook of qualitative research in psychology, 2*, 238e256.
- Cheung, M. (1989). Elderly Chinese living in the united states: Assimilation or adjustment? *Social Work, 34*(5), 457-461.

- Choi, S. L., McDonough, I. M., Kim, M., & Kim, G. (2019). Estate planning among older Americans: The moderating role of race and ethnicity. *Financial Planning Review*, 2, e1058.
- Choi, S. L., McDonough, I. M., Kim, M., & Kim, G. (2020). The association between the number of chronic health conditions and advance care planning varies by race/ethnicity. *Aging & Mental Health*, 24, 453-463.
- Cody, S. L., & Vance, D. E. (2016). The neurobiology of HIV and its impact on cognitive reserve: A review of cognitive interventions for an aging population. *Neurobiology of Disease*, 92, 144-156.
- Conron, K. J., Mimiaga, M. J., & Landers, S. J. (2010). A population-based study of sexual orientation identity and gender differences in adult health. *American Journal of Public Health*, 100(10), 1953-1960.
- Contador, I., Del Ser, T., Llamas, S., Villarejo, A., Benito-León, J., & Bermejo-Pareja, F. (2017). Impact of literacy and years of education on the diagnosis of dementia: A population-based study. *Journal of Clinical and Experimental Neuropsychology*, 39(2), 112-119.
- Cope, D. G. (2014). Methods and meanings: credibility and trustworthiness of qualitative research. In *Oncology nursing forum*, 41(1), 89-91.
- Corbie-Smith, G., Thomas, S. B., & George, D. M. M. S. (2002). Distrust, race, and research. *Archives of Internal Medicine*, 162(21), 2458-2463.
- Creswell, J. W. (2012). *Qualitative inquiry & research design: Choosing among five approaches* (4th ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W. (2014). *A concise introduction to mixed methods research*. SAGE publications.
- Cresswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed method research*. 2nd Sage. Thousand Oaks, CA, 201.
- Crowe, M., Clay, O. J., Martin, R. C., Howard, V. J., Wadley, V. G., Sawyer, P., & Allman, R. M. (2013). Indicators of childhood quality of education in relation to cognitive function in older

- adulthood. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 68(2), 198-204.
- Davenport, L. D. (2016). The role of gender, class, and religion in biracial Americans' racial labeling decisions. *American Sociological Review*, 81, 57–84. <http://dx.doi.org/10.1177/0003122415623286>.
- Deckers, K., van Boxtel, M. P., Schiepers, O. J., de Vugt, M., Munoz Sanchez, J. L., Anstey, K. J., ... & Ritchie, K. (2015). Target risk factors for dementia prevention: a systematic review and Delphi consensus study on the evidence from observational studies. *International journal of geriatric psychiatry*, 30(3), 234-246.
- Dilley, J. A., Simmons, K. W., Boysun, M. J., Pizacani, B. A., & Stark, M. J. (2010). Demonstrating the importance and feasibility of including sexual orientation in public health surveys: health disparities in the Pacific Northwest. *American Journal of Public Health*, 100(3), 460-467.
- Dotson, V. M., & Duarte, A. (2020). The importance of diversity in cognitive neuroscience. *Annals of the New York Academy of Sciences*, 1464(1), 181-191.
- Dziak, J. J., Lanza, S. T., & Tan, X. (2014). Effect size, statistical power, and sample size requirements for the bootstrap likelihood ratio test in latent class analysis. *Structural Equation Modeling*, 21(4), 534–552. <https://doi.org/10.1080/10705511.2014.919819>
- Earnshaw, V. A., Rosenthal, L., Gilstad-Hayden, K., Carroll-Scott, A., Kershaw, T. S., Santilli, A., & Ickovics, J. R. (2018). Intersectional experiences of discrimination in a low-resource urban community: An exploratory latent class analysis. *Journal of Community & Applied Social Psychology*, 28, 80–93. <http://dx.doi.org/10.1002/casp.2342>
- Eid, M., Langeheine, R., & Diener, E. (2003). Comparing typological structures across cultures by multigroup latent class analysis: A primer. *Journal of cross-cultural psychology*, 34(2), 195-210.
- Elder, G. H., Jr. (1998). The life course as developmental theory. *Child Development*, 69(1), 1–12. <http://dx.doi.org/10.1111/j.1467-8624.1998.tb06128.x>

- Emler, C. A. (2016). Social, economic, and health disparities among LGBT older adults. *Generations, 40*(2), 16-22.
- Fredriksen-Goldsen, K. I., Cook-Daniels, L., Kim, H. J., Erosheva, E. A., Emler, C. A., Hoy-Ellis, C. P., ... & Muraco, A. (2014). Physical and mental health of transgender older adults: An at-risk and underserved population. *The Gerontologist, 54*(3), 488-500.
- Fredriksen-Goldsen, K. I., Kim, H. J., Barkan, S. E., Muraco, A., & Hoy-Ellis, C. P. (2013). Health disparities among lesbian, gay, and bisexual older adults: Results from a population-based study. *American journal of public health, 103*(10), 1802-1809.
- Freedman, D., & Manly, J. J. (2018). Assessment of cognition in African American older adults. *APA handbook of dementia*. (pp. 107-123) American Psychological Association.
- Freimuth, V.S., Quinn, S.C., Thomas, S.B., Cole, G., Zook, E., & Duncan, T. (2001). African Americans' views on research and the Tuskegee Syphilis Study. *Social Science and Medicine, 52*(5), 797-808.
- Fyffe, D. C., Mukherjee, S., Barnes, L. L., Manly, J. J., Bennett, D. A., & Crane, P. K. (2011). Explaining differences in episodic memory performance among older African Americans and whites: The roles of factors related to cognitive reserve and test bias. *Journal of the International Neuropsychological Society: JINS, 17*(4), 625.
- Gao, X., Sun, F., & Hodge, D. R. (2019). Elder mistreatment among Chinese American families: Do acculturation and traditionalism matter? *The Journals of Gerontology: Series B, 74*(3), 465-473.
- Garcia, M. A., Reyes, A. M., Downer, B., Saenz, J. L., Samper-Ternent, R. A., & Raji, M. (2017). Age of migration and the incidence of cognitive impairment: A cohort study of elder Mexican Americans. *Innovation in Aging, 1*(3), igx037.
- Glymour, M. M., & Manly, J. J. (2008). Lifecourse social conditions and racial and ethnic patterns of cognitive aging. *Neuropsychology review, 18*(3), 223-254.

- Goodwin, L., Gazard, B., Aschan, L., MacCrimmon, S., Hotopf, M., & Hatch, S. L. (2017). Taking an intersectional approach to define latent classes of socioeconomic status, ethnicity and migration status for psychiatric epidemiological research. *Epidemiology and Psychiatric Sciences*, *9*, 1-12.
- Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist*, *47*(9), 1083.
- Hill, C. V., Pérez-Stable, E. J., Anderson, N. A., & Bernard, M. A. (2015). The National Institute on Aging health disparities research framework. *Ethnicity & disease*, *25*(3), 245-254.
- Howitt, D. (2019). *Introduction to qualitative research methods in psychology: Putting theory into practice*. Pearson UK.
- Islam, N. S., Khan, S., Kwon, S., Jang, D., Ro, M., & Trinh-Shevrin, C. (2010). Methodological issues in the collection, analysis, and reporting of granular data in Asian American populations: historical challenges and potential solutions. *Journal of Health Care for the Poor and Underserved*, *21*(4), 1354–1381. <https://doi.org/10.1353/hpu.2010.0939>
- Jackson, J. S., Govia, I. O., & Sellers, S. L. (2010). Race and ethnic influences over the life-course. In R. H. Binstock & L. K. George (Eds.), *Handbook of aging and the social sciences* (7th ed., pp. 91-103). New York, NY: Academic Press.
- James, C. V., Moonesinghe, R., Wilson-Frederick, S. M., Hall, J. E., Penman-Aguilar, A., & Bouye, K. (2017). Racial/ethnic health disparities among rural adults—United States, 2012–2015. *MMWR Surveillance Summaries*, *66*(23), 1-9.
- Jang, Y., Park, N. S., Kang, S. Y., & Chiriboga, D. A. (2014). Racial/ethnic differences in the association between symptoms of depression and self-rated mental health among older adults. *Community Mental Health Journal*, *50*(3), 325-330.
- Jones, R. N. (2003). Racial bias in the assessment of cognitive functioning of older adults. *Aging & Mental Health*, *7*(2), 83-102.

- Kim, G., Chiriboga, D. A., Bryant, A., Huang, C. H., Crowther, M., & Ma, G. X. (2012). Self-rated mental health among Asian American adults: Association with psychiatric disorders. *Asian American Journal of Psychology, 3*(1), 44.
- Kriegeskorte, N., & Douglas, P. K. (2018). Cognitive computational neuroscience. *Nature neuroscience, 21*(9), 1148-1160.
- Lewis, T. T., & Van Dyke, M. E. (2018). Discrimination and the health of African Americans: The potential importance of intersectionalities. *Current Directions in Psychological Science, 27*(3), 176-182.
- Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., ... & Cooper, C. (2017). Dementia prevention, intervention, and care. *The Lancet, 390*(10113), 2673-2734.
- Logan, R. W. (1933). Educational Segregation in the North. *Journal of Negro Education, 65*-67.
- Manly, J. J., Byrd, D. A., Touradji, P., & Stern, Y. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology, 11*(1), 37-46.
- Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society: JINS, 8*(3), 341-348.
- Martinez-Miller, E. E., Robinson, W. R., Avery, C. L., Yang, Y. C., Haan, M. N., Prather, A. A., & Aiello, A. E. (2020). Longitudinal associations of US acculturation with cognitive performance, cognitive impairment, and dementia: The Sacramento area Latino study on aging. *American Journal of Epidemiology, 189*(11), 1292-1305.
- Mayeda, E. R., Glymour, M. M., Quesenberry Jr, C. P., & Whitmer, R. A. (2017). Heterogeneity in 14-year dementia incidence between Asian American subgroups. *Alzheimer Disease and Associated Disorders, 31*(3), 181.

- Mills, J., Bonner, A., & Francis, K. (2006). The development of constructivist grounded theory. *International journal of qualitative methods*, 5(1), 25-35.
- Mindt, M. R., Arentoft, A., Coulehan, K., & Byrd, D. (2013). Considerations for the neuropsychological evaluation of older ethnic minority populations. *Handbook on the neuropsychology of aging and dementia* (pp. 25-41) Springer.
- Minkler, M., & Wallerstein, N. (Eds.). (2011). Community-based participatory research for health: From process to outcomes. John Wiley & Sons
- Moheet, A., Mangia, S., & Seaquist, E. R. (2015). Impact of diabetes on cognitive function and brain structure. *Annals of the New York Academy of Sciences*, 1353, 60-71.
- Ng, K. P., Chiew, H. J., Lim, L., Rosa-Neto, P., Kandiah, N., & Gauthier, S. (2018). The influence of language and culture on cognitive assessment tools in the diagnosis of early cognitive impairment and dementia. *Expert Review of Neurotherapeutics*, 18(11), 859-869.
- Nicholson, R. A., Kreuter, M. W., Lapka, C., Wellborn, R., Clark, E. M., Sanders-Thompson, V., ... & Casey, C. (2008). Unintended effects of emphasizing disparities in cancer communication to African-Americans. *Cancer Epidemiology and Prevention Biomarkers*, 17(11), 2946-2953.
- Pantoni, L., Poggesi, A., & Inzitari, D. (2007). The relation between white-matter lesions and cognition. *Current opinion in neurology*, 20(4), 390-397.
- Plassman, B. L., Langa, K. M., Fisher, G. G., Heeringa, S. G., Weir, D. R., Ofstedal, M. B., ... & Steffens, D. C. (2007). Prevalence of dementia in the United States: the aging, demographics, and memory study. *Neuroepidemiology*, 29(1-2), 125-132.
- Rieck, J. R., Rodrigue, K. M., Kennedy, K. M., Devous Sr, M. D., & Park, D. C. (2015). The effect of beta-amyloid on face processing in young and old adults: a multivariate analysis of the BOLD signal. *Human brain mapping*, 36(7), 2514-2526.
- Ruis, C., Biessels, G. J., Gorter, K. J., Van Den Donk, M., Kappelle, L. J., & Rutten, G. E. (2009). Cognition in the early stage of type 2 diabetes. *Diabetes care*, 32(7), 1261-1265.

- Salter, P. S., Adams, G., & Perez, M. J. (2018). Racism in the structure of everyday worlds: A cultural-psychological perspective. *Current Directions in Psychological Science*, 27(3), 150-155.
- Samelson, F. (1982). H.H. Goddard and the immigrants. *American Psychologist*, 37(11), 1291–1292.
- Sandelowski, M. (1986). The problem of rigor in qualitative research. *Advances in Nursing Science*, 8(3), 27–37.
- Scharff, D. P., Mathews, K. J., Jackson, P., Hoffsuemmer, J., Martin, E., & Edwards, D. (2010). More than Tuskegee: understanding mistrust about research participation. *Journal of Health Care for the Poor and Underserved*, 21(3), 879-897.
- Schwartz, S. J., Weisskirch, R. S., Hurley, E. A., Zamboanga, B. L., Park, I. J., Kim, S. Y., ... & Greene, A. D. (2010). Communalism, familism, and filial piety: Are they birds of a collectivist feather?. *Cultural Diversity and Ethnic Minority Psychology*, 16(4), 548.
- Scotto Rosato, N., & Baer, J. C. (2012). Latent class analysis: a method for capturing heterogeneity. *Social Work Research*, 36(1), 61-69.
- Shadlen, M. F., Siscovick, D., Fitzpatrick, A. L., Dulberg, C., Kuller, L. H., & Jackson, S. (2006). Education, cognitive test scores, and black-white differences in dementia risk. *Journal of the American Geriatrics Society*, 54(6), 898-905.
- Shields, G. S., & Slavich, G. M. (2017). Lifetime stress exposure and health: A review of contemporary assessment methods and biological mechanisms. *Social and Personality Psychology Compass*, 11(8), e12335.
- Smedley, B. D., Stith, A. Y., & Nelson, A. R. (2003). Institute of Medicine (US) Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care: Unequal Treatment: Confronting Racial and ethnic disparities in health care. Washington, DC: National Academies Press.
- Sofaer, S. (1999). Qualitative methods: what are they and why use them?. *Health Services research*, 34(5 Pt 2), 1101.

- Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, 47(10), 2015-2028.
- Swope, C. B., & Hernández, D. (2019). Housing as a determinant of health equity: A conceptual model. *Social Science & Medicine*, 243, 112571.
- Tang, F., Jang, H., Rauktis, M. B., Musa, D., & Beach, S. (2019). The race paradox in subjective wellbeing among older Americans. *Ageing & Society*, 39(3), 568-589.
- Thompson, B. (1989). Statistical significance, result importance, and result generalizability: Three noteworthy but somewhat different issues. *Measurement and Evaluation in Counseling and Development*, 22(1), 2-6.
- Thompson, H. S., Valdimarsdottir, H. B., Jandorf, L., & Redd, W. (2003). Perceived disadvantages and concerns about abuses of genetic testing for cancer risk: differences across African American, Latina and Caucasian women. *Patient Education and Counseling*, 51(3), 217-227.
- U.S. Census Bureau (2019). The Hispanic Population in the United States: 2019. Retrieved from <https://www.census.gov/data/tables/2019/demo/hispanic-origin/2019-cps.html>
- Wallace, S. P., Cochran, S. D., Durazo, E. M., & Ford, C. L. (2011). The health of aging lesbian, gay and bisexual adults in California. *Policy Brief (UCLA Center for Health Policy Research)*, 1-8.
- Werry, A. E., Daniel, M., & Bergström, B. (2019). Group differences in normal neuropsychological test performance for older non-Hispanic White and Black/African American adults. *Neuropsychology*, 33(8), 1089-1100.
- Windsong, E. A. (2018). Incorporating intersectionality into research design: An example using qualitative interviews. *International Journal of Social Research Methodology*, 21(2), 135-147.
- Wirth, M., Haase, C. M., Villeneuve, S., Vogel, J., & Jagust, W. J. (2014). Neuroprotective pathways: lifestyle activity, brain pathology, and cognition in cognitively normal older adults. *Neurobiology of aging*, 35(8), 1873-1882.

Yaffe, K., Falvey, C., Harris, T. B., Newman, A., Satterfield, S., Koster, A., ... & Simonsick, E. (2013).

Effect of socioeconomic disparities on incidence of dementia among biracial older adults: prospective study. *BMJ*, 347.

Yip, T. (2018). Ethnic/racial identity—A double-edged sword? Associations with discrimination and psychological outcomes. *Current Directions in Psychological Science*, 27(3), 170-175.

Zammit, A. R., Hall, C. B., Lipton, R. B., Katz, M. J., & Muniz-Terrera, G. (2018). Identification of heterogeneous cognitive subgroups in community-dwelling older adults: A latent class analysis of the Einstein aging study. *Journal of the International Neuropsychological Society: JINS*, 24(5), 511-523.