



Cognitive differences across ethnoracial category, socioeconomic status across the Alzheimer's disease spectrum: Can an ability discrepancy score level the playing field?

Ian M. McDonough¹ · Shameka L. Cody² · Erin R. Harrell¹ · Stephanie L. Garrett³ · Taylor E. Popp¹

Accepted: 12 March 2022 / Published online: 25 March 2022
© The Psychonomic Society, Inc. 2022

Abstract

An ability discrepancy (crystallized minus fluid abilities) might be a personally relevant cognitive marker of risk for Alzheimer's disease (AD) and might help reduce measurement bias often present in traditional measures of cognition. In a large national sample of adults aged 60–104 years ($N = 14,257$), we investigated whether the intersectionality of group characteristics previously shown to pose a risk for AD including ethnoracial category, socioeconomic status, and sex (a) differed in ability discrepancy compared to traditional neuropsychological tests and (b) moderated the relationship between an ability discrepancy and AD symptom severity. In cognitively normal older adults, results indicated that across each decade, fluid and memory composite scores generally exhibited large group differences with sex, education, and ethnoracial category. In contrast, the ability discrepancy score showed much smaller group differences, thus removing much of the biases inherent in the tests. Women with higher education differed in discrepancy performance from other groups, suggesting a subgroup in which this score might reduce bias to a lesser extent. Importantly, a greater ability discrepancy was associated with greater AD symptom severity across the AD continuum. Subgroup analyses suggest that this relationship holds for all groups except for some subgroups of Hispanic Americans. These findings suggest that an ability discrepancy measure might be a better indicator of baseline cognition than traditional measures that show more egregious measurement bias across diverse groups of people.

Keywords Ability discrepancy · Alzheimer's disease · Education · Health disparities · Socioeconomic status

Introduction

Alzheimer's disease (AD) is the fifth leading cause of death for older adults and affects one out of every nine older adults in the USA (Alz.org, 2021). Older adults who at one time have normal levels of cognition but who later receive an AD diagnosis are said to be in the preclinical stage of AD. In this preclinical stage, brain pathology develops including amyloid plaques and fibrillar tau (Jack et al., 2018). Tools

to aid early detection of this preclinical stage would allow individuals to engage in lifestyle changes such as cognitive stimulation, mindfulness, and exercise (Deckers et al., 2015) and pharmaceutical treatments (Sperling et al., 2014) to lengthen their healthspan – the time during which they are cognitively healthy (Rowe & Kahn, 1987). While assessing in vivo markers of AD pathology can identify older adults at high risk of eventually being diagnosed with AD, current procedures to do so are expensive or invasive, making such detection inaccessible to much of the population such as working-class individuals or those who do not live close to a hospital. This barrier leaves a glaring necessity for more affordable techniques to identify AD risk consistent with those that detect AD pathology and are associated with progression of AD symptoms.

Much research is now focused on sophisticated neuropsychological tests that can identify subtle and early objective decline in cognition (but still in the normative range of cognition relative to the population) that can predict future

✉ Ian M. McDonough
immcdonough@ua.edu

¹ Department of Psychology, The University of Alabama, Box 870348, Tuscaloosa, AL 35487, USA

² College of Nursing, The University of Alabama, Tuscaloosa, AL, USA

³ Emory University Student Health Services, Atlanta, GA, USA

cognitive impairment or conversion to later AD stages (e.g., Mortamais et al., 2017; Papp et al., 2019; Thomas et al., 2018). While early research used neuropsychological tests to measure single cognitive domains like executive function or episodic memory (e.g., Glisky et al., 1995), more recent research suggests that composite scores that measure multiple cognitive domains might be more sensitive. Such multi-domain composites like the Preclinical Alzheimer's Cognitive Composite (PACC) have been rising in popularity to serve this purpose (Donohue et al., 2014; Lim et al., 2016; Mormino et al., 2017; Papp et al., 2017). Along these lines, proposals have been made to suggest that a common factor across many cognitive domains differs between cognitively normal older adults and those with AD, with single domains like episodic memory offering only a small proportion of unique effect in AD (e.g., Salthouse & Becker, 1998).

However, the level of cognitive performance estimated through both single-domain tests (e.g., episodic memory) and multi-domain composites (e.g., fluid abilities) often differ across diverse populations, potentially influencing the interpretation of the scores. For example, some Black older adults have lower cognitive performance and a higher incidence of an AD diagnosis than non-Hispanic White Americans, but at the same time also have less cognitive decline over time (e.g., Weuve et al., 2018). Researchers have largely attributed this contradictory pattern to some Black older adults being closer to the threshold of cognitive impairment such that slight declines in cognition can increase the likelihood of an AD diagnosis (Weuve et al., 2018).

An alternative explanation is that cognitive tests favor educated, non-Hispanic White males who originally created many of the neuropsychological tests that we use today (Helms, 1992; Jones, 2003). Some studies have questioned the face validity of neuropsychological tests when used in diverse groups because of a potential for a misdiagnosis of AD, although we note not all studies have found that traditional neuropsychological tests differ in minority groups from non-Hispanic Whites (Barnes et al., 2016). One recommendation to account for such measurement differences is to use a “personally relevant” cognitive marker that calibrates test scores (Rentz & Weintraub, 2000; Weintraub et al., 2018). The present study aimed to explore how such a cognitive marker, dubbed a crystallized-fluid ability discrepancy, differs across traditionally marginalized and under-represented groups: ethnoracial minorities, women, and those with lower levels of education.

Diverse characteristics influencing an Alzheimer's disease (AD) diagnosis

Black and Hispanic Americans are twice as likely to be diagnosed with AD and related dementias than non-Hispanic White Americans (Matthews et al., 2018; Perkins et al.,

1997; Steenland et al., 2016; Tang et al., 2001). Traditional cognitive screening tests (e.g., Mini-mental State Examination (MMSE) or Montreal Cognitive Assessment (MoCA)) have well-accepted limitations due to cultural and linguistic elements (e.g., error in correctly identifying “rhinoceros” in a naming subscale). In fact, some of the increase in AD diagnoses might be exacerbated by inappropriate cut-offs used during the screening process for marginalized groups (e.g., Goldstein et al., 2014) or biases during structured interviews (Kiselica, in press). Increasing evidence suggests that ethnoracial differences in cognitive test scores can measure disparate socioeconomic characteristics and arise from differences in family income, education, learning materials, and safe physical environment, to name a few (Cottrell et al., 2015). Higher socioeconomic status (SES) – often measured by education level – has been an indicator of resources that allows one to engage in cognitively rich activities to protect a person from cognitive and functional decline, thus delaying a diagnosis of AD (e.g., Bennett et al., 2003; Roe et al., 2007). This notion has been formalized into the concept of cognitive reserve (Stern, 2012). Thus, minoritized groups with lower SES may have lower cognitive reserve.

Lower SES has been associated with increased stress, which both directly and indirectly leads to deterioration of brain pathways involved in attention and memory (Letang et al., 2021; Nogueira et al., 2016; Zimmerman et al., 2016). Lower SES is also associated with poorer knowledge of beneficial health behaviors (Cubbin & Winkleby, 2005), access to healthcare resources (Adler et al., 1994; Braveman et al., 2010; Kennedy et al., 1998), lower levels of green space, and increased prevalence of fast-food chains (Larson et al., 2009; Powell et al., 2007). All of these factors are sustained and exacerbated by structural institutions that maintain a hierarchy of benefits for upper class non-Hispanic White males at the top and everyone else below, thereby decreasing health generally as well as cognition for under-represented groups (e.g., Barnett et al., 2012; Pohl et al., 2021). Related research has shown that everyday discrimination, which is often higher in minoritized groups, is also related to poorer cognition, further increasing cognitive disparities (e.g., Barnes et al., 2012; Ozier et al., 2019; Sutin et al., 2015; Zahodne et al., 2019). For these reasons, among others, socioeconomic conditions outside of one's control have been proposed as a fundamental cause of many diseases (Link & Phelan, 1995). According to this idea, SES can influence disease processes through a variety of pathways, given its link to such a wide array of financial, social, and cognitive resources.

Late-life cognition and risk for AD not only differs between ethnoracial categories and SES, but also with biological sex (for review, see Buckley et al., 2018). Declines in sex hormones in later life have a direct link to AD risk in men and women. AD pathogenesis is believed to be

regulated by estrogen and progesterone in women and primarily by androgens in men (Vest & Pike, 2013). The sharp reduction in sex hormones for women after menopause might then diminish the neuroprotection that these hormones have in reducing AD pathology, potentially explaining why women are 1.5 times more likely to be diagnosed with AD than men (Andersen et al., 1999; Matthews et al., 2018). Women diagnosed with AD sometimes have more neuropathology, faster rates of brain atrophy, and faster rates of cognitive decline than men (Ardekani et al., 2016; Buckley et al., 2018; Cavedo et al., 2018; Duarte-Guterman et al., 2021; Hohman et al., 2018).

Intersectionality of diverse characteristics

We have briefly summarized how diversity among older adults might lead to differential risk of late-life cognitive decline and risk for AD. Little research has focused on cognitive aging from an intersectional perspective (McDonough et al., 2021). In one example, over 5,000 non-Hispanic White, Black, and Hispanic men and women were tested on their memory and visuospatial abilities (Avila et al., 2019). Cognitive decline was greater for Black women compared to Hispanic men and non-Hispanic women, after adjusting for age and education. The Weathering Hypothesis proposes that social-disadvantaged individuals often stemming from systemic discrimination experience accelerated aging due to chronic and recurring stressors (Geronimus et al., 2006, 2015). Intersectional approaches related to cognitive decline and AD risk are important in the field of cognitive aging given that cumulative disadvantages (e.g., social factors, ethnoracial category, sex, and SES) may help explain how multiple disparities contribute to cognitive decline or “cognitive weathering” across age groups. Indeed, recent research has shown that ethnoracial disparities in a variety of cognitive domains is related to chronic stress and SES even in young adults (Letang et al., 2021).

An argument for a crystallized-fluid ability discrepancy score

The intersectionality of diverse characteristics can complicate the interpretation of low cognitive performance measures as a marker for risk for AD, especially if the measures do not account for educational and cultural biases (especially for non-Western, Educated, Industrialized, Rich, and Democratic (non-WEIRD) populations). At the same time, it may not be realistic to assume any task can be completely unbiased. Thus, a different strategy is to re-calibrate existing cognitive measures so that they can be more person-specific (Weintraub et al., 2018). Such a measure might account for an individual’s education level, linguistic ability, and lifelong intellectual ability within a cross-sectional sample to

infer longitudinal declines in cognition (Deary et al., 2013). In contrast to this idea, most studies investigating cross-sectional differences in cognition confound individual differences in declining cognition with adult level or lifelong intellectual functioning that are affected by education and linguistic ability (Deary et al., 2013). Specifically, an older adult with current low levels of cognition could be at this low level because (a) they declined from previously higher levels, (b) they have always had a low but stable level of cognition performing at their own capacity (i.e., non-artificial low level), or (c) they have an artificially low level of performance because of education, linguistic, cultural, or acculturation differences due to the testing process.

A crystallized-fluid ability discrepancy score might address these issues. This ability discrepancy score is calculated by subtracting one’s fluid ability from one’s crystallized ability (Dierckx et al., 2008; Lezak, 1995; McCarthy et al., 2005), both of which derive from validated cognitive assessments with good reliability. Fluid abilities refer to cognitive domains that require online processing and do not require previous knowledge, including processing speed, executive function, episodic memory, and reasoning (Johnson et al., 2004; Lezak, 1995; Wechsler, 1944). Conversely, a crystallized ability describes knowledge that has been gained through learning or experience and has often been measured through language tasks (Blair & Spreen, 1989; Ekstrom et al., 1976; Wechsler, 1944; Zachary & Shipley, 1986). These two categories of ability often are highly correlated with one another throughout one’s adult lifespan (Cattell, 1971; Deary et al., 2013; Kaufman & Horn, 1996); however, a discrepancy, or asymmetry, begins to emerge in older adults diagnosed with AD and related dementias (O’Carroll & Gilleard, 1986; Wechsler, 1944). Specifically, fluid abilities begin to decline early, whereas some crystallized abilities are relatively spared until later stages.

Extending these ideas to people at risk for AD, McDonough et al. (2016) found that larger ability discrepancy scores were associated with in vivo estimates of amyloid plaques and neurodegeneration – two biomarkers of AD. McDonough and Popp (2020) replicated these findings in a larger, independent sample and further showed that an ability discrepancy better predicted AD biomarkers than a composite episodic memory score and single-domain discrepancy scores (e.g., crystallized-episodic memory). Such consistent associations with AD biomarkers suggest the potential for an ability discrepancy score to aid in the identification of high-risk subgroups in the cognitively normal range. Specifically, we argue that both fluid and crystallized abilities often depend on language (e.g., verbal materials and instructions) and accounting for crystallized abilities helps adjust for levels of fluid ability that might be artificially lower than expected as previously outlined. While promising, previous tests validating this measure have used

majority non-Hispanic White samples with relatively high SES. Thus, it may not be a sufficient predictor of risk for AD or sensitive enough to detect intra-individual variability among the diverse aging population. Instead, it might be useful for detecting broad differences in cognitive decline and ranking individuals by their performance (e.g., large vs. small cognitive discrepancy).

Current study

The present study aimed to build upon the previous research by addressing the following questions. First, does an ability discrepancy reduce pre-existing differences in cognitive assessments often found in groups at risk for AD or other related dementias such as ethn racial category, SES, sex, and their intersection? Second, is an ability discrepancy measure equally predictive of AD symptom severity across varying levels of diverse characteristics? Lastly, does an ability discrepancy predict AD symptom severity even when controlling for a composite score of episodic memory or a fluid composite score?

Method

Study details

Since 2005, the National Alzheimer's Coordinating Center (NACC) has collected the Uniform Data Set (UDS) on participants from approximately 30 past and present US Alzheimer's Disease Research Centers (ADRC). All participants entering the ADRCs follow standard protocols by a trained clinician according to NACC guidelines. The data used in these analyses spanned all three versions of the UDS from September 2005 to December 2016 and included data from 35 different ADRCs.

Participants

Participants in the final sample were included if they had all crystallized measures and at least four of the five fluid measures of interest, years of education to serve as a proxy for SES, self-reported ethn racial category, self-reported sex, and a Clinical Dementia Rating (CDR)[®] dementia staging instrument (Morris, 1993). The CDR is a semi-structured interview given to both participants and informants to assess cognitive functioning and daily functions (memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care). The global CDR score groups people into five categories of dementia staging (0, no impairment; 0.5, questionable impairment; 1, mild impairment; 2, moderate impairment; and 3, severe impairment). The sum of boxes score (CDR-SOB) is obtained by

summing each of the domain box scores ranging from 0 to 18 to form a symptom severity score.

Exclusion criteria were a primary language self-reported as anything other than English, and "other" ethn racial category. Participants differed in their reasoning for coming to the ADRC (e.g., clinical evaluation vs. research participation), which might lead to different profiles of cognition. To reduce this variability, we also restricted participants to those whose primary reason was for research.

The final cross-sectional sample consisted of 14,257 individuals aged 60–104 years. Table S1 (see Online Supplemental Material (OSM)) includes the demographic data for cognitively normal older adults (CDR global score = 0) and those classified as on the AD spectrum (CDR global score > 0). We note that while the percentage of participants representing ethn racial minorities and lower education backgrounds is small (which is often found even in nationally representative datasets; Zahodne et al., 2017), the large sample contains a sufficient number of participants from these backgrounds to test for differences between these categories.

Neuropsychological testing

Cognitive measures were chosen a priori from available fluid and crystallized measures based on McDonough and Popp (2020) to form the ability discrepancy scores. These measures included logical memory (immediate and delayed) that measures recall of details in two stories (Wechsler, 1997), category fluency that measures the number of animals and vegetables that can be named within 1 min (Benton, 1983), Trail Making Test that measures the speed at which letters and numbers can be sequentially connected via a pencil (Reitan, 1979), WAIS-R Digit Symbol that measures the speed of matching a symbol to a digit using a key (Wechsler, 1997), and the 30-item Boston Naming Test that measures the number of named picture objects of varying difficulty level (Kaplan et al., 1983).

For each participant, an ability discrepancy score was calculated by standardizing each measure within the cognitively normal group (CDR = 0). In the whole sample, individual task measures were sometimes missing. Because these scores do not rely on any single task or cognitive domain (cf. McDonough & Popp, 2020; McDonough et al., 2016), all measures that were available were averaged together to form fluid and crystallized ability scores. Crystallized ability was composed of the language tasks (i.e., category fluency and Boston Naming Test) and fluid ability was composed of the other measures. While the tasks used to form this crystallized ability composite are sometimes considered fluid ability scores, they were most representative of language abilities and have been validated to be used as appropriate calibration tasks for the other types of fluid abilities used

here (McDonough & Popp, 2020). Notably, these language tasks load together with more traditional tasks of crystallized ability like word reading (McDonough & Popp, 2020). Moreover, this study showed that including these language tasks into a crystallized ability score to form an ability discrepancy score resulted in greater sensitivity to beta-amyloid accumulation and cortical thinning in AD signature regions than the inclusion of word reading alone.

Discrepancy scores for each subject were calculated by subtracting each fluid composite score from the crystallized composite score. Greater scores represented a larger discrepancy in ability, representing a precipitous decline in cognitive functioning (e.g., Kaufman & Horn, 1996; Matarazzo & Herman, 1985; Schretlen et al., 1994). Lower scores, on the other hand, represented more successful cognitive aging. A separate memory composite score also was calculated by averaging the immediate and delayed logical memory scores to be used as a comparison.

Statistical analyses

In the first analysis, we provide a foundation for the assumption that an ability discrepancy score might minimize existing disparities in cognition between sex, ethnorracial category, and education levels. We also tested the assumption that age cohort does not strongly alter these effects by stratifying the analyses by age decade. For these analyses, we conducted linear mixed models to test the main effects of each demographic factor (between subjects) and composite measure (ability discrepancy, fluid composite, memory composite; within subjects) on performance. Explicitly testing the different composite measures provides evidence as to whether group disparities in commonly used cognitive composites (fluid and memory) are larger than any disparities in an ability discrepancy score. Additional Bayesian tests were conducted to provide evidence that groups differed in each of the composite measures.

Next, hierarchical linear regression models were conducted to investigate the impact of sex, education, Hispanic ethnicity, and race (Black, Asian) on level of ability discrepancy, and whether these measures of diversity might moderate the predictability of ability discrepancy level on dementia severity (CDR-SOB). We included all higher-order interactions of our diversity measures to assess potential differences due to intersectionality of these characteristics. Hierarchical models were used to build each regression model in sequential steps while also allowing us to see whether any missing values in the covariates led to selection bias effects. To the extent that the predictors of interest remained significant in the fully adjusted models compared with the unadjusted models, we would conclude that the results would be robust to both the included covariates and

the slight change in sample characteristics due to listwise removal of participants with missing data. The first step of each regression model included only the key predictors of interest and the outcome variable. The second step included interactions between each of the key predictors (if applicable). Critically, including interactions of diversity characteristics serves as our method of investigating intersectionality. Although intersectionality is sometimes assessed by creating separate subgroups, this method provides a systematic manner to assess multiple combinations of intersecting identities through different orders of interactions. The last step included all of the available covariates that might impact the interpretation of sex, education, or ethnorracial category effects. We set the alpha level to 0.005 in accordance with recent recommendations (Benjamin et al., 2018).

The first hierarchical model tested the extent that level of ability discrepancy differed with sex, education, and ethnorracial category. In this model, each of the measures of diversity and their interactions were entered as the independent variable and ability discrepancy was entered as the outcome variable. Covariates were chosen based on their previous associations with cognitive status such as risk for comorbidity, risk for cardiovascular disease, risk for cognitive impairment other than dementia, and functional status. The specific covariates were dummy coded when appropriate and included: age; year visiting the center; permanent move to nursing home; average number of packs smoked per day; smoked cigarettes in past 30 days; total years smoked cigarettes; stroke (remote, recent, absent); Parkinson's disease present; active depression in the last 2 years; marital status (married, widowed, divorced, separated, never married, living as married); type of residence (private residence, retirement/independent living community, assisted living); level of independence (independent, assistance with complex activities, assistance with basic activities, completely dependent); hypertension (remote, recent, absent); alcohol abuse (remote, recent, absent); and diabetes (remote, recent, absent).

The second and third hierarchical models tested the extent that the level of ability discrepancy predicted AD symptom severity and whether our measures of diversity impacted this discrepancy-cognitive status relationship. Ability discrepancy, the measures of diversity, and their interactions were entered as predictors. CDR-SOB score was entered as an outcome variable. The same covariates were used in these two models as in the first model. In the final model, an episodic memory composite score also was entered as a predictor to test whether controlling for episodic memory eliminated any of the effects. For all models, we report standardized betas and interpret values < 0.1 as very small, between 0.1 and 0.2 as small, between 0.2 and 0.3 as moderate, and greater than 0.3 as large (Gignac & Szodorai, 2016).

Results

Consistent with prior work (e.g., Cattell, 1971; Kaufman & Horn, 1996), the crystallized ability and average fluid ability scores were positively correlated with one another in this cognitively normal sample, $r(8,079) = 0.60$, 95% CI [0.59, 0.61], $p < 0.001$. Subtracting the average fluid composite score from the crystallized ability score resulted in an ability discrepancy score for each subject that ranged from -2.11 to 2.29.

Comparison of diversity characteristics on cognitive measure stratified by age decade

The first set of analyses tested the extent of group differences in cognitive performance as a function of sex, education, and ethnoracial category in cognitively normal individuals ($CDR = 0$) and stratified across age decade (see Figs. S1 and S2 in the OSM). For adults in their 60 s ($N = 2,777$), we found significant interactions between cognitive measure and each of the group characteristics other than Hispanic ethnicity at the set alpha level of 0.005 (Table 1). These interactions were driven by a larger difference among sexes (female > male), education (higher > lower), identifying as Black (Black < White), and identifying as Asian

(Asian > White) for the composite scores than the ability discrepancy score. Thus, group differences were largely reduced when using an ability discrepancy score compared with single and multi-domain composite scores. Congruent with these findings, Bayes factor outputs for the sixties decade suggests anecdotal evidence for no group differences in ability discrepancy scores, but moderate to “extreme” evidence for most of the group differences in fluid and memory composite scores (Table S2, OSM). The exceptions include only anecdotal evidence for group differences on the fluid composite score between Hispanic White and non-Hispanic White Americans and moderate evidence for no differences on the memory composite score between Black Americans and non-Hispanic White Americans. Similar patterns of results were found for adults aged in their seventies ($N = 3,336$), 80 years and older ($N = 1,968$), and when including the whole sample controlling for age ($N = 8,081$). The primary exception was that in the seventies and eighties age groups, the fluid composite did not strongly differ from ability discrepancy for Black Americans.

Overall, most comparisons revealed group differences in fluid and memory composite scores across each of the decades consistent with prior research (e.g., Helms, 1992; Jones, 2003). Critically, nearly every group difference was smaller for ability discrepancy scores in cognitively normal

Table 1 Comparisons between demographic characteristics and cognitive composite score to an ability discrepancy score

	Age group							
	60 s		70 s		80 +		All	
	β	SE	β	SE	β	SE	β	SE
Fluid	.644**	.014	.424**	.014	.158**	.019	.434**	.009
Memory	.651**	.014	.513**	.014	.381**	.019	.528**	.009
Sex	.019	.010	.017	.010	.006	.014	.011	.006
Education	.031*	.010	.021	.010	-.005	.014	.007	.006
Hispanic	-.004	.010	.004	.010	-.015	.014	-.007	.006
Black	.009	.010	-.005	.010	-.004	.013	-.004	.006
Asian	-.009	.010	-.021	.010	-.014	.013	-.017	.006
Fluid × Sex	.044*	.014	.038	.014	.031	.019	.047**	.009
Memory × Sex	.104**	.014	.108**	.014	.107**	.019	.111**	.009
Fluid × Education	.111**	.014	.181**	.014	.192**	.019	.180**	.009
Memory × Education	.135**	.014	.188**	.014	.195**	.019	.182**	.009
Fluid × Hispanic	-.014	.014	-.025	.014	.000	.019	-.007	.009
Memory × Hispanic	-.034	.014	-.024	.014	.005	.019	-.018	.009
Fluid × Black	-.065**	.014	-.034	.014	.011	.019	-.027*	.009
Memory × Black	-.039	.014	-.019	.014	-.001	.019	-.018	.009
Fluid × Asian	.031	.014	.056**	.014	.047	.019	.048**	.009
Memory × Asian	.049**	.014	.083**	.014	.067**	.019	.068**	.009

Notes. * $p < .005$, ** $p < .001$. Reference group for Fluid and Memory composite scores was the ability discrepancy score. The analysis with all age groups also controlled for chronological age

older adults, regardless of age decade (Fig. 1). These results suggest that considering crystallized abilities might level the playing field for assessing baseline cognition across diverse demographic groups.

Intersectional characteristics predicting ability discrepancy, fluid cognition, and episodic memory

We next tested the extent that interactions among these diverse characteristics (thus representing intersectionality) were associated with varying levels of ability discrepancy and compared it with episodic memory. The first HLM only included the main effects of these characteristics in cognitively normal older adults. Standardized beta coefficients and standard errors for each factor can be found in Table S3 (OSM). The first unadjusted model was significant, $F(6, 8,074) = 3.23, MSE = 0.53, p = 0.0036$. Individuals with higher education ($p = 0.0032$) were more likely to have higher ability discrepancy scores. The second unadjusted model included interactions among each factor to identify differences in ability score associated with intersectionality of diverse characteristics. The second model was significant, $F(16, 8,064) = 2.31, MSE = 0.53, p = 0.0022$. Of the main

effects, years of education remained significant ($p = 0.0021$). Of the interactions entered into the model, only the education \times sex interaction reached significance ($p = 0.00066$). As seen in Fig. 2, this interaction occurred because the relationship between years of education and ability discrepancy was stronger for females than males. Specifically, higher education was associated with a larger ability discrepancy score in females ($\beta = 0.033, SE = 0.008, p < 0.001$), and no relationship was found in males ($\beta = -0.009, SE = 0.010, p = 0.37$). The third model included our covariates and was significant, $F(42, 7,468) = 2.11, MSE = 0.53, p < 0.0001$. Neither the size nor the significance of the effects changed from the second to the third model, suggesting the effects from Model 2 were robust. When the episodic memory composite score was included in the model (Model 4, $F(43, 7,467) = 12.91, MSE = 0.51, p < 0.0001$), the education \times sex interaction was still significant ($p = 0.002$), as was the effect of episodic memory ($p < 0.0001$). These findings suggest that intersectionality of sex and education on the ability discrepancy measure were largely independent of traditional cognitive measures of early declines in AD found by episodic memory. Because the ability discrepancy score did, in part, consist of episodic memory, it is not surprising that the episodic

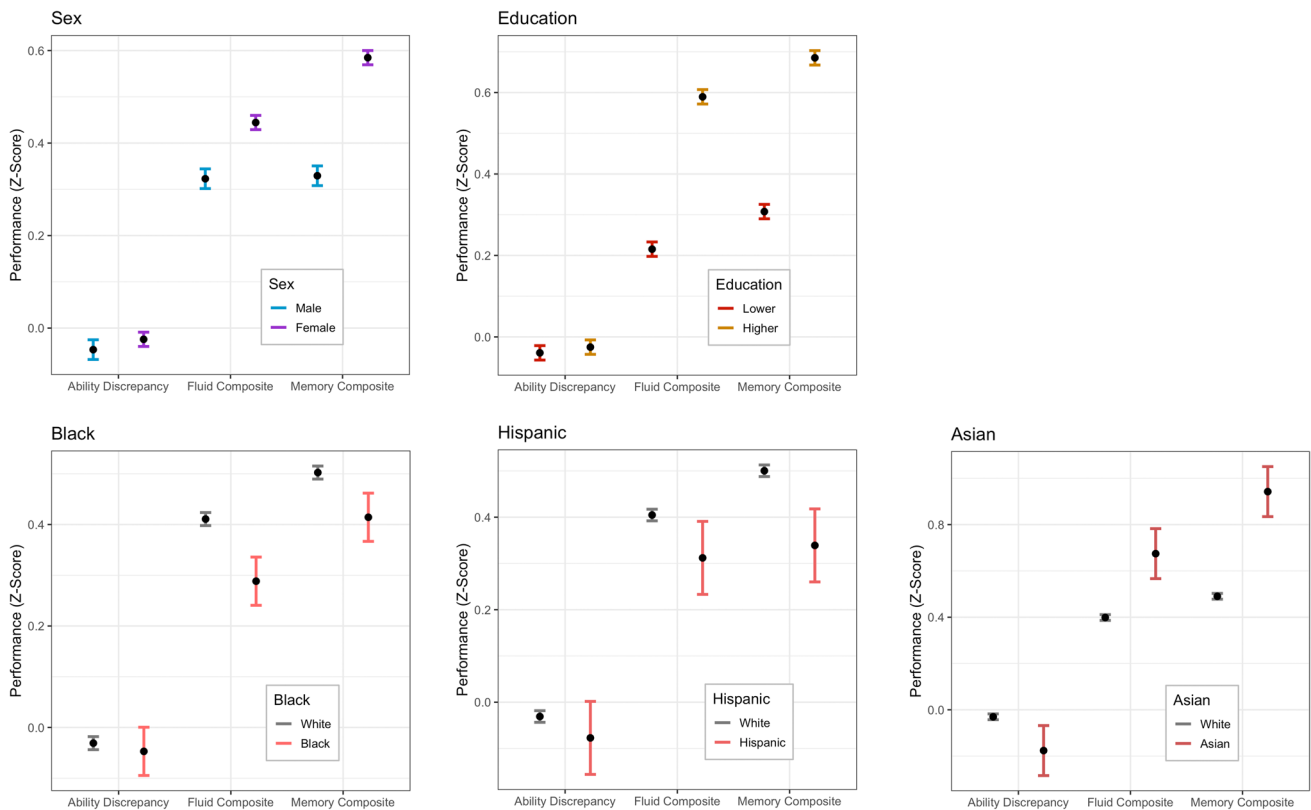


Fig. 1 Comparison between three measures of cognition (ability discrepancy, fluid composite, episodic memory composite) as a function of sex, education, and ethnorracial category. In each panel, disparities

between cognition scores are reduced when using an ability discrepancy score compared to a fluid or memory composite score

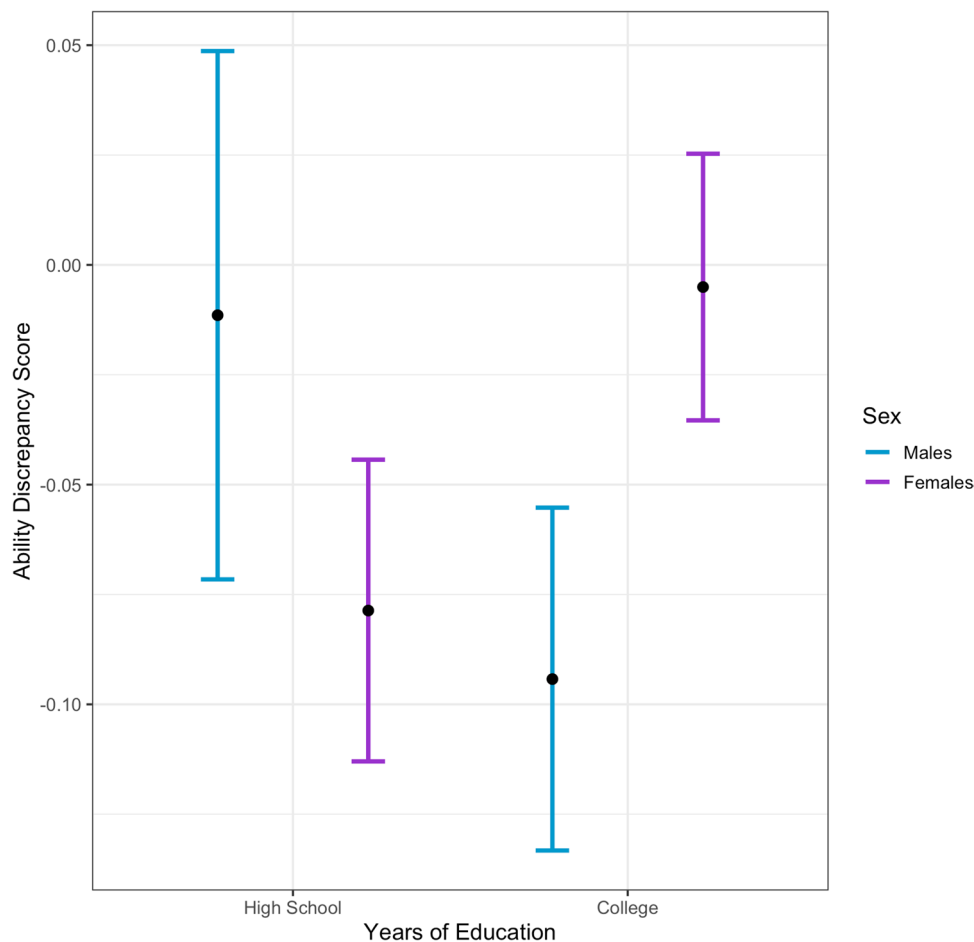


Fig. 2 Interaction of Education and Sex on Level of Ability Discrepancy. While males with lower education had higher ability discrepancy scores than females, males with higher education had lower ability discrepancy scores than females. To the extent that a greater

ability discrepancy score is indicative of greater risk of developing dementia, higher education does not appear to be protective for females

memory composite score also significantly predicted ability discrepancy.

In two final models, we tested the extent that intersectionality differences also might be found on fluid cognition (Model 5) and episodic memory (Model 6) composite scores after controlling for covariates. Each of the models was significant ($ps < 0.001$). Although no significant interactions were found, both models revealed main effects of age ($ps < 0.001$), sex ($ps < 0.001$), years of education ($ps < 0.001$), identifying as Hispanic ethnicity ($ps < 0.005$), and identifying as Asian ($ps < 0.001$). The episodic memory score had an additional main effect of identifying as Black ($ps < 0.001$). While the fluid cognition and episodic memory composite scores revealed similar patterns in comparison with each other, they differed markedly from an ability discrepancy. Notably, whereas the mean composite scores showed many large group differences (as readily apparent in Figs. 1, S1 (OSM), and S2 (OSM)), the ability discrepancy

score showed differences only for females with more education relative to other groups also with more education.

Impact of level of ability discrepancy on AD symptom severity as a function of diversity

The next set of analyses assessed the extent that level of ability discrepancy predicted AD symptom severity across the AD spectrum (Models 7–10). We also tested whether this discrepancy-symptom relationship was moderated by sex, education, and ethnoracial category. The first regression only included the main effect of ability discrepancy. Standardized beta coefficients and standard errors for each factor can be found in Table S4 (OSM). The first model (Model 7) was significant, $F(1, 14,255) = 31.00$, $MSE = 2.24$, $p < 0.0001$. A positive relationship was found between ability discrepancy and CDR-SOB score ($p < 0.0001$). The second model (Model 8) was also significant, $F(30, 14,226) = 13.95$, $MSE = 2.197$,

$p < 0.0001$. The highest order interaction in this model was an ability discrepancy \times sex \times education \times Hispanic ethnicity interaction ($p < 0.001$; Table S4 (OSM)). To better understand this interaction, two additional regressions were conducted separately for non-Hispanic and Hispanic White groups (see Fig. 3). For non-Hispanic White Americans ($N = 12,729$), a main effect of ability discrepancy was found on CDR-SOB score ($\beta = 0.080$, $MSE = 0.020$, $p < 0.001$) but neither the ability discrepancy \times sex ($p = 0.033$), ability discrepancy \times education ($p = 0.34$), nor the ability discrepancy \times education \times sex interactions ($p = 0.27$) reached our threshold of significance. For Hispanic Americans ($N = 343$), we see a different pattern. Only the ability discrepancy \times education \times sex interaction reached significance ($\beta = -0.420$, $MSE = 0.124$, $p = 0.001$), which was due to no significant effect of ability discrepancy on CDR-SOB score for females ($ps > 0.06$) but a nearly significant ability discrepancy \times education interaction for males ($\beta = 0.537$, $MSE = 0.192$, $p = 0.0058$). Hispanic males with at least a college degree did not show a relationship between an ability

discrepancy and AD symptom severity. However, for Hispanic males with a high school degree or lower, a greater ability discrepancy score was associated with less severe AD symptoms (the opposite direction to that in non-Hispanic White Americans and less educated Hispanic females). As shown in Table S4 (OSM), these interaction patterns were still significant after including covariates (Model 9; four-way interaction $p < 0.001$) and episodic memory (Model 10; four-way interaction $p < 0.001$).

In the last two models, we tested the extent that fluid cognition (Model 11) and episodic memory (Model 12) composite scores equally predicted the CDR-SOB score across the diverse groups. Beta values and standard errors can be found in Table S4 (OSM). For the fluid composite score, the highest order interaction was a significant cognitive marker \times sex \times Black interaction ($p < 0.001$). While all groups showed a negative relationship between fluid composite and AD symptom severity, non-Hispanic White males showed a stronger relationship than non-Hispanic White females, whereas Black females showed a stronger

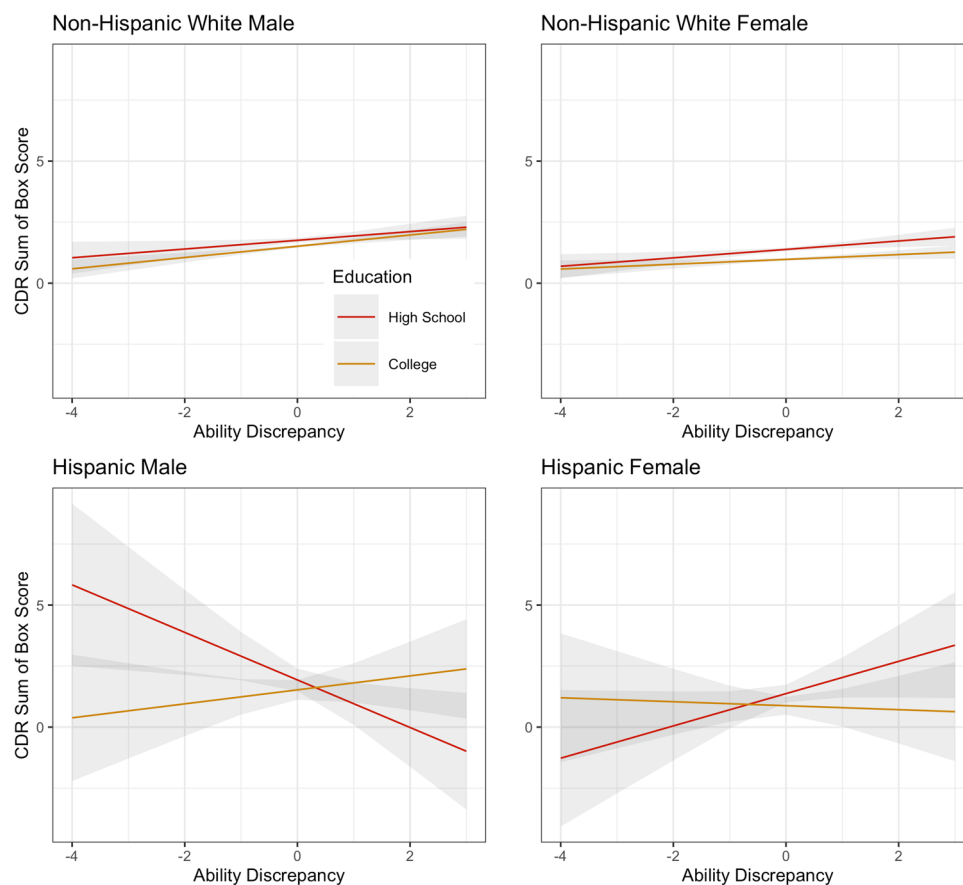


Fig. 3 Four-way interaction between Ability Discrepancy, Sex, Education, and Hispanic ethnicity on CDR Sum of Boxes score. For non-Hispanic White Americans, a higher ability discrepancy was associated with greater Alzheimer’s disease (AD) symptom severity

regardless of sex and educational level. In contrast, Hispanic males with lower education exhibited a relationship in the opposite direction and Hispanic females with higher education exhibited no relationship

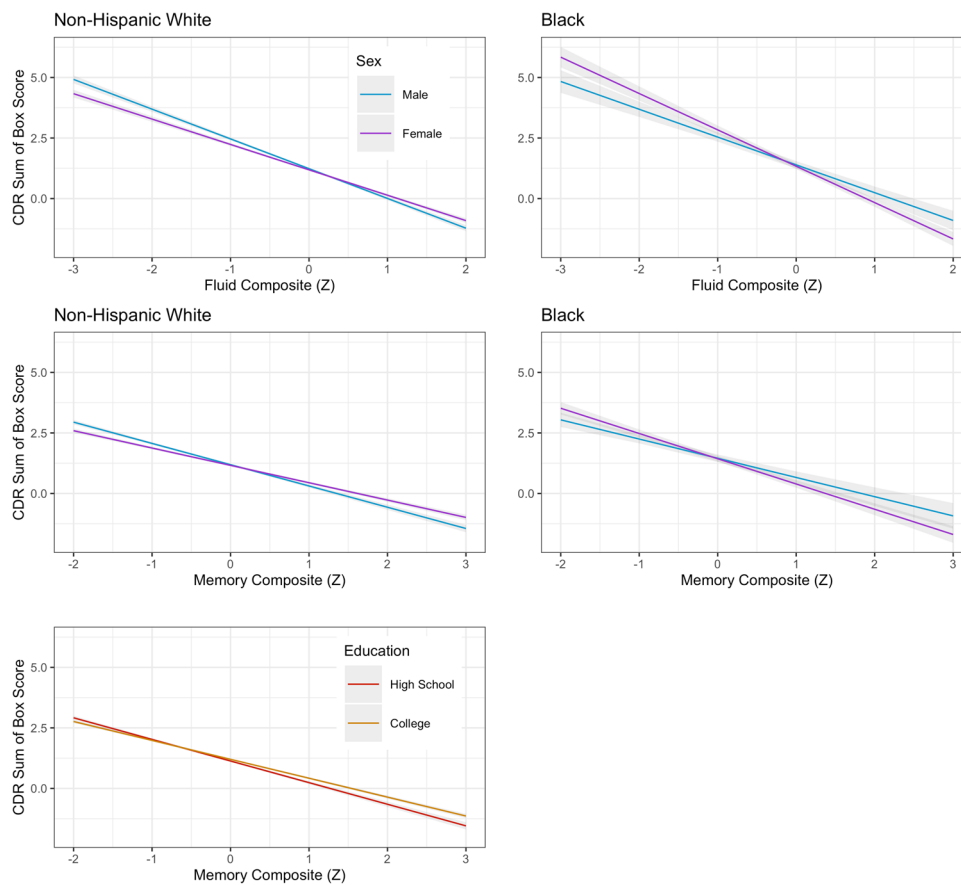


Fig. 4 Interactions between Cognitive Marker, Sex, and Race and between Episodic Memory \times Education on CDR Sum of Boxes score. Non-Hispanic White males showed a stronger relationship than non-Hispanic White females whereas Black females showed a stronger

relationship than Black males. Older adults with lower education had a stronger negative relationship between episodic memory and Alzheimer's disease (AD) symptoms than older adults with higher education

relationship than Black males (see Fig. 4). For the episodic memory composite score, the same three-way interaction between cognitive marker \times sex \times Black was found ($p < 0.001$; see Fig. 4). However, an additional episodic memory \times education interaction was significant ($p < 0.001$), such that older adults with lower education had a stronger negative relationship between episodic memory and AD symptoms than older adults with higher education (see Fig. 4).

Together, these results suggest that none of the cognitive markers tested show the same relationships with AD symptom severity across all groups of older adults. However, the ability discrepancy score complemented the more traditional mean composites with the exception of being less sensitive to AD symptom severity in older Hispanic males with lower education and older Hispanic females with higher education. In contrast, the fluid and episodic memory composite scores were less sensitive (a) in non-Hispanic White females, (b) in Black males, and (c) for more educated older adults (for episodic memory only).

Discussion

The present study investigated whether a crystallized-fluid ability discrepancy measure (a) differed across intersecting identities that have previously been shown to increase the risk for AD in late life, and (b) predicted severity of AD symptoms similarly across the subgroups. We also compared this measure to commonly used mean composite measures of fluid and episodic memory abilities.

An ability discrepancy score levels the playing field across diverse characteristics

First, we found that the fluid and memory composite scores generally exhibited large group differences with sex, education, and ethnoracial category in cognitively normal adults. In contrast, the ability discrepancy score showed much smaller group differences. These patterns were supported by Bayes factor values that consistently indicated evidence toward no group differences in the ability discrepancy score,

but for moderate to extreme group differences in the mean composite scores with the exceptions of differences for Hispanic and Black Americans relative to non-Hispanic White Americans. Comparisons between the three types of measures showed that the ability discrepancy score leveled the playing field for baseline levels of cognition relative to fluid and memory composite measures. One exception was for ethnoracial category, especially in Hispanic Americans. The relatively small benefit of an ability discrepancy score between ethnoracial differences stands in contrast to the large benefits the score has in reducing differences between sexes and education levels. This pattern suggests that although a crystallized-fluid ability discrepancy score reduces baseline differences among many subgroups at risk for AD, this measure might be especially useful in reducing differences due to sex and education.

Baseline cognitive level is an important predictor of dementia risk (e.g., Weuve et al., 2018). Individuals who start out with lower levels of cognition have less room to fall until those cognitive declines begin to interfere with daily functions. For example, Weuve et al. (2018) showed that some Black Americans are at a higher risk of developing a dementia diagnosis than non-Hispanic White Americans, and they argued this pattern arose because of their lower baseline levels of cognition. From this perspective, lower levels of cognition in marginalized groups might be due to lower levels of cognitive reserve. However, this argument assumes that baseline cognition was accurately assessed across all diverse subgroups. Some researchers argue that neuropsychological tests are culturally biased (intentionally or unintentionally) to favor the creator of those tests – namely, educated non-Hispanic White and English-speaking men (Helms, 1992; Jones, 2003). These biases can be manifested in what constitutes intellectual ability, the language standards used in testing instructions and materials, the cultural relevance of the stimuli, and even what a correct answer is assumed to be. Thus, higher performance on neuropsychological tests might be more of a measure of acculturation to the dominant culture than to actual level of cognitive ability. As stated by a recent perspective paper, very little research has been conducted on cognitive reserve in ethnoracial categories (and other marginalized groups) to fully understand how to balance these different perspectives (Babulal et al., 2019). Although the link with cognitive reserve is unclear, the present findings show that an ability discrepancy score is one way to successfully reduce such baseline differences between subgroups. By using language tasks to correct for fluid ability, perhaps one can directly assess how much language is negatively impacting cognitive assessments via differences in acculturation to the testing environment or language standards.

Similar corrections to cognitive measures have been made in health disparities research. Specifically, researchers have advocated for statistically controlling for quality of education through reading level (e.g., Manly et al., 2004). Reading level has been proposed as a better indicator of education quality than years of education because many older ethnoracial minorities grew up in segregated school systems that had fewer hours in a day and fewer school days in a year than schools for non-Hispanic White Americans (e.g., Crowe et al., 2013). Example proxies for education quality include word pronunciation tasks or vocabulary tasks. Such corrections can eliminate or at least attenuate mean levels of cognitive performance between subgroups (e.g., Crowe et al., 2013; Fyffe et al., 2011; Garcia et al., 2021; Manly et al., 2004). Although measures of reading level were not available in this data set, McDonough and Popp (2020) previously showed that composite measure of multiple language abilities (including the Boston Naming Test and Category Fluency) loaded onto the same factor as word reading and were sufficient to create the crystallized portion of the ability discrepancy score that was more sensitive to AD biomarkers than word reading alone. Using objective language tasks to construct an ability discrepancy score, as was done in the present study, likely was a key factor in reducing the ethnoracial and educational level gap found in the mean composite scores.

Ability discrepancies differ at the intersection of education and sex

Our second main finding was that a greater ability discrepancy score (i.e., greater AD risk) was associated with higher education for females but no relationship in males. Instead of interpreting the meaning of an ability discrepancy in highly educated females, this deviation from the other groups might simply indicate that an ability discrepancy does not level the playing field for this subgroup. Females often have better language abilities than males (e.g., Asperholm et al., 2019a, b) that may have inflated their crystallized scores (and/or perhaps increased their cognitive reserve). If so, females would naturally show a greater ability discrepancy score than males. From this perspective, it is worth emphasizing that whereas only females with high education appear to differ in baseline ability discrepancy levels, the mean composite scores were much more susceptible to potentially artificial baseline differences not only in sex and years of education, but also in age and each of the ethnoracial categories assessed. Ultimately, these findings suggest that an ability discrepancy, as constructed here, should be used with caution in research and clinical settings for highly educated females.

Ability discrepancy predicts AD symptom severity

Lastly, we assessed the relationship between ability discrepancy and AD symptom severity across diverse sets of characteristics. This analysis can inform whether an ability discrepancy is sensitive to AD symptom severity across the AD spectrum and test whether this relationship is similar across diverse and intersecting backgrounds. Using the CDR-SOB score as our measure of AD symptom severity, we found that a greater ability discrepancy score was associated with more AD symptoms in unadjusted analyses. When further probing the intersecting effects of subgroups, we found an interaction between ability discrepancy, sex, and education for Hispanic Americans. Breaking down this interaction revealed that Hispanic males with at least a college level of education and Hispanic females with a high school degree or less showed similar positive relationships as the other participants in the sample (i.e., greater ability discrepancy and more severe AD symptoms). However, two subgroups deviated from these patterns: Hispanic males with a high school degree or lower showed the reverse pattern (greater ability discrepancy, fewer AD symptoms) and Hispanic women with at least a college education evidenced no relationship between an ability discrepancy score and AD symptom severity. At minimum, these findings suggest that while an ability discrepancy score is generally predictive of AD symptom severity across many diverse subgroups, this association does not hold for some intersecting identities, and should be considered cautiously by researchers and clinicians.

A broad span of literature suggests that, as a group, those with a lower education level (and lower SES, more generally) and those who identify as being Hispanic have greater prevalence and incidence rates of an AD diagnosis compared with non-Hispanic White Americans (Matthews et al., 2018; Perkins et al., 1997; Steenland et al., 2016; Tang et al., 2001). Although Hispanic Americans, as a group, have lower education levels compared to other ethnoracial categories in the USA (Braveman et al., 2010), our findings suggest that an ability discrepancy measure might be less sensitive to risk for AD in this subgroup. One reason for this decrease in sensitivity is that the cognitive profile of Hispanic Americans may differ in AD compared to non-Hispanic White Americans. Indeed, Hispanic Americans can sometimes exhibit less cognitive impairment than non-Hispanic White Americans despite having autopsy-confirmed AD (Weissberger et al., 2019). One speculation is that bilingualism among some Hispanic Americans might confer a greater degree of cognitive reserve, thereby reducing AD symptoms (e.g., Bialystok et al., 2007; Schweizer et al., 2012). The present data set did not have information about bilingualism or multilingualism to test these ideas.

Comparing ability discrepancy with other cognitive domains

We also examined whether episodic memory, a domain of cognition known to decline early in preclinical AD (e.g., Boraxbekk et al., 2015; Sperling et al., 2011; Weintraub et al., 2018), could explain the relationship between an ability discrepancy and AD symptoms. Controlling for episodic memory did not alter any of the findings. However, episodic memory was associated with AD symptom severity (as measured by the CDR-SOB Score) in these models, suggesting that both an ability discrepancy score and estimates of episodic memory might contribute unique information if used to detect people at risk for AD or used as an outcome variable for potential treatments (see also Salthouse & Becker, 1998). This conclusion is supported by a study that found that greater non-memory impairments at baseline were associated with steeper annual rates of decline in CDR score than baseline memory scores (Scheltens et al., 2018). The authors suggested that non-memory impairments lead to faster AD symptom severity than memory impairments.

Although both episodic memory and fluid cognition were related to AD symptoms, these effects were qualified by interactions with ethnoracial category and sex. Specifically, the associations were relatively weaker for non-Hispanic White females and Black males. These interactions suggest that, like an ability discrepancy, the cognitive markers are not equally sensitive to AD symptoms across subgroups. The reasons, however, for these weaker associations are not clear. Regardless, it is clear that the unique relationships between episodic memory/fluid cognition and AD symptoms should be considered in the context of potential biases in measurement, which is relevant to recent recommendations to include such scores as key outcome measures in clinical trials (e.g., Chhetri et al., 2018; Donohue et al., 2014). The present study suggests that such fluid composite scores might not be equally sensitive to treatment outcomes in all individuals, especially those at most risk for AD such as ethnoracial minorities, women, and those with lower levels of cognitive reserve.

Concerns regarding the formation of an ability discrepancy score

Despite its potential, several concerns exist to its formation and implementation. The crystallized ability score used here was comprised of tasks relying on language performance from which to anchor other fluid cognitive abilities. Our primary argument is that such anchoring can help correct for premorbid levels of intelligence and reduce the influence of language difficulties that can lead to an artificially lower estimation of other fluid abilities. However, such anchoring also provides limitations to understand cognitive decline on

the AD spectrum. Specifically, some productive language tasks including the Boston Naming Test and Verbal Fluency used here have been used as early indicators of AD-related declines in cognition (Jacobs et al., 1995), although sometimes not as early as episodic memory (Hamel et al., 2015; Howieson et al., 2008; Mistridis et al., 2015). From this perspective, one might argue that subtracting out such important cognitive indicators in the discrepancy score used here may contradict its use. Indeed, we have argued previously that multiple measures of receptive language tasks like vocabulary or word pronunciation would make for a more suitable crystallized ability composite (McDonough & Popp, 2020; McDonough et al., 2016). Unfortunately, multiple versions of such receptive language abilities often are not available in standardized neuropsychological batteries. Thus, the present study relied on tasks that greatly depend on language to form these composite scores. Although perhaps counterintuitive, we have previously validated the use of such composites by showing that a discrepancy score using the same productive language tasks was related to amyloid accumulation and neurodegeneration in AD signature regions, while a language discrepancy score using word pronunciation as the only measure of crystallized ability did not correlate with those same AD biomarkers (McDonough & Popp, 2020). This finding also is consistent with a meta-analysis that showed amyloid was not correlated with semantic memory (Hedden et al., 2013) and in autopsy measures showing no relationships between the productive language tasks and amyloid plaques, diffuse senile plaques, or limbic neurofibrillary tangles (Price et al., 2009).

If language tasks can be used as indicators of risk for AD, then how can we interpret an ability discrepancy score that uses language tasks as a method of correction? One possibility is that an ability discrepancy score measures earlier subtle declines in cognition before declines in language abilities. Indeed, multiple longitudinal studies have suggested that declines in language abilities occur after declines in other cognitive domains like episodic memory (Hamel et al., 2015; Howieson et al., 2008; Mistridis et al., 2015). A second possibility is that language deficits in AD may represent a subtype of AD. Converging evidence suggests that one infrequent AD variant (occurring in 19–22.4% of AD cases) increases atrophy primarily in the left temporal lobe and is associated with lower language performance (Vogel et al., 2021; Zhang et al., 2021). This recent research is consistent with lesion-mapping studies and volumetric studies that point selectively to a role of left lateral temporal cortex involved in productive language tasks like the Boston Naming Test (Baldo et al., 2013; Seidenberg et al., 2005). The left temporal AD variant was also associated with fewer participants with abnormal amyloid levels (Zhang et al., 2021), consistent with the lack of association between language deficits and AD pathology found in previous studies (Hedden

et al., 2013; McDonough & Popp, 2020; Price et al., 2009). Thus, an ability discrepancy score, as defined in the present study, may not capture early cognitive decline in this AD subtype, but may capture early cognitive decline in typical AD expressions and perhaps other subtypes not associated with language deficits.

Another criticism is that an ability discrepancy score relies on the use of difference scores (Frazen et al., 1997; Rogosa & Willett, 1983). Difference scores can sometimes reduce between-subject differences and compound measurement error from each score that makes up the subtraction (Hedge et al., 2018). However, we propose that composite fluid and crystallized scores can be used to reduce measurement error in comparison with the single-task scores commonly criticized (McDonough & Popp, 2020). Additionally, even those who criticize the use of difference scores acknowledge that such scores might be useful when subtracting a baseline measure to control for unwanted between-subject variance (Hedge et al., 2018). In this case, a crystallized/fluid ability subtraction is theoretically motivated based on the cognitive domains that often decline early in the AD process (i.e., fluid abilities) versus later (i.e., crystallized abilities).

Study limitations

The study should be interpreted considering its limitations. First, while the sample from the NACC is quite large and diverse, it should not be interpreted as a nationally representative sample. Participants in the sample were selected based on their interest in helping research rather than randomly selected from the community. Second, the sample also consists of residents predominantly in urban areas near medical centers. Thus, the sample contains large numbers of highly educated non-Hispanic White Americans. Third, while this study was the first to examine the predictive value of ability discrepancy scores across diverse groups of older adults, we used baseline data only for these inferences. Longitudinal investigations are needed to further study how diversity metrics influence the predictability of ability discrepancy scores on late-life cognition and conversion to AD. Fourth, several measures were unavailable in the NACC data set including word reading or vocabulary measures to compare with the other language measures, memory recognition scores to complement recall scores, and information about bilingualism or multilingualism. Fifth, all the tests used here were originally developed by educated non-Hispanic White and English-speaking men, thereby constraining how we conceptualize both fluid and crystallized abilities. Novel cognitive tests that take a broader and more inclusive perspective are needed to fully appreciate one's abilities and how they relate to ADRD. Lastly, the effect sizes were relatively small. Other researchers have noted an apparent inverse

relationship between sample size and effect size (e.g., Karlamangla et al., 2014), possibly due to an increase in noise in the measures across sites or increases in heterogeneity of larger samples.

Conclusion

As the number of ethnoracial minorities and SES gaps continues to rise in the USA, increased attention also is being paid to health disparities across these categories and how they intersect with other characteristics like sex. In the case of AD, Black Americans, Hispanic Americans, and females are more likely to be diagnosed with the disease than non-Hispanic White males, but the origins of these increased risks are still being discovered, including possible misdiagnoses underlying these apparent risks (e.g., Goldstein et al., 2014; Kiselica et al., 2021). Research also is slowly uncovering how these health disparities intersect, thus giving more importance to the notion that not all individuals can be simply lumped into one or two categories of people (McDonough et al., 2021). Rather, different types of disparities have the potential to interact and lead to poorer cognitive outcomes (Matthews et al., 2018). However, to understand the origins of these health disparities and eliminate them, a person-specific measure of cognition that accounts for biases in culture and SES is needed. The present study provided initial evidence that an ability discrepancy score might serve as such a measure that is built on the foundation of clinical neuropsychology, and may enhance existing practices by modifying the traditional inherent biases that arise from lack of access, interest, and inclusion of diverse populations.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13421-022-01304-3>.

Authors' contributions IMM assisted with the conception and design of the study; the acquisition, analysis, and interpretation of data; and drafting and revising the manuscript. SLC, SLG, and TEP assisted with the drafting and revising the manuscript. ERH assisted with the revising of the manuscript.

Funding The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA-funded ADRCs: P30 AG019610 (PI Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P50 AG005134 (PI Bradley Hyman, MD, PhD), P50 AG016574 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Thomas Wisniewski, MD), P30 AG013854 (PI M. Marsel Mesulam, MD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P50 AG005131 (PI James Brewer, MD, PhD), P50 AG023501 (PI Bruce Miller, MD), P30

AG035982 (PI Russell Swerdlow, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG049638 (PI Suzanne Craft, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P50 AG033514 (PI Sanjay Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD). The project described was supported by Award Number P30AG031054 from the NIH. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

Code availability Code can be obtained from the corresponding author upon request.

Declarations

Ethics approval All ADCs that contribute data to NACC are approved by their local Institutional Review Board.

The study also was approved by the institutional review board of the University of Alabama and was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Consent to participate Participants provided informed consent at the ADC where they completed their study visits.

Consent for publication Not applicable.

Conflicts of interest/Competing interests All authors declare no financial interest or benefit arising from the direct applications of their research.

References

- Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, S. L. (1994). Socioeconomic status and health: The challenge of the gradient. *American Psychologist*, 49(1), 15.
- Andersen, K., Launer, L. J., Dewey, M. E., Letenneur, L., Ott, A., Copeland, J. R. M., ... Hofman, A. (1999). Gender differences in the incidence of AD and vascular dementia: The EURODEM Studies. *Neurology*, 53(9), 1992–1997.
- Ardekani, B. A., Convit, A., & Bachman, A. H. (2016). Analysis of the MIRIAD data shows sex differences in hippocampal atrophy progression. *Journal of Alzheimer's Disease*, 50(3), 847–857. <https://doi.org/10.3233/JAD-150780>
- Asperholm, M., Nagar, S., Dekhtyar, S., & Herlitz, A. (2019). The magnitude of sex differences in verbal episodic memory increases with social progress: Data from 54 countries across 40 years. *PLoS one*, 14(4), e0214945.
- Asperholm, M., Högman, N., Rafi, J., & Herlitz, A. (2019b). What did you do yesterday? A meta-analysis of sex differences in episodic memory. *Psychological Bulletin*, 145(8), 785.
- Avila, J. F., Vonk, J. M. J., Verney, S. P., Witkiewitz, K., ArceRentería, M., Schupf, N., Mayeux, R., & Manly, J. J. (2019). Sex/gender differences in cognitive trajectories vary as a function of race/ethnicity. *Alzheimer's and Dementia*, 15(12), 1516–1523. <https://doi.org/10.1016/j.jalz.2019.04.006>

- Babulal, G. M., Quiroz, Y. T., Albeni, B. C., Arenaza-Urquijo, E., Astell, A. J., Babiloni, C., ... & O'Bryant, S. E. (2019). Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: update and areas of immediate need. *Alzheimer's & Dementia*, *15*(2), 292–312.
- Baldo, J. V., Arévalo, A., Patterson, J. P., & Dronkers, N. F. (2013). Grey and white matter correlates of picture naming: Evidence from a voxel-based lesion analysis of the Boston Naming Test. *Cortex*, *49*(3), 658–667.
- 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*, *18*(5), 856–865. <https://doi.org/10.1017/s1355617712000628>
- 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. <https://doi.org/10.1017/S1355617715001113>
- Barnett, K., Mercer, S. W., Norbury, M., Watt, G., Wyke, S., & Guthrie, B. (2012). Epidemiology of multimorbidity and implications for health care, research, and medical education: A cross-sectional study. *The Lancet*, *380*(9836), 37–43. [https://doi.org/10.1016/S0140-6736\(12\)60240-2](https://doi.org/10.1016/S0140-6736(12)60240-2)
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., ... & Cesarini, D. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*(1), 6.
- Benton, A. L. (1983). *Contributions to neuropsychological assessment: A clinical manual*. Oxford University Press.
- Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, *45*(2), 459–464.
- Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *Clinical Neuropsychology*, *3*, 129–136.
- Boraxbekk, C. J., Lundquist, A., Nordin, A., Nyberg, L., Nilsson, L. G., & Adolfsson, R. (2015). Free recall episodic memory performance predicts dementia ten years prior to clinical diagnosis: Findings from the Betula Longitudinal Study. *Dementia and Geriatric Cognitive Disorders*, *5*(2), 191–202. <https://doi.org/10.1159/000381535>
- Braveman, P. A., Cubbin, C., Egerter, S., Williams, D. R., & Pamuk, E. (2010). Socioeconomic disparities in health in the United States: What the patterns tell us. *American Journal of Public Health*, *100*(S1), S186–S196. <https://doi.org/10.2105/AJPH.2009.166082>
- Buckley, R. F., Mormino, E. C., Amariglio, R. E., Properzi, M. J., Rabin, J. S., Lim, Y. Y., ... & Doré, V. (2018). Sex, amyloid, and APOE ϵ 4 and risk of cognitive decline in preclinical Alzheimer's disease: findings from three well-characterized cohorts. *Alzheimer's & Dementia*, *14*(9), 1193–1203.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Houghton Mifflin.
- Cavedo, E., Chiesa, P. A., Houot, M., Ferretti, M. T., Grothe, M. J., Teipel, S. J., Lista, S., Habert, M. O., Potier, M. C., Dubois, B., Hampel, H., Bakardjian, H., Benali, H., Bertin, H., Bonheur, J., Boukadida, L., Boukerrou, N., Chiesa, P., Colliot, O., ... Younesi, E. (2018). Sex differences in functional and molecular neuroimaging biomarkers of Alzheimer's disease in cognitively normal older adults with subjective memory complaints. In *Alzheimer's and Dementia* (Vol. 14, Issue 9, pp. 1204–1215). Elsevier Inc. <https://doi.org/10.1016/j.jalz.2018.05.014>
- Chhetri, J. K., de Souto Barreto, P., Cantet, C., Cesari, M., Coley, N., Andrieu, S., & Vellas, B. (2018). Trajectory of the MAPT-PACC-preclinical Alzheimer cognitive composite in the Placebo Group of a Randomized Control Trial: Results from the MAPT Study: Lessons for Further Trials. *The Journal of Prevention of Alzheimer's Disease*, *5*(1), 31–35.
- Cottrell, J. M., Newman, D. A., & Roisman, D. I. (2015). Explaining the black-white gap in cognitive test scores: Toward a theory of adverse impact. *Journal of Applied Psychology*, *100*(6), 1713–1736. <https://doi.org/10.1037/apl0000020>
- 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. <https://doi.org/10.1093/geron/gls122>
- Cubbin, C., & Winkleby, M. A. (2005). Protective and harmful effects of neighborhood-level deprivation on individual-level health knowledge, behavior changes, and risk of coronary heart disease. *American Journal of Epidemiology*, *162*(6), 559–568. <https://doi.org/10.1093/aje/kwi250>
- Deary, I. J., Pattie, A., & Starr, J. M. (2013). The stability of intelligence from age 11 to age 90 years: The Lothian birth cohort of 1921. *Psychological Science*, *24*(12), 2361–2368. <https://doi.org/10.1177/0956797613486487>
- Deckers, K., van Boxtel, M. P., Schiepers, O. J., de Vugt, M., Muñoz Sánchez, J. L., Anstey, K. J., ... Köhler, S. (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. <https://doi.org/10.1002/gps.4245>
- Dierckx, E., Engelborghs, S., De Raedt, R., Van Buggenhout, M., De Deyn, P. P., Verleye, G., ... Ponjaert-Kristoffersen, I. (2008). Differentiation between dementia and depression among older persons: can the difference between actual and premorbid intelligence be useful?. *Journal of geriatric psychiatry and neurology*, *21*(4), 242–249. <https://doi.org/10.1177/0891988708324938>
- Donohue, M. C., Sperling, R. A., Salmon, D. P., Rentz, D. M., Raman, R., Thomas, R. G., ... & Aisen, P. S. (2014). The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. *JAMA neurology*, *71*(8), 961–970.
- Duarte-Guterman, P., Albert, A. Y., Barha, C. K., Galea, L. A. M., on behalf of the Alzheimer's Disease. (2021). Sex influences the effects of APOE genotype and Alzheimer's diagnosis on neuropathology and memory. *Psychoneuroendocrinology*, *129*, 1–10. <https://doi.org/10.1016/j.psyneuen.2021.105248>
- Ekstrom, R., French, J., Harman, H., & Dermen, D. (1976). *Manual for kit of factor referenced cognitive tests*. Educational Testing Service.
- Frazen, M. D., Burgess, E. J., & Smith-Seemiller, L. (1997). Methods of estimating premorbid functioning. *Archives of Clinical Neuropsychology*, *12*(8), 711–738.
- 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.
- Garcia, M. A., Downer, B., Chiu, C. T., Saenz, J. L., Ortiz, K., & Wong, R. (2021). Educational benefits and cognitive health life expectancies: Racial/ethnic, nativity, and gender disparities. *The Gerontologist*, *61*(3), 330–340. <https://doi.org/10.1093/geront/gnaa112>
- Geronimus, A. T., Hicken, M., Keene, D., & Bound, J. (2006). "Weathering" and age patterns of allostatic load scores among blacks and whites in the United States. *American Journal of Public Health*, *96*(5), 826–833. <https://doi.org/10.2105/AJPH.2004.060749>
- Geronimus, A. T., Pearson, J. A., Linnenbringer, E., Schulz, A. J., Reyes, A. G., Epel, E. S., Lin, J., & Blackburn, E. H. (2015).

- Race-ethnicity, poverty, urban stressors, and telomere length in a Detroit community-based sample. *Journal of Health and Social Behavior*, 56(2), 199–224. <https://doi.org/10.1177/0022146515582100>
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74–78.
- Glisky, E. L., Polster, M. R., & Routhieaux, B. C. (1995). Double dissociation between item and source memory. *Neuropsychology*, 9(2), 229–235. <https://doi.org/10.1037/0894-4105.9.2.229>
- Goldstein, F. C., Ashley, A. V., Miller, E., Alexeeva, O., Zanders, L., & King, V. (2014). Validity of the Montreal cognitive assessment as a screen for mild cognitive impairment and dementia in African Americans. *Journal of Geriatric Psychiatry and Neurology*, 27(3), 199–203. <https://doi.org/10.1177/0891988714524630>
- Hamel, R., Köhler, S., Sistermans, N., Koene, T., Pijnenburg, Y., van der Flier, W., ... & Ramakers, I. (2015). The trajectory of cognitive decline in the pre-dementia phase in memory clinic visitors: findings from the 4C-MCI study. *Psychological Medicine*, 45(7), 1509–1519.
- Hedden, T., Oh, H., Younger, A. P., & Patel, T. A. (2013). Meta-analysis of amyloid-cognition relations in cognitively normal older adults. *Neurology*, 80(14), 1341–1348.
- Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, 50, 1166–1186. <https://doi.org/10.3758/s13428-017-0935>
- Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist*, 47(9), 1083.
- Hohman, T. J., Dumitrescu, L., Barnes, L. L., Thambisetty, M., Beecham, G., Kunkle, B., Gifford, K. A., Bush, W. S., Chibnik, L. B., Mukherjee, S., De Jager, P. L., Kukull, W., Crane, P. K., Resnick, S. M., Keene, C. D., Montine, T. J., Schellenberg, G. D., Haines, J. L., Zetterberg, H., ... Zhao, Y. (2018). Sex-specific association of apolipoprotein e with cerebrospinal fluid levels of tau. *JAMA Neurology*, 75(8), 989–998. <https://doi.org/10.1001/jamaneurol.2018.0821>
- Howieson, D. B., Carlson, N. E., Moore, M. M., Wasserman, D., Abendroth, C. D., Payne-Murphy, J., & Kaye, J. A. (2008). Trajectory of mild cognitive impairment onset. *Journal of the International Neuropsychological Society*, 14(2), 192–198.
- Jack Jr, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., ... & Liu, E. (2018). NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14(4), 535–562.
- Jacobs, D. M., Sano, M., Dooneief, G., Marder, K., Bell, K. L., & Stern, Y. (1995). Neuropsychological detection and characterization of preclinical Alzheimer's disease. *Neurology*, 45(5), 957–962.
- Johnson, W., Bouchard, T. J., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: Consistent results from three test batteries. *Intelligence*, 32, 95–107. [https://doi.org/10.1016/S0160-2896\(03\)00062-X](https://doi.org/10.1016/S0160-2896(03)00062-X)
- Jones, R. N. (2003). Racial bias in the assessment of cognitive functioning of older adults. *Aging & Mental Health*, 7(2), 83–102.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test: Assessment of aphasia and related disorders* (p. 2). Lea & Febiger.
- Karlamañgla, A. S., Miller-Martinez, D., Lachman, M. E., Tun, P. A., Koretz, B. K., & Seeman, T. E. (2014). Biological correlates of adult cognition: Midlife in the United States (MIDUS). *Neurobiology of Aging*, 35(2), 387–394.
- Kaufman, A. S., & Horn, J. L. (1996). Age changes on tests of fluid and crystallized ability for women and men on the Kaufman Adolescent and Adult Intelligence Test (KAIT) at ages 17–94 years. *Archives of Clinical Neuropsychology*, 11(2), 97–121.
- Kennedy, B. P., Kawachi, I., Glass, R., & Prothrow-Stith, D. (1998). Income distribution, socioeconomic status, and self-rated health in the United States: Multilevel analysis. *BMJ*, 317(7163), 917–921.
- Kiselica, A. M., Johnson, E., Lewis, K. R., & Trout, K. (2021). Examining racial disparities in the diagnosis of mild cognitive impairment. *Applied Neuropsychology: Adult*, 1–8. <https://doi.org/10.1080/23279095.2021.1976778>
- Larson, N. I., Story, M. T., & Nelson, M. C. (2009). Neighborhood environments: Disparities in access to healthy foods in the US. *American Journal of Preventive Medicine*, 36(1), 74–81. <https://doi.org/10.1016/j.amepre.2008.09.025>
- Letang, S. K., Lin, S. S. H., Parmelee, P. A., & McDonough, I. M. (2021). Ethnoracial disparities in cognition are associated with multiple socioeconomic status-stress pathways. *Cognitive Research: Principles and Implications*, 6(1), 1–17.
- Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). Oxford University Press.
- Lim, Y. Y., Snyder, P. J., Pietrzak, R. H., Ukiqi, A., Villemagne, V. L., Ames, D., ... & Rowe, C. C. (2016). Sensitivity of composite scores to amyloid burden in preclinical Alzheimer's disease: Introducing the Z-scores of Attention, Verbal fluency, and Episodic memory for Nondemented older adults composite score. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 2, 19–26.
- Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior*, 35, 80–94.
- 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.
- Matarazzo, J. D., & Herman, D. O. (1985). Clinical uses of the WAIS-R: Base rates of differences between VIQ and PIQ in the WAIS-R standardization sample. *Handbook of intelligence: Theories, measurements, and applications*, 899–932.
- Matthews, K. A., Xu, W., Gaglioti, A. H., Holt, J. B., Croft, J. B., Mack, D., & McGuire, L. C. (2018). Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged ≥ 65 years. *Alzheimer's & Dementia*. <https://doi.org/10.1016/j.jalz.2018.06.3063>
- McCarthy, F., Burns, W. J., & Sellers, A. H. (2005). Discrepancies between premorbid and current IQ as a function of progressive mental deterioration. *Perceptual and Motor Skills*, 100(1), 69–76. <https://doi.org/10.2466/pms.100.1.69-76>
- McDonough, I. M., & Popp, T. E. (2020). Linear and nonlinear relationships between cognitive subdomains of ability discrepancy and Alzheimer's disease biomarkers. *Neuropsychology*, 34(2), 211–226. <https://doi.org/10.1037/neu0000606>
- McDonough, I. M., Bischof, G. N., Kennedy, K. M., Rodrigue, K. M., Farrell, M. E., & Park, D. C. (2016). Discrepancies between fluid and crystallized ability in healthy adults: A behavioral marker of preclinical Alzheimer's disease. *Neurobiology of Aging*, 46, 68–75. <https://doi.org/10.1016/j.neurobiolaging.2016.06.011>
- McDonough, I. M., Harrell, E. R., Black, S. R., Allen, R. S., & Parmelee, P. A. (2021). Sources of nonreplicability in aging ethnoracial health disparities research. *Psychology and Aging*, 37(1), 60–71. <https://doi.org/10.1037/pag0000606>
- Mistridis, P., Krumm, S., Monsch, A. U., Berres, M., & Taylor, K. I. (2015). The 12 years preceding mild cognitive impairment due to Alzheimer's disease: The temporal emergence of cognitive decline. *Journal of Alzheimer's Disease*, 48(4), 1095–1107.
- Mormino, E. C., Papp, K. V., Rentz, D. M., Donohue, M. C., Amariglio, R., Quiroz, Y. T., ... & Gatchel, J. R. (2017). Early

- and late change on the preclinical Alzheimer's cognitive composite in clinically normal older individuals with elevated amyloid β . *Alzheimer's & Dementia*, 13(9), 1004–1012.
- Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, 43, 2412–2414.
- Mortamais, M., Ash, J. A., Harrison, J., Kaye, J., Kramer, J., Randolph, C., ... Ritchie, K. (2017). Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility. *Alzheimer's and Dementia*, 13(4), 468–492. <https://doi.org/10.1016/j.jalz.2016.06.2365>
- Nogueira, M. L., Lafitte, O., Steyaert, J. M., Bakardjian, H., Dubois, B., Hampel, H., & Schwartz, L. (2016). Mechanical stress related to brain atrophy in Alzheimer's disease. *Alzheimer's & Dementia*, 12, 11–20. <https://doi.org/10.1016/j.jalz.2015.03.005>
- O'Carroll, R. E., & Gilleard, C. J. (1986). Estimation of premorbid intelligence in dementia. *British Journal of Clinical Psychology*, 25(2), 157–158.
- Ozier, E. M., Taylor, V. J., & Murphy, M. C. (2019). The cognitive effects of experiencing and observing subtle racial discrimination. *Journal of Social Issues*, 75(4), 1087–1115. <https://doi-org.libdata.lib.ua.edu/10.1111/josi.12349>
- Papp, K. V., Rentz, D. M., Orlovsky, I., Sperling, R. A., & Mormino, E. C. (2017). Optimizing the preclinical Alzheimer's cognitive composite with semantic processing: The PACC5. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 3(4), 668–677.
- Papp, K. V., Buckley, R., Mormino, E., Maruff, P., Villemagne, V. L., Masters, C. L., ... & Amariglio, R. E. (2019). Clinical meaningfulness of subtle cognitive decline on longitudinal testing in preclinical AD. *Alzheimer's & Dementia*.
- Perkins, P., Annegers, J. F., Doody, R. S., Cooke, N., Aday, L., & Vernon, S. W. (1997). Incidence and prevalence of dementia in a multiethnic cohort of municipal retirees. *Neurology*, 49, 44–50.
- Pohl, D. J., Seblova, D., Avila, J. F., Dorsman, K. A., Kulick, E. R., Casey, J. A., & Manly, J. (2021). Relationship between residential segregation, later-life cognition, and incident dementia across race/ethnicity. *International Journal of Environmental Research and Public Health*, 18(21), 11233. <https://doi.org/10.3390/ijerph182111233>
- Powell, L. M., Chaloupka, F. J., & Bao, Y. (2007). The availability of fast-food and full-service restaurants in the United States. *American Journal of Preventive Medicine*, 33(4), 240–245. <https://doi.org/10.1016/j.amepre.2007.07.005>
- Price, J. L., McKeel Jr, D. W., Buckles, V. D., Roe, C. M., Xiong, C., Grundman, M., ... & Morris, J. C. (2009). Neuropathology of nondemented aging: presumptive evidence for preclinical Alzheimer disease. *Neurobiology of Aging*, 30(7), 1026–1036.
- Reitan, R. (1979). *Manual for administration of neuropsychological test batteries for adults and children*. Reitan Neuropsychology Laboratories.
- Rentz, D. M., & Weintraub, S. (2000). Neuropsychological detection of early probable Alzheimer's disease. In *Early diagnosis of Alzheimer's disease* (pp. 169–189). Humana Press.
- Roe, C. M., Xiong, C., Miller, J. P., & Morris, J. C. (2007). Education and Alzheimer disease without dementia support for the cognitive reserve hypothesis. *Neurology*, 68(3), 223–228. <https://doi.org/10.1212/01.wnl.0000251303.50459.8a>
- Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement*, 20(4), 335–343. <https://doi.org/10.1111/j.1745-3984.1983.tb00211.x>
- Rowe, J. W., & Kahn, R. L. (1987). Human aging: Usual and successful. *Science*, 237(4811), 143–149.
- Salthouse, T. A., & Becker, J. T. (1998). Independent effects of Alzheimer's disease on neuropsychological functioning. *Neuropsychology*, 12(2), 242–252.
- Scheltens, N. M., Tijms, B. M., Heymans, M. W., Rabinovici, G. D., Cohn-Sheehy, B. I., Miller, B. L., ... Van de Flier, W. M. (2018). Prominent non-memory deficits in Alzheimer's disease are associated with faster disease progression. *Journal of Alzheimer's Disease*, 65(3), 1029–1039. <https://doi.org/10.3233/JAD-171088>.
- Schretlen, D., Benedict, R. H., & Bobholz, J. H. (1994). Composite reliability and standard errors of measurement for a seven-subtest short form of the Wechsler Adult Intelligence Scale—Revised. *Psychological Assessment*, 6(3), 188.
- Schweizer, T. A., Ware, J., Fischer, C. E., Craik, F. I., & Bialystok, E. (2012). Bilingualism as a contributor to cognitive reserve: Evidence from brain atrophy in Alzheimer's disease. *Cortex*, 48(8), 991–996.
- Seidenberg, M., Geary, E., & Hermann, B. (2005). Investigating temporal lobe contribution to confrontation naming using MRI quantitative volumetrics. *Journal of the International Neuropsychological Society: JINS*, 11(4), 358.
- Sperling, R., Mormino, E., & Johnson, K. (2014). The evolution of preclinical Alzheimer's disease: Implications for prevention trials. *Neuron*, 84(3), 608–622. <https://doi.org/10.1016/j.neuron.2014.10.038>
- Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., ... & Park, D. C. (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 280–292. <https://doi.org/10.1016/j.jalz.2011.03.003>
- Steenland, K., Goldstein, F. C., Levey, A., & Wharton, W. (2016). A meta-analysis of Alzheimer's disease incidence and prevalence comparing African-Americans and caucasians. *Journal of Alzheimer's Disease*, 50(1), 71–76. <https://doi.org/10.3233/JAD-150778>
- Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology*, 11(11), 1006–1012. [https://doi.org/10.1016/S1474-4422\(12\)70191-6](https://doi.org/10.1016/S1474-4422(12)70191-6)
- Sutin, A. R., Stephan, Y., Carretta, H., & Terracciano, A. (2015). Perceived discrimination and physical, cognitive, and emotional health in older adulthood. *The American Journal of Geriatric Psychiatry*, 23(2), 171–179. <https://doi.org/10.1016/j.jagp.2014.03.007>
- Tang, M. X., Cross, P., Andrews, H., Jacobs, D. M., Small, S., Bell, K., ... Mayeux, R. (2001). Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. *Neurology*, 56, 49–56.
- Thomas, K. R., Edmonds, E. C., Eppig, J., Salmon, D. P., Bondi, M. W., Alzheimer's Disease Neuroimaging Initiative. (2018). Using neuropsychological process scores to identify subtle cognitive decline and predict progression to mild cognitive impairment. *Journal of Alzheimer's Disease*, 64(1), 195–204.
- Vest, R. S., & Pike, C. J. (2013). Gender, sex steroid hormones, and Alzheimer's disease. *Hormones and Behavior*, 63(2), 301–307.
- Vogel, J. W., Young, A. L., Oxtoby, N. P., Smith, R., Ossenkopppele, R., Strandberg, O. T., ... & Hansson, O. (2021). Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nature Medicine*, 1–11.
- Wechsler, D. (1944). *The Measurement of Adult Intelligence*. Williams and Wilkins.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (WAIS-iii)*. Psychological Corporation.
- Weintraub, S., Carrillo, M. C., Farias, S. T., Goldberg, T. E., Hendrix, J. A., Jaeger, J., ... Randolph, C. (2018). Measuring cognition and function in the preclinical stage of Alzheimer's disease. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 4, 64–75. <https://doi.org/10.1016/j.trci.2018.01.003>

- Weissberger, G. H., Gollan, T. H., Bondi, M. W., Nation, D. A., Hansen, L. A., Galasko, D., & Salmon, D. P. (2019). Neuropsychological deficit profiles, vascular risk factors, and neuropathological findings in Hispanic older adults with autopsy-confirmed Alzheimer's disease. *Journal of Alzheimer's Disease*, *67*(1), 291–302.
- Weuve, J., Barnes, L. L., de Leon, C. F. M., Rajan, K. B., Beck, T., Aggarwal, N. T., ... & Evans, D. A. (2018). Cognitive aging in black and white Americans: cognition, cognitive decline, and incidence of Alzheimer disease dementia. *Epidemiology (Cambridge, Mass.)*, *29*(1), 151.
- Zachary, R. A., & Shipley, W. C. (1986). *Shipley Institute of Living Scale: Revised Manual*. Western Psychological Services.
- Zahodne, L. B., Manly, J. J., Smith, J., Seeman, T., & Lachman, M. E. (2017). Socioeconomic, health, and psychosocial mediators of racial disparities in cognition in early, middle, and late adulthood. *Psychology and Aging*, *32*(2), 118. <https://doi.org/10.1037/pag0000154>
- Zahodne, L. B., Sol, K., & Kraal, Z. (2019). Psychosocial pathways to racial/ethnic inequalities in late-life memory trajectories. *The Journals of Gerontology: Series B*, *74*(3), 409–418. <https://doi.org/10.1093/geronb/gbx113>
- Zhang, B., Lin, L., Wu, S., & Al-Masqari, Z. H. (2021). Multiple Subtypes of Alzheimer's Disease Base on Brain Atrophy Pattern. *Brain Sciences*, *11*(2), 278.
- Zimmerman, M. E., Ezzati, A., Katz, M. J., Lipton, M. L., Brickman, A. M., Sliwinski, M. J., & Lipton, R. B. (2016). Perceived stress is differentially related to hippocampal subfield volumes among older adults. *PLoS ONE*, *11*(5), e0154530. <https://doi.org/10.1371/journal.pone.0154530>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open practices statement The data sets generated and analyzed during the current study are available through the publicly available National Alzheimer's Coordinating Center UDS database. The analyses were not pre-registered.