



## Brief Communication

# Construct validation of NIH Toolbox Cognition Battery premorbid cognitive functioning scores in Black and White older Americans with and without mild cognitive impairment

Colt M. Halter<sup>1</sup> , Allison C. Moll<sup>1</sup>, Katherine Kero<sup>2</sup>, Voyko Kavcic<sup>2</sup>, John L. Woodard<sup>1</sup>  and Bruno Giordani<sup>3</sup>

<sup>1</sup>Department of Psychology, Wayne State University, Detroit, MI, USA, <sup>2</sup>Institute of Gerontology, Wayne State University, Detroit, MI, USA and <sup>3</sup>Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA

### Abstract

**Objective:** Valid estimates of premorbid cognitive functioning (PMIQ) are crucial for the assessment of older adults at risk for Alzheimer's disease. We investigated the relationship between the NIH Toolbox-Cognition Battery's (NIHTB-CB) Oral Reading Recognition (ORR) subtest and Wechsler Test of Adult Reading scores (WTAR, convergent validity). We also compared ORR to NIHTB-CB Flanker scores, where null relationships were expected (discriminant validity). **Methods:** The WTAR and NIHTB-CB were administered to 130 cognitively normal (CN) and 113 participants with mild cognitive impairment (MCI). Participants were community-dwelling, older Black and White adults, ages 55–88 years. Data analysis used uncorrected standard scores and Bayesian bivariate correlations. Supplemental materials include intraclass correlations. **Results:** ORR and WTAR scores were strongly positively associated, while ORR and Flanker scores were unrelated. This pattern held when restricting analyses to the two cognitive status groups, the two racial groups, and the four race-by-diagnosis subgroups. **Conclusion:** The findings demonstrate convergent and discriminant validity and support NIHTB-CB ORR scores as valid estimates of scores on a PMIQ measure in older Black and White adults with and without MCI.

**Keywords:** Convergent; discriminant; validity; computerized; MCI; assessment

(Received 19 February 2023; final revision 16 June 2023; accepted 21 June 2023; First Published online 21 July 2023)

### Introduction

Alzheimer's disease (AD) is a rising global concern, with approximately 50 million people living with dementia (Nichols et al., 2022). In the United States, AD presents additional social problems as older Black adults are at twice the risk for being diagnosed compared to older White adults (Weuve et al., 2018). Given its degenerative course, significant focus has been placed on the early detection of dementia and AD. Evidence of cognitive decline compared to prior functioning is one of the essential criteria for diagnosing dementia (Hugo & Ganguli, 2014). In the absence of baseline cognitive testing, which is most often the case, estimates of premorbid cognitive functioning (PMIQ) help to determine the presence and magnitude of decline.

Previous work has indicated that reading ability is a better estimate of cognitive reserve than total years of education within racially diverse populations, as reading ability may reflect the quality of education received (Manly et al., 2002). Reading tests are also relatively resistant to the effects of brain injury and cognitive dysfunction, and as a result, they are often used to estimate PMIQ. A recent review of PMIQ assessments in those with mild and moderate

dementia highlighted the need for more research demonstrating the psychometric properties of such assessments (Overman et al., 2021). This brief report addresses these literature gaps by providing psychometric evidence for a relatively new measure of PMIQ, Oral Reading Recognition (ORR) scores from the NIH Toolbox Cognition Battery (NIHTB-CB), in a diverse sample.

The NIHTB-CB is a recently developed computerized measure, now adapted to be administered via iPad, that assesses fluid and crystallized cognition (Weintraub et al., 2014). While iPad administration facilitates the economic accessibility of the NIHTB-CB for use with minority and underserved populations, the validity of test scores in this modality must still be established. Some evidence suggests the battery has clinical utility in assessing older adults (Hackett et al., 2018; Scott et al., 2019). Other research in healthy young and older adults has broadly reported adequate construct validity for the crystallized composite scores but poor to acceptable validity for fluid cognition scores (Buckley et al., 2017; Loring et al., 2019; Ott et al., 2022). While measures of crystallized cognition are expected to remain relatively stable over time in those with MCI and measures of fluid cognition are expected to decline over time, Kairys and colleagues (2022) found a different pattern

**Corresponding author:** Colt M. Halter; Email: coltmhalter@wayne.edu

**Cite this article:** Halter C.M., Moll A.C., Kero K., Kavcic V., Woodard J.L., & Giordani B. (2024) Construct validation of NIH Toolbox Cognition Battery premorbid cognitive functioning scores in Black and White older Americans with and without mild cognitive impairment. *Journal of the International Neuropsychological Society*, 30: 194–198, <https://doi.org/10.1017/S1355617723000425>

Copyright © INS. Published by Cambridge University Press 2023. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is unaltered and is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use or in order to create a derivative work.

using the NIHTB-CB. In a sample of older Black adults, these researchers found that crystallized scores were more predictive of MCI than fluid scores (Kairys et al., 2022). Taken together, the extant literature generally supports the NIHTB-CB as an adequate measure of crystallized reasoning with relatively poorer validity as a measure of fluid reasoning. Notwithstanding, the literature suggests some further need to investigate the validity and clinical utility of NIHTB-CB scores in racial minorities.

Performance on the ORR test from the NIHTB-CB may estimate PMIQ. However, no extant research demonstrates the validity of the NIHTB-CB ORR scores when administered on the iPad as measures of PMIQ in a large sample of Black and White participants with and without MCI. Validating performance on this instrument as a measure of PMIQ is essential because it may help to characterize baseline cognitive functioning. Additionally, valid tests of PMIQ that are easy to administer and economically accessible may help to alleviate racial disparities in AD as they may become part of more routine and affordable care.

Therefore, in a sample of older Black and White adults with and without MCI, we investigated whether scores on the NIHTB-CB ORR test were related to performance on the Wechsler Test of Adult Reading (WTAR), an established word recognition test that estimates PMIQ (Green et al., 2008). Convergent validity is the extent to which scores from a measure positively correlate with scores from a similar measure, whereas discriminant validity is the extent to which scores from a given measure do not correlate with scores from a dissimilar measure (Campbell & Fiske, 1959). As evidence for convergent validity, we hypothesized positive correlations between scores on the NIHTB-CB ORR scores and WTAR performance. We expected a null relationship between NIHTB-CB ORR and Flanker performance as a confirmation of discriminant validity.

## Methods

A total of 243 community-dwelling older adults were recruited across two sites: the Healthier Black Elders Center (Mitchell et al., 2020) at the Wayne State University Institute of Gerontology and the Michigan Alzheimer's Disease Research Center (MADRC). Participants were co-enrolled in allied studies, and all recruitment and baseline measurements occurred between March 2017 and March 2022. The recruitment period was the same for both sites. Inclusion and exclusion criteria followed the MCI criteria established by the 2011 National Institute of Health and Alzheimer's Association (Albert et al., 2011). Participants were excluded if they suffered from significant psychiatric, medical, or neurological deficits other than MCI that may have impaired cognitive ability. Participants were also excluded if they had physical limitations that precluded neuropsychological assessment. A consensus conference determined MCI diagnoses following NACC guidelines outlined previously (Rahman-Filipiak et al., 2018). The MCI diagnostic group included all-cause MCI. All participants provided informed consent, and the University of Michigan Medical School and Wayne State University IRBs approved data collection. All procedures complied with the Declaration of Helsinki.

Neuropsychological assessments were completed at the MADRC. All participants received the National Alzheimer's Coordinating Center (NACC) Uniform Data Set, Version 3 (UDS-3) assessments and locally added measures (e.g., the WTAR and the NIHTB-CB). Only baseline measurements were included in the current study. All testing occurred on the same day, and the WTAR preceded the NIHTB-CB.

The WTAR is a standardized reading test and may serve as a measure of PMIQ, as demonstrated by Green et al. (2008). A page of words with irregular pronunciations is presented, and the participant is asked to read each word aloud. The word difficulty increases throughout the list, and the examiner determines the number of correctly read items (Green et al., 2008).

Similarly, the NIHTB-CB ORR test is a standardized reading assessment designed to assess crystallized cognition, administered via iPad. Single, irregularly spelled words appear in the center of the iPad screen and are read aloud by the participant. Items are presented sequentially using computer adaptive testing (i.e., item difficulty is automatically adjusted according to participant performance). The NIHTB-CB Flanker test is a standardized assessment of inhibitory control and attention. It consists of trials in which a central arrow pointing left or right is presented on the iPad, flanked by arrows in a congruent or incongruent direction. The participant is asked to indicate the direction of the central arrow for each trial, and scores are determined by accuracy and response time (Weintraub et al., 2014).

Bayesian ANOVAs and chi-square tests of independence were used for demographic analyses. To evaluate the influence of demographic predictors on both NIHTB-CB and WTAR equally in our sample, we used uncorrected standard scores for all neuropsychological measures. Bayesian bivariate correlations were used to demonstrate convergent and discriminant validity. Intraclass correlation coefficients (ICC) were included as supplemental analyses. Bayesian analyses yield Bayes Factors ( $BF_{10}$ ) that estimate the relative support for the alternative and null hypotheses. As  $BF_{10}$  values exceed 3, they provide increasing evidence for rejecting the null hypothesis. When  $BF_{10}$  values fall below 1/3 and approach 0, they suggest evidence for accepting the null hypothesis.  $BF_{10}$  values between 1/3 and 3 indicate indeterminate support for either hypothesis (Kruschke, 2021). Analyses were done in JASP version 0.16.3 (JASP Team, 2022). Default priors were used: ANOVAs used the uniform distribution as the prior; the chi-square prior concentration was set to 1; correlations used a stretched beta prior width of 1.

## Results

### Demographics

The sample's sex and racial compositions were approximately 75% female and 63% Black. Age ranged from 55 to 88 years, with an average age of 71.4 ( $SD = 7.3$ ). Years of education ranged from eight to 20 years, with a mean of 15.7 years ( $SD = 2.5$ ). Participants comprised two diagnostic groups: 130 were CN, and 113 were diagnosed with MCI. This composition formed four race-by-diagnosis subgroups: 74 CN Blacks, 79 MCI Blacks, 56 CN Whites, and 34 MCI Whites.

$2 \times 2$  Bayesian ANOVAs were conducted to determine whether there were any differences in age or years of education, respectively, depending on cognitive status (CN and MCI) and race (Black and White). Results indicated that age did not differ depending on an interaction effect between race and cognitive status ( $BF_M = 0.11$ ) nor did age differ by race alone ( $BF_M = 0.05$ ). Unsurprisingly, there was a main effect indicating age differed across cognitive status ( $BF_M = 12.20$ ). A post hoc Bayesian *t*-test showed that the MCI group was 2.7 years older than the CN group, corresponding to a small effect size ( $BF_{10} = 9.03$ , Median  $\delta = 0.37$ , 95% CI [0.12, 0.62]).

Similarly, years of education did not differ depending on an interaction effect of race and cognitive status ( $BF_M = 0.58$ ).

**Table 1.** Bayesian correlation comparisons

	Whole Sample (N = 243)	CN (n = 130)	MCI (n = 113)	Black (n = 153)	White (n = 90)	Black CN (n = 74)	White CN (n = 56)	Black MCI (n = 79)	White MCI (n = 34)
<b>ORR +</b>	$r = .90^{***}$	$r = .91^{***}$	$r = .89^{***}$	$r = .86^{***}$	$r = .90^{***}$	$r = .88^{***}$	$r = .89^{***}$	$r = .82^{***}$	$r = .91^{***}$
<b>WTAR</b>	[.87, .92]	[.87, .93]	[.83, .92]	[.81, .90]	[.84, .93]	[.82, .92]	[.80, .93]	[.72, .88]	[.81, .95]
<b>ORR + Flanker</b>	$r = .11$	$r = .01$	$r = .11$	$r = .08$	$r = -.08$	$r = -.08$	$r = -.08$	$r = .13$	$r = -.15$
	[-.01, .24]	[-.16, .18]	[-.07, .29]	[-.08, .23]	[-.28, .12]	[-.30, .15]	[-.33, .18]	[-.09, .34]	[-.44, .19]
<b>WTAR + Flanker</b>	$r = .14$	$r = .08$	$r = .08$	$r = .10$	$r = -.07$	$r = .00$	$r = -.03$	$r = .07$	$r = -.16$
	[.01, .26]	[-.09, .24]	[-.11, .25]	[-.06, .25]	[-.27, .14]	[-.22, .22]	[-.28, .23]	[-.15, .28]	[-.46, .18]

\* $BF_{10} > 10$ , \*\*  $BF_{10} > 30$ , \*\*\*  $BF_{10} > 100$ .

**Table 2.** Bayesian correlations of PMIQ scores and demographics

	Whole Sample (N = 243)	CN (n = 130)	MCI (n = 113)	Black (n = 153)	White (n = 90)	Black CN (n = 74)	White CN (n = 56)	Black MCI (n = 79)	White MCI (n = 34)
<b>WTAR +</b>	$r = .47^{***}$	$r = .44^{***}$	$r = .48^{***}$	$r = .51^{***}$	$r = .42^{***}$	$r = .50^{***}$	$r = .51^{***}$	$r = .46^{***}$	$r = .34$
<b>Education</b>	[.36, .56]	[.28, .56]	[.32, .60]	[.38, .61]	[.23, .57]	[.30, .65]	[.28, .67]	[.25, .61]	[.00, .59]
<b>WTAR +</b>	$r = .10$	$r = .14$	$r = .16$	$r = .03$	$r = .27^\dagger$	$r = .07$	$r = .30$	$r = .10$	$r = .31$
<b>Age</b>	[-.03, .22]	[-.03, .30]	[-.02, .33]	[-.13, .18]	[.07, .45]	[-.16, .29]	[.04, .51]	[-.12, .31]	[-.04, .57]
<b>ORR + Education</b>	$r = .44^{***}$	$r = .41^{***}$	$r = .44^{***}$	$r = .46^{***}$	$r = .38^{***}$	$r = .48^{***}$	$r = .40^*$	$r = .38^{**}$	$r = .37$
	[.33, .53]	[.25, .54]	[.27, .57]	[.32, .57]	[.18, .54]	[.28, .63]	[.15, .59]	[.17, .55]	[.04, .61]
<b>ORR +</b>	$r = .07$	$r = .11$	$r = .13$	$r = -.03$	$r = .28^\dagger$	$r = .00$	$r = .27$	$r = .02$	$r = .36$
<b>Age</b>	[-.05, .20]	[-.07, .27]	[-.05, .31]	[-.19, .13]	[.08, .46]	[-.23, .22]	[.00, .49]	[-.20, .24]	[.02, .61]

$^\dagger BF_{10} > 3$ , \*  $BF_{10} > 10$ , \*\*  $BF_{10} > 30$ , \*\*\*  $BF_{10} > 100$ .

Evidence for differences in years of education between racial groups ( $BF_M = 1.05$ ) and cognitive status groups ( $BF_M = 1.42$ ) was inconclusive. Lastly, Bayesian chi-square tests of independence revealed no conclusive evidence that the proportion of MCI status was related to race ( $BF_{10} = 2.02$ ).

**Validity coefficients**

Bayesian bivariate correlations in the entire sample of participants indicated that ORR scores were strongly positively associated with WTAR scores ( $r = .90$ ,  $BF_{10} = 5.94 \times 10^{85}$ , 95% CI [.87, .92]). ORR scores were not conclusively associated with Flanker test scores ( $r = .11$ ,  $BF_{10} = 0.39$ , 95% CI [-.01, .24]). WTAR scores and Flanker test scores were also not conclusively correlated ( $r = .14$ ,  $BF_{10} = 0.76$ , 95% CI [.01, .26]). ICCs were included as supplemental analyses and showed this same pattern of results (see Supplemental Material). Further correlational analyses showed that years of education were moderately positively correlated with performance on the WTAR ( $r = .47$ ,  $BF_{10} = 5.38 \times 10^{11}$ , 95% CI [.36, .56]) and the ORR test ( $r = .44$ ,  $BF_{10} = 6.47 \times 10^9$ , 95% CI [.33, .53]). However, age was not associated with WTAR scores ( $r = .10$ ,  $BF_{10} = 0.25$ , 95% CI [-.03, .22]) or with ORR test scores ( $r = .07$ ,  $BF_{10} = 0.15$ , 95% CI [-.05, .20]). The null relations between age and WTAR scores and age and ORR scores remained for all subgroup analyses except when restricting the sample to White participants only, where weak relations emerged ( $r = .27$ ,  $BF_{10} = 3.69$ , 95% CI [.07, .45] and  $r = .28$ ,  $BF_{10} = 4.87$ , 95% CI [.08, .46]).

**Restricting analyses to cognitive status, racial, and race-by-diagnosis groups**

The same Bayesian bivariate correlations were used when restricting the sample to each cognitive status group (CN and MCI), the racial groups (Black or White participants only), and the four race-by-diagnosis subgroups. Results are summarized in

Table 1. These analyses were consistent with the validity coefficients found in the whole sample. The correlations between the neuropsychological assessment scores and demographic covariates are found in Table 2.

**Group differences**

Bayesian correlation coefficients estimate the correlation parameter value and the 95% most credible values for that parameter. Therefore, comparisons across the subgroups of the data are possible, as nonoverlapping 95% CIs indicate different correlation parameters.

Comparing the correlation parameter estimates for the relationship between WTAR scores and the ORR scores revealed that the 95% CIs for the whole sample overlapped with those of the two diagnostic groups and the four race-by-diagnosis subgroups. Thus, the relationship between WTAR performance and ORR scores is consistent across cognitive status, race, and race-by-diagnosis subgroups, suggesting convergent psychometric validity (see Table 1). Further, this pattern of overlapping 95% CIs also held across the correlation estimates between ORR scores and Flanker scores when restricting the sample to diagnostic and race-by-diagnosis groups. These results provide evidence of construct validity for the NIHTB-CB ORR as a measure of PMIQ in older Black and White adults with and without MCI.

**Discussion**

In support of our hypotheses, the present study indicated that in a sample of older Black and White adults who were either CN or diagnosed with MCI, scores on the NIHTB-CB ORR test were strongly positively correlated to an established neuropsychological measure of PMIQ and unrelated to Flanker test scores. These results provide evidence of psychometric convergent and discriminant validity, respectively, of the ORR scores as a measure of PMIQ in this population. The validity coefficients remained

consistent even when restricting the whole sample to individual cognitive status groups, racial groups, and race-by-diagnosis subgroups.

Age was uncorrelated with scores on the WTAR or the ORR test in this group of older participants over 55 years, except when restricting the sample to White participants only where weak correlations were found. Expectedly, years of education was moderately correlated with both WTAR scores and NIHTB-CB ORR test performance. The current findings are consistent with previous research on reading ability as a better predictor of neuropsychological functioning than years of education (Manly et al., 2005). The present findings also extend the previous research that suggested adequate construct validity for the NIHTB-CB crystallized measures by using an established language functioning measure and expanding the sample to a clinical context (Buckley et al., 2017; Ott et al., 2022).

To our knowledge, this study is the first to have validated NIHTB-CB ORR test scores as a measure of PMIQ in a racially diverse sample of CN older adults and older adults with MCI. The current findings show that NIHTB-CB ORR test scores are valid for estimating the WTAR scores of healthy aging adults and those with MCI. Importantly, these findings demonstrate that the NIHTB-CB ORR test performance functions equally well across racial and race-by-diagnosis subgroups, likely tapping into the same cognitive domain. Given Black Americans' disproportionate risk of developing AD (Weuve et al., 2018), validating accessible and practical neuropsychological assessments is an essential step forward for early AD detection. The present findings address these concerns and demonstrate that performance on the iPad NIHTB-CB may be used to estimate baseline cognitive functioning in the context of age-related cognitive decline in older, community-dwelling adults.

This study had several limitations. Sampling bias may play a role in interpreting the results, given that our sample was more female than the general population. Including another paper-and-pencil measure would have bolstered the evidence of discriminant validity. MCI participants were all-cause type, which could obfuscate differences in the relationship between ORR and WTAR scores across MCI subtypes. Nevertheless, the present study establishes robust validity coefficients (in the .80–.90 range), demonstrating the ability of NIHTB-CB ORR scores to estimate PMIQ.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/S1355617723000425>

**Acknowledgments.** The authors have no conflicts of interest to disclose. This research was partially supported by funding from NIH/NIA grant P30 AG053760 to the MADRC (PI, H. Paulson) and from NIA/NIH, R01 AG 054484 (PI, V. Kavcic).

## References

- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to 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), 270–279. <https://doi.org/10.1016/j.jalz.2011.03.008>
- Buckley, R. F., Sparks, K. P., Papp, K. V., Dekhtyar, M., Martin, C., Burnham, S., Sperling, R. A., Rentz, D. M. (2017). Computerized cognitive testing for use in clinical trials: A comparison of the NIH Toolbox and Cogstate C3

- Batteries. *The Journal of Prevention of Alzheimer's Disease*, 4(1), 3–11. <https://doi.org/10.14283/jpad.2017.1>
- Campbell, D. T., Fiske, D. W. 'Convergent and discriminant validation by the multitrait-multimethod matrix.' *Psychological Bulletin* 1959, 56(2), 81–105. <https://doi.org/10.1037/h0046016>
- Green, R. E. A., Melo, B., Christensen, B., Ngo, L.-A., Monette, G., & Bradbury, C. (2008). Measuring premorbid IQ in traumatic brain injury: An examination of the validity of the Wechsler Test of Adult Reading (WTAR). *Journal of Clinical and Experimental Neuropsychology*, 30(2), 163–172. <https://doi.org/10.1080/13803390701300524>
- Hackett, K., Krikorian, R., Giovannetti, T., Melendez-Cabrero, J., Rahman, A., Caesar, E. E., Chen, J. L., Hristov, H., Seifan, A., Mosconi, L., Isaacson, R. S. 'Utility of the NIH Toolbox for assessment of prodromal Alzheimer's disease and dementia'. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 2018, 10(1), 764–772. <https://doi.org/10.1016/j.dadm.2018.10.002>
- Hugo, J., & Ganguli, M. (2014). Dementia and cognitive impairment: Epidemiology, diagnosis, and treatment. *Clinics in Geriatric Medicine*, 30(3), 421–442. <https://doi.org/10.1016/j.cger.2014.04.001>
- JASP Team (2022). *JASP (Version 0.16.3) [Computer Software]*.
- Kairys, A., Daugherty, A., Kavcic, V., Shair, S., Persad, C., Heidebrink, J., Bhaumik, A., Giordani, B. (2022). Laptop-administered NIH Toolbox and cogstate brief battery in community-dwelling Black adults: Unexpected pattern of cognitive performance between MCI and healthy controls. *Journal of the International Neuropsychological Society*, 28(3), 239–248. <https://doi.org/10.1017/S135561772100028X>
- Kruschke, J. K. (2021). Bayesian analysis reporting guidelines. *Nature Human Behaviour*, 5(10), 1282–1291. <https://doi.org/10.1038/s41562-021-01177-7>
- Loring, D. W., Bowden, S. C., Staikova, E., Bishop, J. A., Drane, D. L., & Goldstein, F. C. (2019). NIH Toolbox picture sequence memory test for assessing clinical memory function: Diagnostic relationship to the Rey Auditory Verbal Learning Test. *Archives of Clinical Neuropsychology*, 34(2), 268–276. <https://doi.org/10.1093/arclin/acy028>
- Manly, J. J., Jacobs, D. M., Touradjji, 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*, 8(3), 341–348. <https://doi.org/10.1017/S1355617702813157>
- Manly, J. J., Schupf, N., Tang, M.-X., & Stern, Y. (2005). Cognitive decline and literacy among ethnically diverse elders. *Journal of Geriatric Psychiatry and Neurology*, 18(4), 213–217. <https://doi.org/10.1177/0891988705281868>
- Mitchell, J., Perry, T., Rorai, V., Ilardo, J., Lichtenberg, P., & Jackson, J. (2020). Building and sustaining a community advisory board of African American older adults as the foundation for volunteer research recruitment and retention in health sciences. *Ethnicity & Disease*, 30(Suppl), 755–764. <https://doi.org/10.18865/ed.30.S2.755>
- Nichols, E., Steinmetz, J. D., Vollset, S. E., Fukutaki, K., Chalek, J., Abd-Allah, F., Abdoli, A., Abualhasan, A., Abu-Gharbieh, E., Akram, T. T., Al Hamad, H., Alahdab, F., Alanezi, F. M., Alipour, V., Almustanyir, S., Amu, H., Ansari, I., Arabloo, J., & Ashraf, T. (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the Global Burden of Disease Study 2019. *The Lancet Public Health*, 7(2), e105–e125. [https://doi.org/10.1016/S2468-2667\(21\)00249-8](https://doi.org/10.1016/S2468-2667(21)00249-8)
- Ott, L. R., Schantell, M., Willett, M. P., Johnson, H. J., Eastman, J. A., Okelberry, H. J., Wilson, T. W., Taylor, B. K., May, P. E. (2022). Construct validity of the NIH toolbox cognitive domains: A comparison with conventional neuropsychological assessments. *Neuropsychology*, 36(5), 468–481. <https://doi.org/10.1037/neu0000813>
- Overman, M. J., Leeworthy, S., & Welsh, T. J. (2021). Estimating premorbid intelligence in people living with dementia: A systematic review. *International Psychogeriatrics*, 33(11), 1145–1159. <https://doi.org/10.1017/S1041610221000302>
- Rahman-Filipiak, A. M., Giordani, B., Heidebrink, J., Bhaumik, A., & Hampstead, B. M. (2018). Self- and informant-reported memory complaints: Frequency and severity in cognitively intact individuals and those with mild cognitive impairment and neurodegenerative dementias. *Journal of Alzheimer's Disease: JAD*, 65(3), 1011–1027. <https://doi.org/10.3233/JAD-180083>

- Scott, E. P., Sorrell, A., Benitez, A. 'Cognition battery in healthy older adults: Reliability, validity, and agreement with standard neuropsychological tests'. *Journal of the International Neuropsychological Society* 2019, 25(08), 857–867. <https://doi.org/10.1017/S1355617719000614>
- Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Slotkin, J., Carlozzi, N. E., Bauer, P. J., Wallner-Allen, K., Fox, N., Havlik, R., Beaumont, J. L., Mungas, D., Manly, J. J., Moy, C., Conway, K., Edwards, E., Nowinski, C. J., Gershon, R. (2014). The cognition battery of the NIH Toolbox for assessment of neurological and behavioral function: Validation in an adult sample. *Journal of the International Neuropsychological Society*, 20(6), 567–578. <https://doi.org/10.1017/S1355617714000320>
- Weuve, J., Barnes, L. L., Mendes de Leon, C. F., Rajan, K. B., Beck, T., Aggarwal, N. T., Hebert, L. E., Bennett, D. A., Wilson, R. S., Evans, D. A. 'Cognitive aging in black and white Americans: Cognition, cognitive decline, and incidence of Alzheimer disease dementia'. *Epidemiology (Cambridge, Mass.)* 2018, 29(1), 151–159. <https://doi.org/10.1097/EDE.00000000000000747>