

The Multilingual Naming Test (MINT) as a Measure of Picture Naming Ability in Alzheimer's Disease

Alena Stasenکو¹ , Diane M. Jacobs², David P. Salmon² and Tamar H. Gollan^{3,*}

¹San Diego State University/University of California, San Diego Joint Doctoral Program in Clinical Psychology, USA

²Department of Neurosciences, Shiley-Marcos Alzheimer's Disease Research Center, University of California, 9444 Medical Center Dr #1-100, La Jolla, CA 92037, USA

³Department of Psychiatry, University of California, 9500 Gilman Drive, La Jolla, San Diego, CA 92093, USA

(RECEIVED July 21, 2018; FINAL REVISION March 31, 2019; ACCEPTED April 29, 2019; FIRST PUBLISHED ONLINE June 28, 2019)

Abstract

Objective: The present study investigated the ability of the Multilingual Naming Test (MINT), a picture naming test recently added to the National Alzheimer's Coordinating Center's (NACC) Uniform Data Set neuropsychological test battery, to detect naming impairment (i.e., dysnomia) across stages of Alzheimer's disease (AD). **Method:** Data from the initial administration of the MINT were obtained on NACC participants who were cognitively normal ($N = 3,981$) or diagnosed with mild cognitive impairment ($N = 852$) or dementia ($N = 1,148$) with presumed etiology of AD. Dementia severity was rated using the Clinical Dementia Rating (CDR) scale. **Results:** Cross-sectional multiple regression analyses revealed significant effects of diagnostic group, sex, education, age, and race on naming scores. Planned comparisons collapsing across age and education groups revealed significant group differences in naming scores across levels of dementia severity. ROC curve analyses showed good diagnostic accuracy of MINT scores for distinguishing cognitively normal controls from AD dementia, but not from MCI. Within the cognitively normal group, there was a robust interaction between age and education such that naming scores exhibited the most precipitous drop across age groups for the least educated participants. Additionally, education effects were stronger in African-Americans than in Whites (a race-by-education interaction), and race effects were stronger in older than in younger age groups (a race-by-age interaction). **Conclusions:** The MINT successfully detects naming deficits at different levels of cognitive impairment in patients with MCI or AD dementia, but comparison to age, sex, race, and education-corrected norms to determine impairment is essential.

Keywords: Naming, MINT, Dysnomia, Alzheimer's disease, MCI

While the clinical hallmark of Alzheimer's disease (AD) is a deficit in episodic memory related to pathology in mesial temporal lobe structures (e.g., hippocampus, entorhinal cortex; Hyman et al., 1984), impaired semantic memory with resulting word-finding difficulty is also evident early in the disease (Barbeau et al., 2012). Patients with AD often report word-finding difficulty for names of people, places, and objects, and are impaired on formal tests of confrontation picture naming. Studies using the Boston Naming Test (BNT; Kaplan et al., 1983), for example, have shown impaired picture naming in patients with AD dementia (Williams et al., 1989) or Mild Cognitive Impairment (MCI), a condition that is often a prodromal stage of AD (Willers et al., 2008; but see Balthazar et al., 2007). Research findings generally suggest

that the picture-naming deficit in AD and MCI results from semantic network degradation that causes difficulty accessing knowledge of concepts (Garrard et al., 2005; Hodges et al., 1991, 1992; Huff et al., 1986; Lukatela et al., 1998; Mulatti et al., 2014). However, decline in attention, working memory, lexical retrieval, or the ability to ignore interference from a previous response may also contribute (Balota & Duchek, 1991; Rogers et al., 2006).

Findings regarding the impact of normal aging on naming ability are mixed—some studies suggest that naming remains relatively stable throughout the lifespan (Hashimoto et al., 2016; LaBarge et al., 1986) while others show a decline in naming performance with aging (Au et al., 1995; Ivnik et al., 1995; Mitrushina & Satz, 1995; Van Gorp et al., 1986). Studies that show a decline have attributed it to an age-related decrease in efficiency of lexical access (e.g., Barresi et al., 2000; Burke & Shafto, 2004; Kavé et al., 2010) or phonological processing (Rizio et al., 2017).

*Correspondence and reprint requests to: Tamar H. Gollan, Department of Psychiatry, University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093, USA. E-mail: tgollan@ucsd.edu

A meta-analysis that included eleven published studies of picture naming supported this possibility, concluding that there is an age-related decline in lexical retrieval that begins after age 70 (Feyereisen, 1997; see also MacKay et al., 2002; Zec et al., 2005 for similar findings). A second meta-analysis, in contrast, determined that no strong conclusions could be drawn regarding age-related decline in naming ability or lexical retrieval due to methodological limitations of most of the studies such as small sample sizes and differences in age ranges, tasks, study designs, and statistical analyses (Goulet et al., 1994; for a review of these issues see Connor et al., 2004).

The dysnomia that often characterizes early AD makes picture naming an integral component of neuropsychological assessment of aging, MCI, and AD dementia. Many picture naming tasks are designed with an ascending difficulty of items. This design enhances the efficiency of test administration, but makes it inappropriate to use a direct translation into a language different from the one in which the test was developed since the features that determine item difficulty, such as familiarity, frequency of use, and the meaning of items and their names, differ across languages and cultures. To broaden the ability to detect naming ability across various languages, the Multilingual Naming Test (MINT; Gollan et al., 2012) was developed for use in a number of languages (e.g., English, Spanish, Mandarin, Hebrew) with roughly equivalent overall difficulty of items across languages. Several studies have validated the MINT as a bilingual proficiency measure (i.e., by testing with the MINT in both languages) in cognitively normal Chinese–English children (Sheng et al., 2014), young adult Spanish–English and Chinese–English bilinguals (Gollan et al., 2012; Tomoschuk et al., 2018) and older Spanish–English bilinguals (Gollan et al., 2012). Clinical validation of the MINT was carried out in one study which showed that an abbreviated (32-item) version of the MINT, which mainly includes only the more difficult items from the full version, detected mild dysnomia in English monolingual patients with MCI or mild-to-moderate AD dementia, and provided evidence that the loss was semantic in nature (Ivanova et al., 2013). This same study also found that the MINT detected naming impairments in Spanish–English bilinguals with AD, but only when they were tested in their dominant language.

The abbreviated (32-item) MINT recently replaced the short (30-item) version of the BNT in the National Alzheimer's Disease Coordinating Centers' (NACC) Uniform Data Set (UDS) neuropsychological test battery. A cross-walk study that included both measures showed that the MINT and BNT were highly correlated ($r = 0.76$; Monsell et al., 2016). Preliminary normative data for the MINT have been recently published for the UDS (v.3) (Weintraub et al., 2018 see https://www.alz.washington.edu/WEB/npsych_means.html), but the ability of the MINT to capture impairments in language abilities across the various stages of cognitive impairment associated with AD has not been determined. In addition, it has not been determined if the MINT, like the BNT and other naming tests (Ashaie &

Obler, 2014; Welch et al., 1996; but see Connor et al., 2004), is subject to an age-by-education interaction effect in which naming scores decrease the most in the oldest-old with the lowest level of education. Thus, the goals of the present study were to (1) examine group differences on the MINT across increasing levels of cognitive impairment in a large sample of patients with MCI or dementia presumably due to AD, and cognitively healthy controls (>5,000 participants) and (2) assess the effects of sex, age, race, and education on MINT scores in normal aging, MCI, and AD dementia.

METHODS

Participants and Inclusion/Exclusion Criteria

Participants were tested on the MINT during their approximately annual UDS evaluation as part of their participation in longitudinal studies of AD at one of approximately 29 NIA-funded Alzheimer's Disease Centers (ADCs) across the United States. Our analyses included data from March 2015 through May 2017 and included only the first administration of version 3 of the UDS.

The sample was restricted to individuals clinically diagnosed as cognitively normal or with MCI or dementia due to AD based upon diagnostic research criteria (Albert et al., 2011; McKhann et al., 2011). Within each of the three clinically diagnosed groups, dementia severity was further stratified using the Clinical Dementia Rating (CDR), a semi-structured clinical interview to rate cognitive and functional performance (Morris, 1993). Participants with AD dementia were grouped into CDR scores of 0.5 (very mild), 1 (mild), or 2 (moderate). Participants with a CDR score of 3 were excluded due to insufficient sample size. Cognitively normal participants with CDR scores of >0 were excluded. MCI participants had CDR scores of either 0 ($n = 51$) or 0.5 ($n = 801$). Because the clinician's diagnosis and the CDR rating were conducted independently at some sites, CDR scores of 0.5 occurred in both the MCI and AD groups.

We included participants with concomitant medical or psychiatric conditions if those conditions were not *primary* or *contributing* causes of the observed cognitive impairment, with the exception that a few conditions judged to be *contributing* only were allowed (depression, anxiety, systemic illness, and current medications). Participants with current alcohol abuse were excluded. Only primary English speakers and participants with complete data for all variables of interest (i.e., age, sex, education, and MINT score) were included. When a participant's education level was listed as greater than 20 years, it was re-coded as a maximum of 20. Only participants age 50 and older were included in the analyses to better match age distributions across groups.

Table 1 shows the demographic characteristics and significance tests with corresponding effect sizes of the final groups used in the analyses. Controls had significantly higher education than both MCI and AD, and MCI had significantly higher education than AD. Controls were significantly younger than

Table 1. Demographic characteristics of participants and associated *p*-values and effect sizes*

	Diagnosis Group			Significance Test : <i>p</i> -value (effect size)		
	Controls (<i>n</i> = 3,981)	MCI (<i>n</i> = 852)	AD (<i>n</i> = 1,148)	Controls vs MCI	Controls vs AD	MCI vs AD
Age (Mean; <i>SD</i>)	72.5 (9.1)	75.3 (8.8)	75.0 (9.6)	<.001 (0.31)	<.001 (0.33)	0.49 (0.03)
Education (Mean; <i>SD</i>)	16.3 (2.5)	16.0 (2.7)	15.4 (2.9)	<.05 (0.12)	<.001 (0.32)	<.001 (0.21)
Sex (% Female)	65.3	48.0	51.7	<.001 (0.14)	<.001 (0.11)	0.11 (0.04)
Handedness (% Right)	89.6	88.3	87.7	0.24 (0.02)	0.07 (0.03)	0.71 (0.01)
Race (% White)	82.4	83.3	88.4	0.51 (0.01)	<.001 (0.07)	0.001 (0.07)
Race (% African-American)	15.2	13.8	8.6	0.30 (0.02)	<0.001 (0.08)	<0.001 (0.08)
Ethnicity (% Non-Hispanic)	97.3	97.3	97.1	0.98 (0.001)	0.77 (0.004)	0.81 (0.01)

**p*-values and Cohen's *d* effect sizes correspond to *t*-tests for age and education; *p*-values and Cramer's *V* effect sizes correspond to Pearson's *Chi*-square tests for categorical variables (sex, handedness, race, and ethnicity)

both MCI and AD; however, MCI and AD did not differ on age. Chi-square tests showed that controls had a higher percentage of female participants than the MCI and AD groups, and controls had a slightly lower percent of White participants than the AD group; notably the effect sizes of these differences were small (Cramer's $V \leq 0.14$). Most of the participants in the overall sample were White (84%), and there were 14% African-American or Black, 1% Asian, and 1% Other Race, which reflects the overall distribution within the national ADCs. Informed consent was obtained at the individual ADCs, as approved by individual Institutional Review Boards (IRBs), and sharing of de-identified UDS data was approved by the University of Washington's IRB (the site of the NACC data repository).

Dependent Variable—MINT Scores

The abbreviated MINT consists of 32 black-and-white line drawings of objects that are presented one at a time to be named beginning with relatively easier items and ending with relatively more difficult to name items. If an individual encounters difficulty identifying an object, a semantic or phonemic cue is provided. The number of spontaneous correct responses and correct responses following a semantic cue are summed to give a total score. Thus, the abbreviated MINT total scores range from 0 to 32 and these were used as the primary outcome measure.

Statistical Methods

Correlation and simultaneous multiple regression analyses (with model comparisons) were carried out to assess the relationship between diagnostic group, age, education, sex, and race in an additional analysis, on MINT total score. Inclusion of variables in the current models was based on previous research showing that naming tests are often affected by age and education, as well as the interaction between them. Sex was examined as a main effect given a previous finding

of a sex effect on the MINT (Weintraub et al., 2018). Race effects were also considered given previous research showing that disparities in quality of education and early life experiences can adversely impact cognitive test performance of older African-Americans (Manly et al., 2004; Manly et al., 2002; Sisco et al., 2015; Welsh et al., 1995). Age and education level were modeled as continuous predictors. Sex (female, male), race (White, African-American), and diagnostic group (normal cognition, MCI, AD with CDR 0.5, AD with CDR 1, AD with CDR 2) were dummy-coded and modeled as categorical predictors (nominal or ordinal, respectively). Before computing interaction terms, age and education were centered to avoid multicollinearity. Age was modeled as a quadratic term for the cognitively normal group because a polynomial curve provided a better model fit, whereas education was modeled as a linear term because a polynomial term did not significantly improve model fit. Participants were stratified into four age (<65, 65–75, 76–85, 86+) and three education (≤ 12 , 13–15, 16+) groups for the purposes of conducting *t*-tests comparing age- and education-matched groups and providing tables with means, standard deviations, and cut-off scores corresponding to percentiles.

To identify the potential diagnostic utility of the MINT, the area under the curve (AUC), sensitivity, specificity, and ideal hypothetical cut-off MINT scores for predicting group membership were calculated separately for cognitively normal versus MCI and cognitively normal versus AD using Receiver Operating Characteristic (ROC) curve analyses. The AUC measure provides an overall indication of the diagnostic accuracy; a minimum value of >0.75 was considered clinically significant (Fan et al., 2006). Sensitivity (i.e., the true positive rate or hit) and specificity (i.e., the true negative rate or correct rejection) were calculated and sensitivity was plotted as a function of specificity. Thus, the ROC curves illustrate diagnostic accuracy for all possible cut-off scores (from which an optimal cut-off can then be determined).

All data were analyzed using SPSS, version 24 (IBM; Chicago, IL) and R (version 3.3.3; R Core Team, 2016)

Table 2. Pearson bivariate correlations between MINT total score and age and education, separately by diagnosis group and race

		All participants	Normal Cognition	MCI and AD
Overall sample	Age	-.151***	-.147***	-.088***
	Education	.202***	.300***	.121***
	<i>n</i>	5,981	3,981	2,000
White	Age	-0.133***	-0.140***	-0.069**
	Education	0.176***	0.212***	0.117***
	<i>n</i>	5,005	3,280	1,725
African-American	Age	-0.303***	-0.311***	-0.225**
	Education	0.230***	0.291***	0.181*
	<i>n</i>	824	607	217

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

was used for graphs and robust regressions. The statistical results from parametric tests (i.e., regression, t -tests) need to be interpreted with caution because the MINT scores and the standardized residuals were not normally distributed, and there were unequal variances between groups. Given the very large sample sizes, however, violations of the normality assumption may not noticeably impact results, and transformations may actually bias estimates (Schmidt & Finan, 2018). Therefore, we chose to present non-transformed data, and used non-parametric tests such as Mann–Whitney U tests and robust regressions, whenever applicable (see below).

RESULTS

Regression Analyses

Correlations between age and education and total MINT score, shown for the overall sample and separately for Whites and African-Americans, the two most-represented groups in our overall sample, are presented in Table 2 and results of the regression analyses are summarized in Table 3. When all participants were included, multiple regression analyses showed significant associations between naming ability and diagnostic group, sex, education, and age (all $ps < .001$; Table 3a). This full model accounted for 39% of the variance in MINT scores ($F(1, 5,975) = 760.0$; $MSE = 13.5$; $p < .001$). MINT scores decreased as dementia severity and age increased, and increased as education level increased, and were higher for males than females. The main effect of sex can be visualized in Table 4. The main effects of age and education were qualified by a significant interaction between them, such that the association between naming ability and age depended on participants' education level ($p = .008$). When this analysis was carried out within the cognitively normal group only, the same patterns of results were obtained (see Table 3b). When the analysis was carried out within only the MCI and AD groups, there were significant associations between MINT score and diagnostic group, sex, age, and education, but the interaction between

age and education did not reach significance ($p = .25$; see Table 3c). Given the small effect size and to overcome the limitations of parametric tests, we tested the robustness of the age-by-education interaction in the cognitively normal group with a linear regression method using MM-estimation ('robustbase' package in R) as well as with a Weighted Least Squares (WLS) regression, which is more equipped for heteroscedasticity. Both analyses showed that the age-by-education interaction was indeed robust and persistent ($\beta = 0.05$; $p < 0.001$; and $\beta = 0.04$; $p = 0.01$, for robust and WLS regression, respectively), and all other main effects remained significant. Finally, the analyses reported in Table 3 were repeated using a 4 standard deviation trim of MINT and Montreal Cognitive Assessment (MOCA) scores within age, education, and race groups (given significantly different education levels among Whites and African-Americans)—all main effects and interactions remained the same.

Planned comparisons collapsing across age and education using Mann–Whitney U tests (which do not assume a normal distribution or homogeneity of variances between groups) revealed that MINT scores were significantly different for cognitively normal versus MCI ($U = 1,079,054$; $p < .001$; $r = .24$), MCI versus AD with CDR 0.5 ($U = 120,490$; $p < .001$; $r = .21$), AD with CDR 0.5 versus AD with CDR 1 ($U = 79,209$; $p < .001$; $r = .21$), and AD with CDR 1 versus AD with CDR 2 ($U = 39,912$; $p < .001$; $r = .25$). Figure 1 illustrates the main effects of diagnostic group plotted separately for age (1A) and education (1B), showing that the MINT was associated with impairments in naming across dementia severity regardless of age or education. Figure 2 illustrates the age-by-education interaction in cognitively healthy participants, showing that MINT scores were lowest in the oldest-old with the lowest education level. Table 5 presents means and standard deviations, when stratified by age, education, and dementia severity groups. Independent sample t -tests were carried out to compare the mean naming scores between adjacent diagnostic groups matched by age and education (e.g., MCI vs. normal cognition, AD with CDR 0.5 vs. MCI, etc). In general, the more severely impaired diagnostic group performed worse on the MINT than the next least impaired group, except in the youngest and oldest dementia groups. MINT scores of cognitively normal participants showed a near-ceiling effect (see Tables 5 and 6; Figures 1 and 2) that was particularly evident in individuals with higher education (i.e., 13 years and more). Collapsing across age and education, 32% of the cognitively normal sample obtained a perfect score of 32 points and 81% of the sample scored 29 or above. Table 6 presents cut scores on the MINT for the bottom 2%, 10%, and 15% of individuals in the cognitively normal group, stratified by age and education. In light of the observed ceiling effects, these values can be used to supplement the normative data provided by Weintraub et al. (2018) to aid clinical interpretation of MINT scores.

Table 3. Multiple regression results showing the association between MINT scores and diagnostic group, sex, education, and age shown as three separate models for (a) all participants, (b) cognitively normal only, and (c) MCI and AD groups

	Variable	B (SE)	β	95% CI	p
a. All participants	Diagnosis Group	-2.389 (.042)	-.584	-2.472, -2.306	< .001
	Sex	1.113 (.099)	.116	.918, 1.308	< .001
	Education	.185 (.018)	.105	.149, .221	< .001
	Age	-.038 (.005)	-.075	-.048, -.028	< .001
	Age x Education	.005 (.002)	.027	.001, .009	.008
b. Cognitively normal only	Sex	.779 (.070)	.163	.641, .917	< .001
	Education	.232 (.013)	.260	.206, .258	< .001
	Age	-.034 (.004)	-.138	-.042, -.027	< .001
	Age ² *	-.002 (.000)	-.109	-.003, -.002	< .001
	Age x Education	.005 (.001)	.054	.002, .008	< .001
c. MCI and AD	Diagnosis group	-2.755 (.119)	-.455	-2.988, -2.522	< .001
	Sex	1.639 (.257)	.128	1.134, 2.143	< .001
	Education	.093 (.046)	.041	.003, .182	.042
	Age	-.049 (.014)	-.070	-.076, -.022	< .001
	Age x Education	.006 (.005)	.023	-.004, .015	.246

Sex: 0 = female; 1 = male

Diagnosis group: 0 = normal cognition; 1 = MCI; 2 = AD (CDR = 0.5); 3 = AD (CDR = 1); 4 = AD (CDR = 2)

*A quadratic term for age produced a better model fit.

Table 4. Means (and standard deviations) of MINT scores for females and males by diagnosis group

Diagnosis Group	Females	Males
Cognitively Normal	29.7 (2.4)	30.6 (1.7)
MCI	27.5 (3.7)	28.6 (3.7)
AD (CDR = .5)	25.5 (4.7)	26.4 (6.0)
AD (CDR = 1)	22.2 (6.4)	24.5 (6.7)
AD (CDR = 2)	16.3 (8.4)	20.7 (8.8)

Race Effects

To consider if the observed age-by-education interaction differed by race (White, African-American), and whether the MINT may require separate norms for African-Americans, race was added as an additional factor to the regression within the cognitively normal group (given a sufficient sample size of African-American cognitively healthy participants; $n = 607$). Individuals in other race categories were excluded: that is Asian: $n = 53$; Native-Hawaiian/Pacific Islander: $n = 3$; American Indian/Alaska Native: $n = 22$; Other: $n = 10$; Unknown: $n = 6$. We tested a three-way interaction between age, education, and race (and included all lower order two-way interactions). Table 7 shows the results of this regression. Similar to the previous analysis, there were main effects of age, sex, and education. The previously robust interaction between age and education (see Table 3) became marginally significant in this more complex model. African-American cognitively normal elders on average had lower MINT scores than cognitively normal Whites ($M = 27.8$ vs 30.4). Further, education effects were significantly stronger in the African-American than in the White group, an education-by-race

interaction, and race differences on the MINT were larger in the oldest participants, an age-by-race interaction (see Table 7). The three-way interaction between age, education, and race was not significant ($p = 0.51$).

When the same analysis was restricted to patients only (MCI and AD), there was a main effect of race such that African-American patients on average had lower MINT scores than Whites ($M = 24.1$ vs 25.5; $B = -1.484$; $SE = 0.419$; $\beta = -0.073$; $p < 0.001$; 95% CI = $-2.306, -0.662$). Higher order interactions were not significant ($ps \geq 0.12$). Given the observed main effect of race and significant two-way interactions in cognitively normal participants, Table 8 shows means and standard deviations separately for Whites (Table 8a) and African-Americans (Table 8b). Table 9 shows percentile cut-offs separately for Whites (9a) and African-Americans (9b). Table 9b is stratified by two (instead of four) age groups because of insufficient data in some cells for African-Americans.

Receiver Operating Characteristic (ROC) curves

Figure 3 depicts ROC curves for classifying participants based on MINT scores. MINT performance provided good diagnostic accuracy for classifying probable AD versus cognitively normal participants (AUC = 0.85; $SE = 0.01$; $p < 0.001$; 95% CI = 0.84, 0.87). A cut-off MINT score of less than or equal to 28.5 produced a sensitivity and specificity of 75% and 81%, respectively, while a cut-off score of less than or equal to 29.5 produced sensitivity and specificity of 83% and 70%, respectively. In contrast, MINT performance did not provide acceptable diagnostic accuracy for classifying MCI versus cognitively normal participants (AUC = 0.68; $SE = 0.01$; $p < 0.001$; 95% CI = 0.66, 0.70).

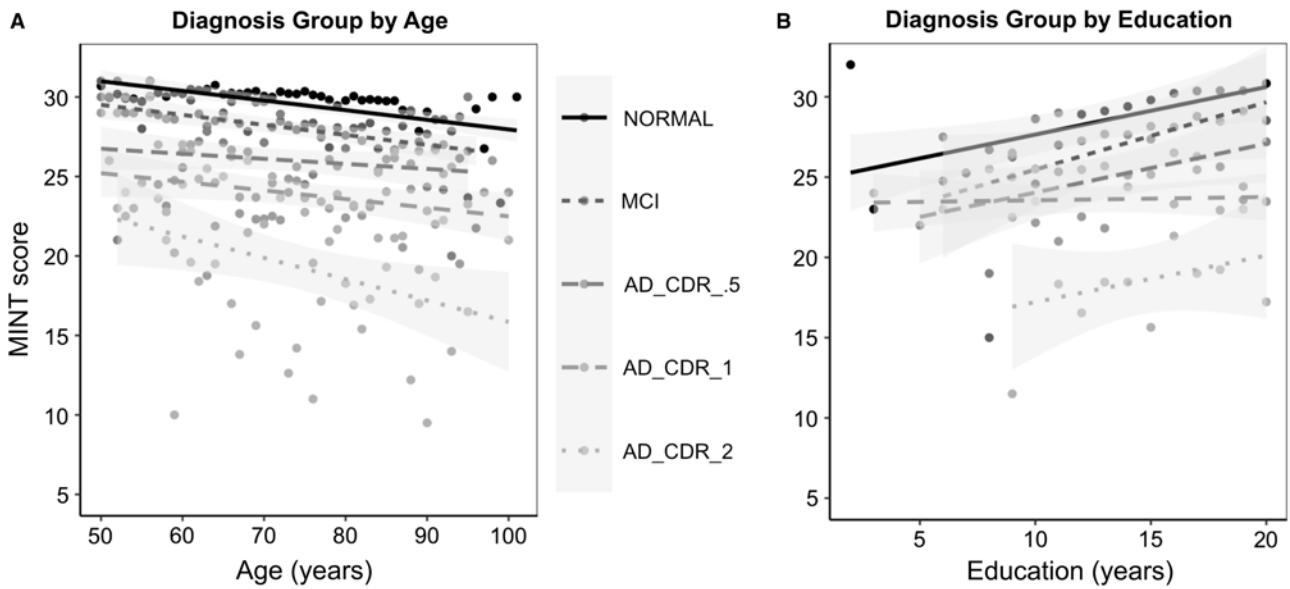


Fig. 1. Graphs depicting the MINT’s ability to detect differences in picture naming ability across level of cognitive impairment plotted separately for (A) age and (B) education. The black line represents the cognitively normal group, with increasing levels of cognitive impairment represented by lighter shading.

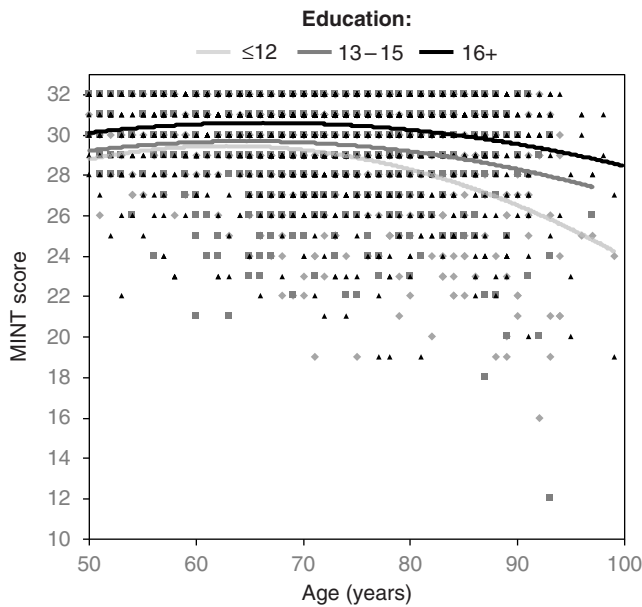


Fig. 2. Scatterplots showing that MINT naming scores were lowest in individuals who are oldest and with the lowest level of education in the cognitively normal elderly. Education level was modeled as a continuous predictor for regression analyses but for visualization purposes was stratified into three levels. Scatter points for 16+ years of education are depicted in black triangles, 13–15 in dark gray squares, and ≤12 years in light gray diamonds.

DISCUSSION

Results of this study demonstrate that the MINT can detect group differences in naming ability at different stages of cognitive impairment associated with AD regardless of age

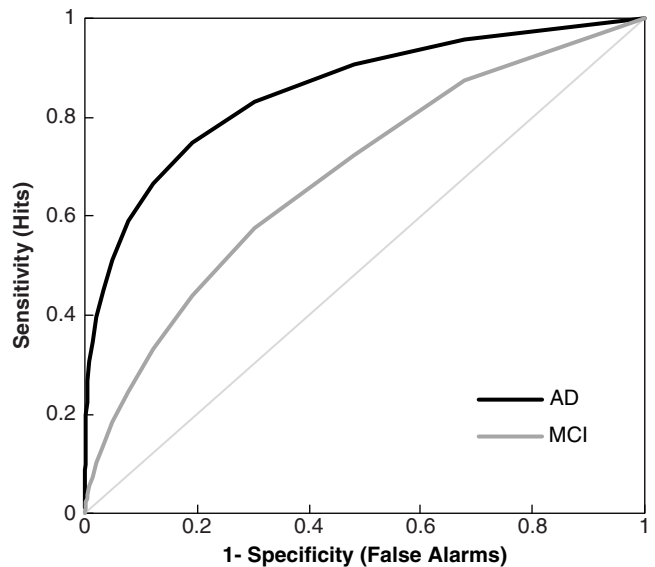


Fig. 3. Receiver Operating Characteristic (ROC) curves comparing sensitivity and specificity of MINT scores in discriminating between cognitively normal participants from those with AD (in black) and those with MCI (in dark gray).

and education level. Within age- and education-matched subgroups, participants with MCI generally scored significantly worse than healthy controls, participants with mild dementia scored worse than those with MCI, and participants with moderate dementia scored worse than those with mild dementia. Of clinical relevance, MINT scores (collapsed across age and education levels) provided good diagnostic accuracy for distinguishing AD from cognitively normal participants, but

Table 5. Mean MINT score and (standard deviation^a), and accompanying *n*, grouped by age, education, diagnosis, and CDR score^b

Age	Education	Normal Cognition		MCI		Dementia due to AD					
		<i>M</i> (<i>SD</i>)	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>n</i>	CDR = 0.5		CDR = 1		CDR = 2	
						<i>M</i> (<i>SD</i>)	<i>n</i>	<i>M</i> (<i>SD</i>)	<i>n</i>	<i>M</i> (<i>SD</i>)	<i>n</i>
50-64	≤ 12	29.8 (1.8)	69	26.8 (2.6) ^{***}	16	27.1 (3.0)	13	23.7 (5.8) [^]	22	20.3 (3.0) ^b	4
	13-15	29.5 (2.5)	144	28.5 (3.6) [*]	23	27.1 (3.4)	15	27.5 (4.3) ^b	8	20.7 (9.7) [^]	11
	16+	30.6 (1.8)	491	28.4 (5.5) ^{***}	61	26.5 (6.9) [^]	46	23.0 (8.5) [*]	47	22.0 (6.3)	15
65-75	≤ 12	28.9 (2.7)	183	28.2 (2.9) [^]	63	25.6 (5.4) ^{**}	38	23.6 (6.1)	36	17.5 (10.3) ^{**}	18
	13-15	29.6 (2.2)	313	27.9 (2.9) ^{***}	54	24.0 (5.9) ^{***}	29	27.5 (4.3)	35	14.6 (9.2) ^{**}	11
	16+	30.5 (1.8)	1351	28.9 (2.9) ^{***}	206	26.2 (5.7) ^{***}	92	23.0 (8.5) ^{**}	103	19.9 (7.4) [*]	26
76-85	≤ 12	28.5 (2.8)	156	26.6 (4.5) ^{**}	46	24.6 (3.5) [^]	28	21.7 (5.9) [*]	62	15.6 (9.5) ^{***}	29
	13-15	29.3 (2.1)	179	27.5 (3.8) ^{***}	52	26.5 (3.4)	22	24.5 (5.8) [^]	43	18.4 (6.9) ^{**b}	9
	16+	30.3 (2.0)	752	28.1 (3.7) ^{***}	225	26.0 (5.6) ^{***}	72	23.3 (6.5) ^{**}	119	19.4 (9.3) ^{**}	49
86+	≤ 12	26.3 (3.9)	56	25.9 (4.6)	21	23.4 (6.0) ^b	7	22.2 (5.4)	22	16.9 (7.2) [*]	13
	13-15	28.3 (4.0)	63	27.2 (3.5)	21	26.4 (4.9) ^b	5	20.4 (6.0) ^{^ b}	9	15.0 (9.8) ^b	4
	16+	29.5 (2.6)	224	27.7 (4.1) ^{***}	64	27.1 (4.5)	18	24.0 (7.6) [^]	43	19.7 (10.0) [^]	25

^a The standard deviations in this table should be interpreted with caution given ceiling effects on MINT scores (for review of these issues see Uttil, 2005).

^b These cells have < 10 participants contributing to the mean

Asterisks denote a significant difference in scores between two adjacent diagnostic groups (i.e., comparison with the group to the left). * = *p* < .05; ** = *p* < .01; *** = *p* < .001; ^ = marginal significance.

Table 6. MINT cut-off scores associated with number of cognitively normal participants (*n* = 3,981) falling below < 2%, < 10%, < 15% of the sample, and within normal limits (WNL) grouped by age and education*. A “-” is placed wherever less than 10 participants contributed to a cell.

Age	Education	<i>n</i>	< 2 %	< 10 %	< 15 %	WNL
50-64	≤ 12	69	-	-	-	28
	13-15	144	-	25	27	28
	16+	491	-	27	28	29
65-75	≤ 12	183	-	24	26	27
	13-15	313	-	26	27	28
	16+	1351	25	27	28	29
76-85	≤ 12	156	-	23	25	26
	13-15	179	-	25	26	27
	16+	752	24	27	28	29
86+	≤ 12	56	-	-	-	21
	13-15	63	-	24	26	27
	16+	224	-	25	26	27

*For example, fewer than 10% of cognitively normal participants who were between 50 and 64 years old, and with 13-15 years of education had a MINT score of 25 or lower.

Table 7. Multiple regression results showing the association between MINT scores and sex, education, age, and race in cognitively normal elders (*n* = 3,887)

Variable	<i>B</i> (<i>SE</i>)	β	95% CI	<i>p</i>
Sex	0.650 (0.065)	0.137	0.522, 0.778	<.001
Education	0.134 (0.014)	0.151	0.107, 0.162	<.001
Age	-0.030 (0.004)	-0.121	-0.037, -0.023	<.001
Age ²	-0.002 (0.000)	-0.104	-0.003, -0.002	<.001
Race	-2.116 (0.094)	-0.341	-2.302, -1.931	<.001
Age x Education	0.003 (0.002)	0.028	0.000, 0.006	.087*
Education x Race	0.136 (0.032)	0.073	0.073, 0.200	<.001
Age x Race	-0.050 (0.010)	-0.082	-0.070, -0.030	<.001
Age x Education x Race	-0.002 (0.003)	-0.012	-0.008, 0.004	.511

Sex: 0 = female; 1 = male.

Race: 0 = White; 1 = African-American.

*Note this interaction was significant in a model without the three-way interaction when including all participants; see Table 3b.

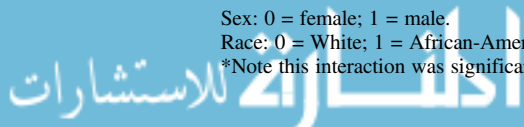


Table 8. Mean MINT score and (standard deviation^a), and accompanying *n*, grouped by age, education and diagnosis and shown separately for Whites (*n* = 5,005) and African-Americans (*n* = 824)^b

	Age	Education	Normal Cognition		MCI		Dementia due to AD	
			<i>M(SD)</i>	<i>n</i>	<i>M(SD)</i>	<i>n</i>	<i>M(SD)</i>	<i>n</i>
a. White	50-64	≤ 12	30.0 (1.6)	50	28.8 (1.2) ^b	6	24.4 (5.4)	33
		13-15	30.5 (1.4)	94	29.3 (2.7)	18	25.1 (7.0)	33
		16+	30.9 (1.4)	428	28.7 (5.7)	54	24.5 (7.6)	102
	65-75	≤ 12	29.8 (2.0)	114	28.6 (3.1)	40	23.3 (7.5)	79
		13-15	30.1 (1.9)	225	28.6 (2.6)	39	22.4 (7.3)	67
		16+	30.7 (1.6)	1159	29.1 (2.9)	183	24.2 (6.4)	195
	76-85	≤ 12	29.4 (2.1)	110	27.4 (4.0)	35	21.2 (7.4)	95
		13-15	29.8 (1.9)	143	28.3 (3.3)	40	24.3 (5.9)	70
		16+	30.5 (1.7)	678	28.3 (3.7)	203	23.5 (7.3)	217
	86+	≤ 12	28.3 (2.7)	28	26.5 (4.8)	17	21.6 (6.6)	31
		13-15	29.5 (3.2)	49	27.4 (3.6)	19	20.4 (8.1)	14
		16+	29.9 (2.1)	202	28.0 (4.1)	56	23.8 (7.8)	79
b. African-American	50-64	≤ 12	29.2 (2.2)	18	25.3 (2.5) ^b	8	21.0 (–) ^b	1
		13-15	27.4 (3.0)	44	29.5 (0.7) ^b	2	27.0 (–) ^b	1
		16+	28.6 (2.4)	45	27.0 (4.4) ^b	5	20.8 (11.9) ^b	4
	65-75	≤ 12	27.1 (3.0)	62	27.0 (2.3)	19	20.6 (8.2) ^b	8
		13-15	28.3 (2.3)	79	25.9 (3.2)	14	19.8 (5.4) ^b	5
		16+	29.2 (2.2)	158	27.5 (2.7)	20	25.0 (6.6)	21
	76-85	≤ 12	26.3 (3.0)	44	24.3 (5.2)	10	19.7 (6.0)	20
		13-15	27.3 (1.8)	34	25.1 (4.8)	11	26.3 (3.2) ^b	3
		16+	28.2 (3.0)	61	26.9 (3.2)	18	21.8 (6.5)	18
	86+	≤ 12	24.1 (3.9)	27	23.3 (2.5) ^b	4	18.6 (6.5) ^b	9
		13-15	24.1 (3.6)	14	25.5 (3.5) ^b	2	22.8 (5.6) ^b	4
		16+	25.6 (3.7)	21	24.8 (3.9) ^b	5	20.4 (11.5) ^b	5

^a The standard deviations in these tables should be interpreted with caution given ceiling effects on MINT scores (for review of these issues see Uttl, 2005)

^b These cells have < 10 participants contributing to the mean

not for distinguishing MCI from cognitively normal participants. Finally, we found robust associations between MINT scores and sex and race, as well as an interaction between age and education in cognitively healthy adults, such that MINT scores were lowest in the oldest and least-educated adults. Although the effect size of this interaction was statistically small, the finding was robust and clinically meaningful (e.g., in Table 6 there was a 6-point difference on the MINT between the least and most educated in the oldest-old group; while there was only a 1-point difference between the least and most educated in the youngest group using the ‘within normal limits’ cut-off). However, additional studies are needed to assess this age-by-education interaction with a more sensitive picture naming test. MINT scores were near ceiling levels for most cognitively normal participants in our sample (i.e., 81% of participants had a MINT score of 29 and above), particularly in White adults with 13 or more years of education, so it could have underestimated their naming difficulties. Overall, the MINT has the potential to detect deficits in naming ability in those with MCI or mild AD dementia and to track word-finding difficulty throughout the course of AD, but demographic characteristics and ceiling effects need to be considered.

Lower education was associated with poorer performance on the MINT—those with the lowest level of education (high school or less) had the lowest naming scores, followed by those with some college, and then those with 16 or more years of education. Notably, education effects were stronger in the African-American group, likely reflecting discordance between years of education and quality of education in ethnic minorities and immigrants. Older age was also associated with poorer naming ability—those in the oldest-old group (86+) had the lowest MINT scores, followed by those in the next oldest groups (76–85 and 65–75), and then those in the youngest group (<65). Similar age and education effects have been reported for the BNT (Albert et al., 1988; Fastenau et al., 1998; LaBarge et al., 1986; Nicholas et al., 1985).

In follow-up analyses including only the White and African-American groups (the next largest race represented in the sample with *n* = 217 and *n* = 607 in the patient and control groups, respectively)—and controlling for age, education, and sex—there was a robust main effect of race such that on average African-Americans named fewer pictures than Whites, consistent with some previous studies (Roberts & Hamsher, 1984; Welsh et al., 1995; but see Manly et al., 1998 in which differences disappeared after accounting for acculturation), and a robust interaction between

Table 9. MINT cut-off scores associated with number of cognitively normal participants falling below or within a given percentile range or within normal limits (WNL), grouped by age and education* and presented separately for (a) Whites ($n = 3,280$) and (b) African-Americans ($n = 607$). A “-” is placed wherever less than 10 participants contributed to a cell.

	Age	Education	n	< 2%	< 10%	< 15%	WNL
a. White	50-64	≤ 12	50		–	–	28
		13-15	94	–	28	28	29
		16 +	428	–	28	29	30
	65-75	≤ 12	114		–	27	28
		13-15	225	–	28	28	29
		16 +	1159	25	28	28	29
	76-85	≤ 12	110		–	26	27
		13-15	143	–	–	27	28
		16 +	678	25	27	28	29
	86 +	≤ 12	28	–	–	–	26
		13-15	49	–	27	27	28
		16 +	202	–	26	27	28
	Age	Education	n	6-10%	11-15%	WNL	
b. African-American	< 70	≤ 12	42	–	–	25	
		13-15	84	–	24	25	
		16 +	128	25	26	27	
	70 +	≤ 12	109	20	21	22	
		13-15	87	23	24	25	
		16 +	157	23	24	25	

*For example, fewer than 10% of cognitively normal Whites who were between 50 and 64 years old, and with 13-15 years of education had a MINT score of 28 or lower. 6-10% of cognitively normal African-Americans who were less than 70 years old, and with 16+ years of education had a MINT score of 25 or lower.

education and race in the cognitively normal group, suggesting that education level affected MINT performance relatively more in African-Americans than Whites. These findings suggest that items on naming tests may often be culturally and/or linguistically biased, and reflect differences in quality of education and early life experiences. In line with this, we also observed a robust age-by-race effect, such that the biggest race differences on the MINT were observed in the oldest cognitively healthy elderly. The oldest African-Americans may have received the lowest quality of education compared to those educated in the post-segregation era (Aiken-Morgan et al., 2015; Anderson, 1988; Manly, 2005 for review), and therefore may have had the lowest degree of exposure to and familiarity with items on the MINT. Consideration of reading abilities rather than education has been shown to attenuate differences in neuropsychological test performance between elderly African-Americans and Whites (Manly et al., 2002, 2004), but reading scores were not available for the current sample.

A particularly notable finding was the age-by-education interaction effect on MINT scores in the cognitively normal group—individuals with low education level and older age exhibited the lowest naming ability on the MINT. This finding is consistent with previous reports of an interaction between age and education on the BNT in cognitively normal elderly (Ashaie & Obler, 2014; Welch et al., 1996). The effects of education on cognition might not be linear; once

a critical number of years of schooling (a threshold) is reached, an individual may subsequently be better able to continue to accrue vocabulary knowledge throughout a lifetime of experiences. By contrast, individuals with lower (or poorer quality) years of education may fail to accrue additional knowledge at the same rate because of insufficient basis of knowledge from which to build. On this view, the interaction effect implies that while higher (post-high school) education levels might have some impact on cognition, education more drastically affects naming ability at the lowest levels (e.g., Ashaie & Obler, 2014; Welch et al., 1996). Very low education levels might also limit cognitive reserve, which refers to the ability to maintain normal levels of cognitive performance despite the presence of pathology due to protective factors and compensatory processes (Mortimer, 1997; Stern, 2002; for reviews see Fratiglioni & Wang, 2007, Meng & D’Arcy, 2012). Education is widely regarded as a protective factor and a proxy for cognitive reserve (for reviews see Caamaño-Isorna et al., 2006; Meng & D’Arcy, 2012). Because age is a risk factor for AD, individuals with the lowest education level may be in the earliest stages of AD and may eventually progress to MCI or AD dementia. However, the cognitive reserve hypothesis is speculative and must be confirmed with follow-up longitudinal studies examining these age and education interaction effects in ‘robust normal control groups’ (Sliwinski et al., 1996;

Edmonds et al., 2015) who do not subsequently decline on neuropsychological measures over the subsequent 1–2 years or have biomarker evidence that they do not harbor AD pathology.

Across all groups, MINT scores were higher for males than for females. This finding is consistent with the report of the preliminary norming study for the UDS neuropsychological test battery (Weintraub et al., 2018), and with studies that show males generally outperform females on the BNT and other picture naming tasks in both AD and normal control groups (Hall et al., 2012; Laiacona et al., 1998; Randolph et al., 1999). A possible explanation for this difference is the higher number of non-living ($n = 24$) than living ($n = 8$) items on the MINT. A number of studies have demonstrated robust sex-by-category interaction effects in picture naming with males performing better than females with non-living items, and females performing better than males with living items (Laiacona et al., 1998; Laws, 1999; McKenna & Parry, 1994). This interaction is most likely due to sex-related differences in item familiarity. A similar sex-related interaction was observed across semantic fluency tasks where cognitively normal males retrieved more names of tools than cognitively normal females, whereas the females retrieved more names of fruit than the males (Capitani et al., 1999).

Limitations and Future Directions

This study revealed several psychometric limitations of the 32-item MINT, the most significant being the near-ceiling effect in cognitively normal elderly individuals. Although the MINT, like the BNT, was not developed to detect subtle naming deficits (e.g., the BNT was developed to assess anomia in aphasia), these limitations reduce the MINT's ability to detect very early decline in naming ability that might be associated with AD. ROC curve analyses showed that the MINT was reasonably sensitive in discriminating cognitively normal participants from those with AD, but not from those with MCI. To at least partially address these psychometric limitations, we provided MINT cut-off scores in terms of percentiles (see Tables 6 and 9)—an approach that may be more useful than means and standard deviations for determining impairment. The observed ceiling effect on the MINT may have occurred because the 32 items from the original 68-item MINT were selected for their sensitivity based on a fairly small sample ($N = 130$; Ivanova et al., 2013) and may not generalize well to other independent samples. It should be noted, however, that the 32 items were chosen to be matched in difficulty to the BNT items; MINT items that were not included tended to be easier items (as the MINT was designed to assess a wider range of naming abilities in the non-dominant language). Thus, an alternative selection of items from the full MINT would likely increase rather than reduce ceiling effects.

There are several aspects of the NACC cohort that could have influenced the results. First, more males than females in the dementia group may have artificially inflated the means

for that group in light of findings that males score higher than females on the MINT. The NACC cohort is also highly educated on average and predominantly English-speaking and Caucasian, thus reducing the generalizability of these findings to other populations. Despite these psychometric and sample limitations, the current analyses reveal that the MINT detects deficits in naming ability across a range of levels of cognitive impairment, including mild-to-moderate stages of dementia. Given that almost all items were named by cognitively normal individuals (producing ceiling effects on the test), the failure of patients to name more difficult items is likely due to semantic loss commonly seen in AD (Garrard et al., 2005; Hodges et al., 1991; Ivanova et al., 2013) rather than unfamiliarity with difficult items.

Because the MINT was originally developed for use in multiple languages, it would be fruitful to examine the sensitivity of the MINT to AD-related language decline in non-English speaking populations, and to determine the impact of demographic factors on MINT performance in these populations. The impact of bilingualism on MINT performance should also be examined given the known effects of bilingualism on picture naming performance (e.g., Gollan et al., 2007). Our finding of a significant effect of race suggests that comprehensive normative data including this variable should be considered in future work. Additional studies are needed to validate the utility of the MINT in participants with lower levels of education with additional measures of quality of education that work better for ethnic minorities (e.g., a test of single word reading; Manly et al., 2002).

Finally, longitudinal studies are needed to determine if the MINT effectively tracks the progression of naming deficits through the stages of AD from preclinical disease through frank dementia. A biomarker-defined diagnosis of AD in cognitively normal individuals (i.e., preclinical AD; Sperling et al., 2011) is becoming increasingly considered as part of the AD spectrum (Jack et al., 2016, 2018), so future assessment of the MINT's utility in biomarker-defined preclinical AD is warranted. It would also be interesting to examine the impact of co-morbid vascular risk factors (e.g., hypertension, hyperlipidemia, white matter changes on MRI) on MINT performance. Subcortical white matter changes in individuals with AD, for example, may be related to decline in attention, working memory, retrieval processes, or inefficient processing of visual stimuli (Prins et al., 2005) that could impact picture naming ability.

In conclusion, the MINT effectively detects naming impairments in mild-to-moderate AD dementia. There are, however, significant effects of age, education, sex, and race on MINT performance in cognitively normal elderly and in those with AD. This suggests that consideration of demographically corrected normative data is essential for accurately determining naming impairment in AD.

ACKNOWLEDGEMENTS

We wish to thank the UCSD Shiley-Marcos Alzheimer's Disease Research Center (P50 AG05131), and Jörg Matt,

Marc Norman, Bob Heaton, and Lily Kamalyan for helpful discussion. AS was funded by a Ruth L. Kirschstein National Research Service Award (NRSA) Individual Predoctoral Fellowship from the NIA (F31 AG058379-02). THG was funded by the National Institute on Deafness and Other Communication Disorders (011492). DMJ was funded by P50 AG05131. DPS was funded by P50 AG05131, RO1 grant AG049810 and the Helen A. Jarret Chair for Alzheimer's Disease Research, and is a paid consultant for Takeda Pharmaceuticals, Inc. and Aptinyx, Inc. The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA-funded ADCs: 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).

CONFLICTS OF INTEREST

The authors have nothing to disclose.

REFERENCES

- Aiken-Morgan, A.T., Gamaldo, A.A., Sims, R.C., Allaire, J.C., & Whitfield, K.E. (2015). Education desegregation and cognitive change in African American older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 70, 348–356.
- 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, 270–279.
- Albert, M.S., Heller, H.S., & Milberg, W. (1988). Changes in naming ability with age. *Psychology and Aging*, 3, 173–178.
- Anderson, J.D. (1988). *The education of Blacks in the South, 1860–1935*. Chapel Hill: University of North Carolina Press.
- Ashaie, S., & Obler, L. (2014). Effect of age, education, and bilingualism on confrontation naming in older illiterate and low-educated populations. *Behavioural Neurology*, 2014, 1–10.
- Au, R., Joung, P., Nicholas, M., Obler, L.K., Kass, R., & Albert, M.L. (1995). Naming ability across the adult life span. *Aging and Cognition*, 2, 300–311.
- Balota, D.A., & Duchek, J.M. (1991). Semantic priming effects, lexical repetition effects, and contextual disambiguation effects in healthy aged individuals and individuals with senile dementia of the Alzheimer type. *Brain and Language*, 40, 181–201.
- Balthazar, M.L.F., Martinelli, J.E., Cendes, F., & Damasceno, B.P. (2007). Lexical semantic memory in amnesic mild cognitive impairment and mild Alzheimer's disease. *Arquivos De Neuropsiquiatria*, 65, 619–622.
- Barbeau, E.J., Didic, M., Joubert, S., Guedj, E., Koric, L., Felician, O., Ranjeva, J.P., Cozzone, P., & Ceccaldi, M. (2012). Extent and neural basis of semantic memory impairment in mild cognitive impairment. *Journal of Alzheimer's Disease*, 28, 823–837.
- Barresi, B.A., Nicholas, M., Tabor Connor, L., Obler, L.K., & Albert, M.L. (2000). Semantic degradation and lexical access in age-related naming failures. *Aging, Neuropsychology, and Cognition*, 7, 169–178.
- Burke, D.M., & Shafto, M.A. (2004). Aging and language production. *Current Directions in Psychological Science*, 13, 21–24.
- Capitani, E., Laiacona, M., & Barbarotto, R. (1999). Sex affects word retrieval of certain categories in semantic fluency tasks. *Cortex*, 35, 273–278.
- Caamaño-Isorna, F., Corral, M., Montes-Martínez, A., & Takkouche, B. (2006). Education and dementia: a meta-analytic study. *Neuroepidemiology*, 26, 226–232.
- Connor, L.T., Spiro, A., Obler, L.K., & Albert, M.L. (2004). Change in object naming ability during adulthood. *The Journals of Gerontology*, 59, 203–209.
- Edmonds, E.C., Delano-Wood, L., Clark, L.R., Jak, A.J., Nation, D.A., McDonald, C.R., Libon, D.J., Au, R., Galasko, D., Salmon, D.P., Bondi, M.W., & Alzheimer's Disease Neuroimaging Initiative. (2015). Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. *Alzheimer's & Dementia*, 11, 415–424.
- Fan J., Upadhye S., & Worster A. (2006). Understanding receiver operating characteristic (ROC) curves. *CJEM*, 8, 19–20.
- Fastenau, P.S., Denburg, N.L., & Mauer, B.A. (1998). Parallel short forms for the Boston Naming Test: psychometric properties and norms for older adults. *Journal of Clinical and Experimental Neuropsychology*, 20, 828–834.
- Feyereisen, P. (1997). A meta-analytic procedure shows an age-related decline in picture naming: comments on Goulet, Ska, and Kahn (1994). *Journal of Speech, Language, and Hearing Research*, 40, 1328–1333.
- Fratiglioni, L., & Wang, H.-X. (2007). Brain reserve hypothesis in dementia. *Journal of Alzheimer's Disease*, 12, 11–22.
- Garrard, P., Lambon Ralph, M.A., Patterson, K., Pratt, K.H., & Hodges, J.R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: a new approach. *Brain and Language*, 93, 79–94.

- Gollan, T.H., Fennema-Notestine, C., Montoya, R.I., & Jernigan, T.L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society, 13*, 197–208.
- Gollan, T.H., Weissberger, G.H., Runnqvist, E., Montoya, R.I., & Cera, C.M. (2012). Self-ratings of spoken language dominance: A Multi-Lingual Naming Test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Bilingualism, 15*, 594–615.
- Goulet, P., Ska, B., & Kahn, H.J. (1994). Is there a decline in picture naming with advancing age? *Journal of Speech and Hearing Research, 37*, 629–644.
- Hall, J.R., Vo, H.T., Johnson, L.A., Wiechmann, A., & O'Bryant, S.E. (2012). Boston Naming Test: Sex differences in older adults with and without Alzheimer's dementia. *Psychology, 3*, 485–488.
- Hashimoto, N., Johnson, B., & Peterson, A. (2016). The effects of thematic relations on picture naming abilities across the lifespan. *Neuropsychology, Development, and Cognition, 23*, 499–512.
- Hodges, J.R., Salmon, D.P., & Butters, N. (1991). The nature of the naming deficit in Alzheimer's and Huntington's disease. *Brain, 114*, 1547–1558.
- Hodges, J.R., Salmon, D.P., & Butters, N. (1992). Semantic memory impairment in Alzheimer's disease: Failure of access or degraded knowledge? *Neuropsychologia, 30*, 301–314.
- Huff, F.J., Corkin, S., & Growdon, J.H. (1986). Semantic impairment and anomia in Alzheimer's disease. *Brain and Language, 28*, 235–249.
- Hyman, B.T., Van Hoesen, G.W., Damasio, A.R., & Barnes, C.L. (1984). Alzheimer's disease: cell-specific pathology isolates the hippocampal formation. *Science, 225*, 1168–1170.
- Ivanova, I., Salmon, D.P., & Gollan, T.H. (2013). The Multilingual Naming Test in Alzheimer's disease: Clues to the origin of naming impairments. *Journal of the International Neuropsychological Society, 19*, 272–283.
- Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, E.G., & Petersen, R.C. (1995). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE token, WRAT-R reading, AMNART, Stroop, TMT, JLO. *The Clinical Neuropsychologist, 10*, 262–278.
- Jack, C.R., Bennett, D.A., Blennow, K., Carrillo, M.C., Dunn, B., Haeberlein, S.B., Holtzman, D.M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J.L., Montine, T., Phelps, C., Rankin, K.P., Rowe, C.C., Scheltens, P., Siemers, E., Snyder, H.M., Sperling, R., & Contributors. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association, 14*, 535–562.
- Jack, C.R., Bennett, D.A., Blennow, K., Carrillo, M.C., Feldman, H.H., Frisoni, G.B., Hampel, H., Jagust, W.J., Johnson, K.A., Knopman, D.S., Petersen, R.C., Scheltens, P., Sperling, R.A., & Dubois, B. (2016). A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology, 87*, 539–547.
- Kaplan E.F., Goodglass H., & Weintraub S. (1983). *The Boston naming test* (2nd ed.). Philadelphia: Lea & Febiger.
- Kavé, G., Knafo, A., & Gilboa, A. (2010). The rise and fall of word retrieval across the lifespan. *Psychology and Aging, 25*, 719–724.
- LaBarge, E., Edwards, D., & Knesevich, J.W. (1986). Performance of normal elderly on the Boston Naming Test. *Brain and Language, 27*, 380–384.
- Laiacona, M., Barbarotto, R., & Capitani, E. (1998). Semantic category dissociations in naming: is there a sex effect in Alzheimer's disease? *Neuropsychologia, 36*, 407–419.
- Laws, K.R. (1999). Sex affects naming latencies for living and non-living things: Implications for familiarity. *Cortex, 35*, 729–733.
- Lukatela, K., Malloy, P., Jenkins, M., & Cohen, R. (1998). The naming deficit in early Alzheimer's and vascular dementia. *Neuropsychology, 12*, 565–572.
- Mackay, A.I., Connor, L.T., Albert, M.L., & Obler, L.K. (2002). Noun and verb retrieval in healthy aging. *Journal of the International Neuropsychological Society, 8*, 764–770.
- Manly, J. (2005). Advantages and disadvantages of separate norms for African Americans. *The Clinical Neuropsychologist, 19*, 270–275.
- 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*, 37–46.
- Manly, J.J., Miller, S.W., Heaton, R.K., Byrd, D., Reilly, J., Velasquez, R.J., Saccuzzo, D.P., & Grant, I. (1998). The effect of African-American acculturation on neuropsychological test performance in normal and HIV-positive individuals. *The HIV Neurobehavioral Research Center (HNRC) Group. Journal of the International Neuropsychological Society: JINS, 4*(3), 291–302.
- Manly, J.J., Jacobs, D.M., Touradji, P., Small, S.A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society, 8*, 341–348.
- McKenna, P., & Parry, R. (1994). Category specificity in the naming of natural and man-made objects: Normative data from adults and children. *Neuropsychological Rehabilitation, 4*, 255–281.
- McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C., Morris, J.C., Rossor, M.N., Scheltens, P., Carrillo, M.C., Thies, B., Weintraub, S., & Phelps, C.H. (2011). The diagnosis of dementia 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*, 263–269.
- Meng, X., & D'Arcy, C. (2012). Education and dementia in the context of the Cognitive Reserve Hypothesis: A systematic review with meta-analyses and qualitative analyses. *PLoS ONE, 7*, e38268.
- Mitrushina, M., & Satz, P. (1995). Repeated testing of normal elderly with the Boston Naming Test. *Aging, 7*, 123–127.
- Monsell, S.E., Dodge, H.H., Zhou, X.-H., Bu, Y., Besser, L.M., Mock, C., Hawes, S.E., Kukull, W.A., Weintraub, S., & Neuropsychology Work Group Advisory to the Clinical Task Force. (2016). Results from the NACC Uniform Data Set Neuropsychological Battery crosswalk study. *Alzheimer Disease and Associated Disorders, 30*, 134–139.
- Morris, J.C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology, 43*, 2412–2414.
- Mortimer, J.A. (1997). Brain reserve and the clinical expression of Alzheimer's disease. *Geriatrics, 52*, S50–S53.
- Mulatti, C., Calia, C., De Caro, M.F., & Della Sala, S. (2014). The cumulative semantic interference effect in normal and pathological ageing. *Neuropsychologia, 65*, 125–130.

- Nicholas, M., Obler, L., Albert, M., & Goodglass, H. (1985). Lexical retrieval in healthy aging. *Cortex*, *21*, 595–606.
- Prins, N.D., van Dijk, E.J., den Heijer, T., Vermeer, S.E., Jolles, J., Koudstaal, P.J., & Hofman, A. Breteler, M.M. (2005). Cerebral small-vessel disease and decline in information processing speed, executive function and memory. *Brain*, *128*, 2034–2041.
- Randolph, C., Lansing, A.E., Ivnik, R.J., Cullum, C.M., & Hermann, B.P. (1999). Determinants of confrontation naming performance. *Archives of Clinical Neuropsychology*, *14*, 489–496.
- R Core Team (2016). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rizio, A.A., Moyer, K.J., & Diaz, M.T. (2017). Neural evidence for phonologically based language production deficits in older adults: An fMRI investigation of age-related differences in picture-word interference. *Brain and Behavior*, *7*, e00660.
- Roberts, R.J., & Hamsher, K.D. (1984). Effects of minority status on facial recognition and naming performance. *Journal of Clinical Psychology*, *40*, 539–545.
- Rogers, T.T., Ivanoiu, A., Patterson, K., & Hodges, J.R. (2006). Semantic memory in Alzheimer's disease and the frontotemporal dementias: a longitudinal study of 236 patients. *Neuropsychology*, *20*, 319–335.
- Schmidt, A.F., & Finan, C. (2018). Linear regression and the normality assumption. *Journal of Clinical Epidemiology*, *98*, 146–151.
- Sheng, L., Lu, Y., & Gollan, T.H. (2014). Assessing language dominance in Mandarin-English bilinguals: Convergence and divergence between subjective and objective measures. *Bilingualism: Language and Cognition*, *17*, 364–383.
- Sisco, S., Gross, A.L., Shih, R.A., Sachs, B.C., Glymour, M.M., Bangen, K.J., Benitez, A., Skinner, J., Schneider, B.C., & Manly, J.J. (2015). The role of early-life educational quality and literacy in explaining racial disparities in cognition in late life. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *70*, 557–567.
- Sliwinski, M., Lipton, R.B., Buschke, H., & Stewart, W. (1996). The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *The Journals of Gerontology*, *51*, 217–225.
- Sperling, R.A., Aisen, P.S., Beckett, L.A., Bennett, D.A., Craft, S., Fagan, A.M., Iwatsubo, T., Jack, C.R. Jr., Kaye, J., Montine, T.J., Park, D.C., Reiman, E.M., Rowe, C.C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M.C., Thies, B., Morrison-Bogorad, M., Wagster, M.V., & Phelps, C.H. (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*, 280–292.
- Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*, 448–460.
- Tomoschuk, B., Ferreira, V.S., & Gollan, T.H. (2018). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition*, *22*, 516–536.
- Uttl, B. (2005). Measurement of Individual Differences: Lessons From Memory Assessment in Research and Clinical Practice. *Psychological Science*, *16*, 460–467.
- Van Gorp, W.G., Satz, P., Kiersch, M.E., & Henry, R. (1986). Normative data on the Boston Naming Test for a group of normal older adults. *Journal of Clinical and Experimental Neuropsychology*, *8*, 702–705.
- Weintraub, S., Besser, L., Dodge, H.H., Teylan, M., Ferris, S., Goldstein, F.C., Giordani, B., Kramer, J., Loewenstein, D., Marson, D., Mungas, D., Salmon, D., Welsh-Bohmer, K., Zhou, X.H., Shirk, S.D., Atri, A., Kukull, W.A., Phelps, C., & Morris, J.C. (2018). Version 3 of the Alzheimer Disease Centers' Neuropsychological Test Battery in the Uniform Data Set (UDS). *Alzheimer Disease and Associated Disorders*, *32*, 10–17.
- Welch, L.W., Doineau, D., Johnson, S., & King, D. (1996). Educational and sex normative data for the Boston Naming Test in a group of older adults. *Brain and Language*, *53*, 260–266.
- Welsh, K.A., Fillenbaum, G., Wilkinson, W., Heyman, A., Mohs, R.C., Stern, Y., Harrell, L., Edland, S.D., & Beekly, D. (1995). Neuropsychological test performance in African-American and white patients with Alzheimer's disease. *Neurology*, *45*, 2207–2211.
- Willers, I.F., Feldman, M.L., & Allegri, R.F. (2008). Subclinical naming errors in mild cognitive impairment: A semantic deficit? *Dementia & Neuropsychologia*, *2*, 217–222.
- Williams, B.W., Mack, W., & Henderson, V.W. (1989). Boston Naming Test in Alzheimer's disease. *Neuropsychologia*, *27*, 1073–1079.
- Zec, R., Markwell, S., Burkett, N., & Larsen, D. (2005). A longitudinal study of confrontation naming in the "normal" elderly. *Journal of the International Neuropsychological Society*, *11*, 716–726.

