

## RESEARCH ARTICLE

# The MINT Sprint 2.0: A picture naming test for detection of naming impairments in Alzheimer's disease and in preclinical AD

Tamar H. Gollan<sup>1</sup> | Dalia L. Garcia<sup>2</sup> | Alena Stasenko<sup>1</sup> | Mayra Murillo<sup>1</sup> | Chi Kim<sup>3</sup> | Douglas Galasko<sup>3</sup> | David P. Salmon<sup>3</sup><sup>1</sup>Department of Psychiatry, University of California, San Diego, La Jolla, California, USA<sup>2</sup>Joint Doctoral Program in Language and Communicative Disorders, San Diego State University/University of California, San Diego, La Jolla, California, USA<sup>3</sup>Department of Neurosciences, University of California, San Diego and Shiley-Marcos Alzheimer's Disease Research Center, La Jolla, California, USA**Correspondence**

Tamar H. Gollan, Department of Psychiatry, University of California, 9500 Gilman Drive, La Jolla 92093, CA, USA.

E-mail: [tgollan@ucsd.edu](mailto:tgollan@ucsd.edu)**Funding information**

NIH, Grant/Award Numbers: P30-AG062429, AG076415, NS119285, AG077915; NSF BCS, Grant/Award Number: 1923065

**Abstract****INTRODUCTION:** Evidence on the onset of naming deficits in Alzheimer's disease (AD) is mixed. Some studies showed an early decline, but others did not. The present study introduces evidence from a novel naming test.**METHODS:** Cognitively normal ( $n = 138$ ), mild cognitive impairment (MCI;  $n = 21$ ), and Alzheimer's disease (AD;  $n = 31$ ) groups completed an expanded Multilingual Naming Test with a time-pressured administration procedure (MINT Sprint 2.0). Cerebrospinal fluid biomarkers classified participants as true controls ( $n = 61$ ) or preclinical AD ( $n = 26$ ).**RESULTS:** Total correct MINT Sprint 2.0 scores exhibited good sensitivity and specificity ( $>0.85$ ) for discriminating true controls from cognitively impaired (MCI/AD) groups and showed significant differences between true controls and preclinical AD groups. Time measurement did not improve classification, but percent resolved scores exhibited promise as an independent AD marker.**DISCUSSION:** Naming deficits can be detected in the earliest stages of AD with tests and procedures designed for this purpose.**KEYWORDS**

aging, cerebrospinal fluid biomarkers, multilingual naming test (MINT), preclinical Alzheimer's disease, picture naming, speeded naming

## 1 | BACKGROUND/INTRODUCTION

Word finding difficulty is often one of the earliest reported symptoms of Alzheimer's disease (AD), though typically it is not diagnostic because naming failures (especially for proper names) also increase with healthy aging.<sup>1-5</sup> While the outcome might be the same—inability to retrieve names—the underlying cause of naming deficits might differ in AD versus in healthy aging. In healthy aging, naming failures are most commonly attributed to later processing stages such as activa-

tion of phonological representations.<sup>1,6</sup> By contrast, naming failures in AD may arise from deficits in all stages of lexical retrieval including activation of semantic knowledge.<sup>7,8</sup> Supporting this view, individuals with AD are less likely than cognitively healthy older adults to resolve naming failures when provided with a phonological cue<sup>9</sup> and report less access to partial phonological information about words they are attempting to retrieve when stuck in a tip-of-the-tongue (TOT) state of temporary retrieval failure<sup>10,11</sup> (e.g., they are less likely to report “it starts with a P”).

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.

**RESEARCH IN CONTEXT**

- 1. Systematic review:** The existing literature presents mixed evidence on the emergence of naming deficits in Alzheimer's disease (AD). Some studies suggest naming deficits (and language deficits in general) only emerge later in the course of the disease (long after memory impairments), whereas other studies found significant naming deficits even in preclinical AD.
- 2. Interpretation:** We introduce a novel and more powerful time-pressured naming test with a unique administration procedure that reveals naming deficits can be detected even in individuals classified as cognitively normal but who are at risk for AD based on cerebrospinal fluid (CSF) biomarkers. Additionally, cognitively healthy older adults who do not have AD risk (i.e., true controls based on AD biomarkers) are more likely to rapidly resolve failed naming responses without a cue than individuals at risk for AD and than individuals with cognitive impairment. These results suggest that subtle impairments in semantic knowledge occur very early in the course of AD.
- 3. Future directions:** Additional work is needed to replicate these findings with a larger number of individuals at risk for AD, to confirm the eventual diagnosis of AD in individuals with positive CSF biomarkers, to test the diagnostic power of the new naming test in multilinguals (and the other languages for which the test was developed), and to better characterize how time pressure, aging, and AD jointly affect naming ability.

While deficits in episodic and semantic memory are often found in a preclinical or prodromal stage of AD,<sup>12–14</sup> anomia and frank aphasia typically emerge later in disease progression (with the exception of the relatively rare clinical dementia syndrome of primary progressive aphasia; see review by Weintraub et al.<sup>15</sup>). However, evidence from formal objective tests of naming is mixed. Some studies have found that picture naming scores are not impaired in those with mild cognitive impairment (MCI<sup>16</sup>) or that they do not provide good diagnostic accuracy for distinguishing cognitively normal individuals from those with MCI.<sup>17</sup> Other studies suggest that picture naming is sensitive for detecting those with MCI who will soon progress to AD dementia<sup>18</sup> or an even earlier stage, before a diagnosis of MCI (perhaps especially in participants with less than college education level<sup>19,20</sup>). Relatively few studies have investigated this question and even fewer have asked if naming impairments might be identified during *preclinical* AD when there is no clinical diagnosis, but risk of cognitive impairment is apparent through biomarkers.

In the present study, we examined the effects of AD on the ability to perform the Multilingual Naming Test (MINT) Sprint 2.0, a speeded confrontation naming test.<sup>21</sup> We hypothesized that the MINT Sprint

2.0 would be more sensitive than traditional naming tests for identifying early cognitive changes of AD because of three unique features: (1) a larger number of items to increase reliability, (2) a time-pressured administration procedure that taxes lexical retrieval processes, and (3) a multi-staged administration procedure that measures momentary retrieval failures and their resolution. We further predicted that the different MINT Sprint 2.0 scores might be differentially sensitive to different stages of AD. We hypothesized that the total correct naming score would be the best measure overall for detecting naming impairments at all stages of AD because it requires success at all processing levels including activation of semantics, lexical retrieval, and phonological retrieval. We hypothesized that efficiency scores, which jointly consider naming success and speed,<sup>22</sup> and a percent resolution (PR) score, which reflects the ability to resolve momentary naming failures, would be most sensitive for detecting preclinical AD in which complete naming failures are less likely.

**2 | METHODS****2.1 | Participants**

The current study included 190 monolingual English-speaking participants from the longitudinal study of the University of California, San Diego (UCSD) Alzheimer's Disease Research Center (ADRC). We focused primarily on a subset of participants contrasting (1) cognitively impaired ( $n = 52$ ) versus true controls (i.e., cognitively intact and negative for cerebrospinal fluid [CSF] biomarkers of AD;  $n = 61$ ), and (2) true controls versus preclinical AD based on the presence of CSF biomarkers ( $n = 26$ ). Table 1 shows participant characteristics (see [supporting information](#) for larger sample characteristics). ADRC participants complete annual evaluations with a clinical and medical history, brief medical examination, neurological and neuropsychological assessment, screening for depression and other psychiatric symptoms, assessment of functional activities of daily living, and laboratory tests. Results are reviewed by at least two board-certified neurologists and a neuropsychologist to reach a consensus clinical diagnosis. This is a two-step process with classification as cognitively normal, MCI,<sup>23</sup> or dementia followed by assignment of a presumed etiology based on published criteria for AD,<sup>24</sup> dementia with Lewy bodies (DLB<sup>25</sup>), frontotemporal dementia (FTD<sup>26</sup>), or other neurological disease. Only patients with dementia presumed to be due to AD were included. Dementia severity was classified as mild ( $\geq 120$  points) or moderate ( $< 120$  points) based on Dementia Rating Scale (DRS) scores (range 0–144). The MINT Sprint 2.0 scores were not considered in reaching the diagnosis.

A research lumbar puncture (LP) was completed for CSF biomarker analysis on a subset of participants at one of the ADRC evaluations. Biomarker information was not used in making the clinical diagnosis. Research LP and preanalytical preparation and storage of CSF were performed using standardized procedures<sup>27</sup> following recommended best practices.<sup>28</sup> CSF (15–25 mL) was collected in the morning after overnight fasting. Samples were processed, aliquoted into 500  $\mu$ L

**TABLE 1** Characteristics and test scores for (a) true controls versus (b) preclinical AD, and (c) cognitively impaired participants with MCI or AD (groups [a] and [c] were included in ROCs; see Table 4, Figure 1).

| Type of characteristic     | Demographic/measure               | All cognitively normal participants with CSF biomarker data |       |   |       |                    |       |
|----------------------------|-----------------------------------|---|-------|---|-------|--------------------|-------|
|                            |                                   | True controls, biomarker negative (n = 61)                  |       | Preclinical AD, biomarker positive (n = 26) |       | MCI/AD (n = 52)    |       |
|                            |                                   | M   | SD    | M   | SD    | M                  | SD    |
| Demographic                | <b>Age</b>                        | 77.3 <sup>a</sup>   | 5.1   | 80.0 <sup>b</sup>                           | 6.2   | 79.0 <sup>a</sup>  | 7.2   |
|                            | Education                         | 17.5 <sup>a</sup>   | 2.2   | 17.1 <sup>a</sup>                           | 1.6   | 17.2 <sup>a</sup>  | 2.0   |
|                            | Male/female                       | 28/33   |       | 13/13                                       |       | 28/24              |       |
| Cognitive status           | MMSE                              | 28.9 <sup>a</sup>   | 1.3   | 28.8 <sup>a</sup>                           | 1.4   | 23.9 <sup>b</sup>  | 4.2   |
|                            | MoCA                              | 25.9 <sup>a</sup>   | 2.4   | 25.3 <sup>a</sup>                           | 2.7   | 18.2 <sup>b</sup>  | 5.2   |
|                            | DRS                               | 139.4 <sup>a</sup>  | 3.4   | 138.0 <sup>a</sup>                          | 4.8   | 122.9 <sup>b</sup> | 15.9  |
| CSF biomarker <sup>1</sup> | <b>Tau/Aβ42</b>                   | .3 <sup>a</sup>   | 0.1   | 1.4 <sup>b</sup>                            | 2     | 1.3 <sup>b</sup>   | 1.0   |
|                            | <b>Tau</b>                        | 284.4 <sup>a</sup>  | 126.8 | 576.0 <sup>b</sup>                          | 235.4 | 541.1 <sup>b</sup> | 343.1 |
|                            | <b>Aβ42</b>                       | 954.4 <sup>a</sup>  | 424.1 | 534.2 <sup>b</sup>                          | 182.2 | 510.5 <sup>b</sup> | 223.1 |
|                            | CSF to cognitive test lag (years) | 2.9 <sup>a</sup>  | 1.6   | 3.4 <sup>a</sup>                            | 3.1   | 3.0 <sup>a</sup>   | 2.8   |
| Existing naming test       | 32-item MINT                      | 31.2 <sup>a</sup>   | 1.4   | 30.6 <sup>a</sup>                           | 1.8   | 26.6 <sup>b</sup>  | 5.9   |
| MINT Sprint 2.0 scores     | First pass correct                | 74.5 <sup>a</sup>   | 3.9   | 72.7 <sup>a</sup>                           | 5.4   | 62.6 <sup>b</sup>  | 13.9  |
|                            | Second pass correct               | 3.7 <sup>a</sup>  | 2.1   | 4.0 <sup>a</sup>                            | 2.7   | 6.8 <sup>b</sup>   | 4.6   |
|                            | <b>Total correct</b>              | 78.1 <sup>a</sup>   | 2.3   | 76.7 <sup>b</sup>                           | 3.5   | 69.3 <sup>b</sup>  | 11.8  |
|                            | Percent resolved (PR)             | 71.0 <sup>a</sup>   | 24.5  | 62.4 <sup>a</sup>                           | 25.4  | 44.1 <sup>b</sup>  | 17.8  |
|                            | Efficiency first pass             | 2.2 <sup>a</sup>  | 0.7   | 2.3 <sup>a</sup>                            | 0.6   | 4.5 <sup>b</sup>   | 4.2   |
|                            | Efficiency total                  | 3.2 <sup>a</sup>  | 1.4   | 3.8 <sup>a</sup>                            | 1.5   | 7.4 <sup>b</sup>   | 5.2   |
|                            | Minutes first pass                | 2.0 <sup>a</sup>  | 0.5   | 2.1 <sup>a</sup>                            | 0.4   | 2.9 <sup>b</sup>   | 1.0   |
|                            | Minutes total                     | 3.1 <sup>a</sup>  | 1.2   | 3.6 <sup>a</sup>                            | 1.2   | 5.7 <sup>b</sup>   | 2.5   |

Note: Values in the same row that do not share the same superscript are significantly different at  $p < 0.05$  (2-tailed); bolding highlights significant differences between biomarker negative versus positive cognitively normal controls.

Abbreviations: Aβ, amyloid beta; AD, Alzheimer's disease; CSF, cerebrospinal fluid; DRS, Dementia Rating Scale; MCI, mild cognitive impairment; MINT, Multilingual Naming Test; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; ROC, receiver operating characteristic; SD, standard deviation.

<sup>1</sup>Biomarker data missing for two participants with MCI and two with AD. Efficiency = completion time/percent correct.

fractions in polypropylene microtubes, snap-frozen, and stored at  $-80^{\circ}\text{C}$  until assayed. Samples were analyzed on an automated Lumipulse platform using assays developed with established monoclonal antibodies<sup>29</sup> (Fujirebio Inc.). CSF AD biomarkers included amyloid beta (Aβ)1-40, Aβ1-42, total tau, and phosphorylated tau 181 (p-tau181). The ratio of total tau over Aβ1-42 (tau/Aβ42) was used as a composite biomarker of AD. A cut-point for biomarker positivity based on Lumipulse data was derived for the ratio from CSF samples from  $n = 462$  UCSD ADRC participants who ranged from cognitively normal to severely demented. A tau/Aβ42 cut-point of .609 was derived using mixture model analysis. This cut-point is consistent with a published Lumipulse assay cut-point (tau/Aβ42 > 0.540) for AD biomarker positivity derived against clinical reads of amyloid positron emission tomography (PET) scans<sup>29</sup> with validation in multiple cohorts. In the current study, classification as biomarker negative was made only if CSF had been collected within 5 years of MINT Sprint 2.0 testing, and we controlled for the age of LP in critical comparisons. This interval

was allowed given the stability of CSF biomarkers over several years.<sup>30</sup> Classification as biomarker positive in a small number of participants was based on CSF collected > 5 years prior to testing (assuming someone remains positive over time). There was no significant difference between biomarker negative versus positive groups in the time interval between CSF collection and MINT Sprint 2.0 testing. Most participants with probable AD ( $n = 24/31$ ) were biomarker positive (four had not contributed CSF, and three were negative within 3 years of testing). About half of those with MCI ( $n = 9/21$ ) were biomarker positive (six were negative, four had not contributed CSF, and two were negative but > 5 years had passed since CSF contribution).

## 2.2 | Mint Sprint 2.0

The MINT Sprint 2.0 test and materials are available at the following link on the [Open Science Framework](#) (OSF). Characteristics for English

**TABLE 2** Characteristics of the English names for all 80 MINT Sprint 2.0 items, the 32-item MINT (currently part of the UDS), and the full 60-item BNT.

|                              | Frequency <sup>a,b</sup> |      | Length in phonemes <sup>c</sup> |     | Length in syllables |     |
|------------------------------|--------------------------|------|---------------------------------|-----|---------------------|-----|
|                              | M                        | SD   | M                               | SD  | M                   | SD  |
| MINT Sprint 2.0 <sup>d</sup> | 39.6                     | 78.1 | 4.1                             | 1.5 | 1.5                 | 0.7 |
| 32-item MINT                 | 8.5                      | 9.0  | 4.4                             | 1.3 | 1.5                 | 0.6 |
| 60-item BNT                  | 23.6                     | 88.0 | 5.2                             | 1.9 | 2.0                 | 0.9 |

Abbreviations: BNT, Boston Naming Test; MINT, Multilingual Naming Test; SD, standard deviation; UDS, Uniform Data Set.

<sup>a</sup>For responses with multiple correct answers, the frequency of the most common answer was taken.

<sup>b</sup>The word "level" from the MINT Sprint 2.0 was excluded because the CELEX (Baayen et al.<sup>31</sup>) does not discriminate between the two different nouns represented by the word (i.e., the tool, which is depicted in the MINT Sprint 2.0, vs. the position on a scale).

<sup>c</sup>Length in phonemes was calculated using CLEARPOND (Marian et al.<sup>32</sup>).

<sup>d</sup>The MINT Sprint 2.0 accepts either "mortar" or "pestle" as correct responses for the picture. Here, we used the frequency of "mortar."

MINT Sprint 2.0 items are shown in Table 2<sup>31,32</sup> with other tests for comparison. The MINT Sprint 2.0 has eight rows of 10 color pictures of objects (each approximately 1 to 1.5 square inches) simultaneously presented on a 17 × 13 inch laminated card (the original test was presented in PowerPoint<sup>21</sup>). The bottom rows contain more difficult items drawn from studies designed to elicit TOT states.<sup>4,33</sup> Like the original MINT tests,<sup>21,34</sup> the MINT Sprint 2.0 was designed to be equally valid in English, Spanish, Mandarin, and Hebrew. The MINT Sprint 2.0 contains all the original 68 MINT items with the exception of 4 items that were replaced for being too difficult ("porthole"), too easy ("hand"), or poorly matched across languages (i.e., "king" and "witch" are rare in Mandarin). Additionally, "mortar" and "pestle" are credited as 1 point in the MINT Sprint 2.0 if either is produced correctly (instead of requiring both words because the Spanish word for "pestle" is rare). In the 2.0 version, we also replaced a few of the original MINT Sprint<sup>21</sup> items to avoid pictures with easily guessed cognate names (e.g., "gyroscope" is "giróscopo"), which can affect naming in bilinguals.<sup>35,36</sup> MINT Sprint 2.0 items are ordered by difficulty, collapsing across languages, based on existing Spanish-English data from Garcia and Gollan,<sup>21</sup> and pilot data from approximately 10 native Hebrew speakers and 10 native Mandarin speakers. Items were swapped locally (moving as little as possible) to avoid having consecutive words beginning with the same sound or rhyming in any of the four languages.

To induce a sense of time pressure, participants are told they have 3 minutes to name as many pictures as they can as quickly as possible starting at the top left corner and making their way across each row. If participants take longer than 3 to 4 seconds on any given picture the examiner says, "keep going" and encourages them to not spend too much time on any one picture. In the first pass, participants are allowed to go back to name items they previously skipped. The 3-minute cutoff is not imposed (i.e., participants are given as much time as they need), but most participants require < 3 minutes to complete their first pass (initial attempt) through the grid. Instructions are: "I am going to show you eight rows of pictures. Starting at the top left, try to name each picture in English from first to last going as quickly as you can without making errors. If you come across one you don't know or can't remember say "don't know" out loud and keep going. If the name comes to

mind later, you can go back and tell us. You will have 3 minutes to name as many pictures as you can."

After participants indicate they are finished, they are prompted to try again to name only items they had skipped or named incorrectly during the first pass. This second attempt only at items that were previously missed is called the second pass. Instructions are: "Now let's see if you can get some of the ones you missed. If you still don't know, just say 'don't know' and we'll move on quickly. I'm going to point out some objects that you either skipped or weren't quite right. Please let me know if the name comes to mind."

The examiner then points to missed items and asks the participant to try again. No semantic or phonemic cues are provided. If the response was incorrect (e.g., "tomato" instead of "apple") the examiner says, "Take a closer look at this one. Do you have a different name for that?" If a superordinate or subordinate response was provided, the examiner says, "Do you have a more specific name/more general name?" If the participant failed to notice an arrow pointing to a critical part of the picture (e.g., if they said "window" instead of "blinds") the examiner says, "See what the arrow is pointing at here. Do you have a name for that?" Finally, if the participant skipped an item or said, "don't know" the examiner says, "Did you see this one?" or "Do you know this one?" Time to complete the first pass and the entire test is recorded.

MINT Sprint 2.0 scores include: (1) number of correct responses on the first pass; (2) number of correct responses on the second pass; (3) total correct (the sum of [1] and [2]); (4) a PR score, which is the number of correct responses on the second pass divided by the number of items missed on the first pass (multiplied by 100); (5) a first pass efficiency score,<sup>37</sup> which is completion time (in minutes) divided by percent correct on the first pass; and (6) a total efficiency score, which is total administration time (in minutes) divided by total percent correct. The dominant response, dominant response variants (e.g., "plane" for "airplane"), picture-specific variants (e.g., "Monarch" for "butterfly"), responses including the target name (e.g., "wishing well" for "well"), and regional variants (e.g., "torch" for "flashlight") are counted as correct. Items named correctly on the first pass (even if produced out of order) are credited in the first pass score. Correct names produced only after prompting are credited in the second pass score.

### 2.3 | Procedure

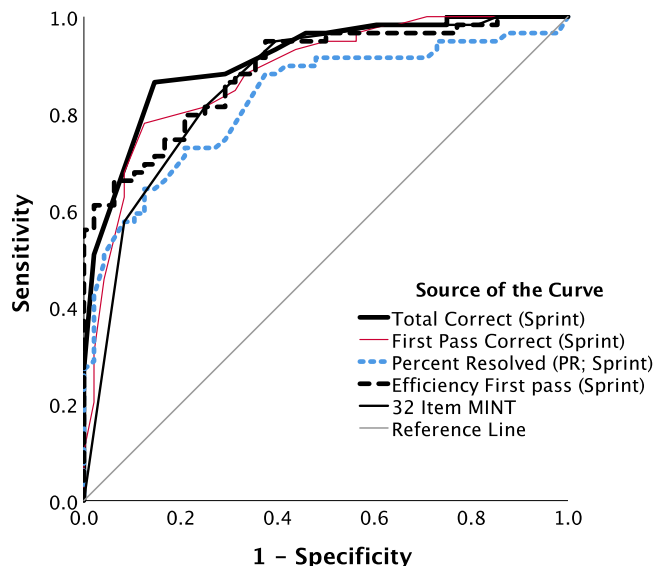
Participants completed the MINT Sprint 2.0 as part of the annual ADRC evaluation, which included procedures of the Uniform Data Set (UDS) to collect systematic information from all federally funded ADRCs.<sup>38,39</sup> The MINT Sprint 2.0 was administered immediately before the 32-item MINT because the latter includes phonological cues that might have facilitated retrieval of overlapping items if the 32-item test had been administered first. Participants were tested individually in a quiet well-lighted room. Audio of the MINT Sprint 2.0 administration was recorded. During testing the examiner recorded pictures named correctly on each pass using a rectangular scoresheet that reproduced the spatial grid and order of pictures on the testing card. The acceptable correct name(s) are printed in the appropriate scoresheet location with room to write any incorrect responses. Response codes were used to record responses (e.g., check mark for correct, a circle for second pass prompt needed). This format facilitated scoring in time with the rapid pace of naming. Responses and completion time were checked against the audio recording by a separate examiner.

### 2.4 | Statistical analyses

Data used for analyses are accessible at the OSF link provided above. Four MINT Sprint 2.0 scores of primary interest were (1) number correct on the first pass, (2) first pass efficiency score, (3) total correct (first plus second pass correct), and (4) PR. First pass correct scores reflect the ability to name pictures under time pressure. First pass efficiency scores jointly consider speed and accuracy in the time-pressured portion of the test. By contrast, total correct scores minimize the effects of time pressure. The number correct on the second pass depends on the number of failed items on the first pass (participants who fail more items on the first pass have more opportunities to gain items on the second pass). Therefore, we focused on PR scores, which reflect the proportion of retrieval failures that were only temporarily failed (and then resolved on the second pass). Minutes needed to complete the first pass and the entire test show average administration time, but we anticipated these values would not be as useful as the first pass efficiency score. The total efficiency score was not of primary interest because the second pass was not time pressured.

We began by exploring the data with Pearson bivariate correlations using all participants tested ( $n = 190$ ; see [supporting information](#)) to determine which (if any) measures might provide novel or relatively independent measures of cognitive status (Table 3). To examine the diagnostic potential of the MINT Sprint 2.0 we focused on three subgroups (see Table 1) including (1) true controls (biomarker-negative participants), (2) preclinical AD (biomarker-positive participants), and (3) cognitively impaired participants which included  $n = 21$  with MCI,  $n = 18$  with mild AD, and  $n = 13$  with moderate AD (see [supporting information](#) for additional information).

To examine the diagnostic potential, we used receiver operating characteristic (ROC) curves and calculated area under the curve (AUC)



**FIGURE 1** Receiver operating characteristic curves classifying participants as being either mild cognitive impairment/Alzheimer's disease or true controls without elevated risk (i.e., biomarker negative; see Tables 1 and 4). MINT, Multilingual Naming Test

to compare individuals with cognitive impairment (MCI/AD) versus true (biomarker negative) controls on four key MINT Sprint 2.0 scores and on the 32-item MINT score for comparison. Optimal sensitivity and specificity were determined by the Youden  $J$  index (Figure 1, Table 4). With this same subset of the data, we also used logistic regression to determine whether adding measures derived from the second pass of the MINT Sprint 2.0 improved on the ability of the first pass to differentiate cognitively impaired (MCI/AD) from true controls.

We then examined sensitivity of the MINT Sprint 2.0 to preclinical AD in two additional analyses. First, we compared the MINT Sprint 2.0 measures in true controls versus preclinical AD groups using linear regression controlling for age and lag between cognitive testing and LP. Second, we compared biomarker negative versus positive controls using simple  $t$  tests in case-control matched groups (matched for age, education, and biological sex). In both analyses, we generated ROC curves to examine the diagnostic sensitivity of MINT Sprint 2.0 scores for detecting preclinical AD. Data were analyzed using Statistical Package for Social Sciences (SPSS) v28.

## 3 | RESULTS

Table 1 shows that true controls (cognitively normal, biomarker negative) and preclinical AD (cognitively normal, biomarker positive) groups had higher scores on cognitive status tests (Mini-Mental State Examination [MMSE],<sup>40</sup> Montreal Cognitive Assessment [MoCA],<sup>41</sup> DRS<sup>42</sup>), and higher naming test scores than cognitively impaired participants. The mean interval between CSF collection and MINT Sprint 2.0 testing was approximately 3 years and did not differ across groups. On average, true controls completed the MINT Sprint 2.0 in just over

**TABLE 3** Pearson bivariate correlations between test scores in all participants tested on the MINT Sprint 2.0 ( $n = 190$ ; see supporting information).

|                                    | DRS                | MMSE         | MoCA   | 32-item MINT | Minutes first pass | Minutes total | First pass correct | Second pass correct | Total correct | Percent resolved (PR) | Efficiency first pass |
|------------------------------------|--------------------|--------------|--------|--------------|--------------------|---------------|--------------------|---------------------|---------------|-----------------------|-----------------------|
| MMSE                               | <b>0.851</b>       |              |        |              |                    |               |                    |                     |               |                       |                       |
| MoCA                               | 0.834              | <b>0.833</b> |        |              |                    |               |                    |                     |               |                       |                       |
| 32-item MINT <sup>a</sup>          | 0.657              | 0.588        | 0.639  |              |                    |               |                    |                     |               |                       |                       |
| Minutes first pass <sup>a</sup>    | -0.641             | -0.594       | -0.663 | -0.602       |                    |               |                    |                     |               |                       |                       |
| Minutes total <sup>a</sup>         | -0.636             | -0.600       | -0.664 | -0.667       | <b>0.817</b>       |               |                    |                     |               |                       |                       |
| First pass correct                 | 0.715              | 0.639        | 0.696  | 0.893        | -0.629             | -0.755        |                    |                     |               |                       |                       |
| Second pass correct <sup>b</sup>   | -0.595             | -0.552       | -0.602 | -0.561       | 0.517              | 0.689         | <b>-0.836</b>      |                     |               |                       |                       |
| Total correct                      | 0.658              | 0.575        | 0.618  | <b>0.929</b> | -0.563             | -0.636        | <b>0.944</b>       | -0.616              |               |                       |                       |
| Percent resolved (PR)              | 0.230 <sup>c</sup> | 0.241        | 0.289  | 0.457        | -0.371             | -0.461        | 0.387              | -0.004 <sup>d</sup> | 0.553         |                       |                       |
| Efficiency first pass <sup>a</sup> | -0.697             | -0.617       | -0.617 | -0.848       | 0.658              | 0.560         | -0.849             | 0.679               | <b>-0.902</b> | -0.336                |                       |
| Efficiency total <sup>a</sup>      | -0.747             | -0.673       | -0.693 | -0.893       | 0.729              | 0.828         | -0.931             | 0.694               | -0.922        | -0.433                | <b>0.908</b>          |

Abbreviations: DRS, Dementia Rating Scale; MINT, Multilingual Naming Test; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment.

Note: With just two marked exceptions, all correlations in this table are significant at the  $p < 0.001$  level. Bolded correlations highlight the strongest correlation each measure exhibits (fewer than 11 values are bolded because in some cases the highest correlation matched in both directions, for example, of all included tests the DRS was most strongly correlated with the MMSE, which was also most strongly correlated with the DRS).

<sup>a</sup> $n = 186$ .

<sup>b</sup> $n = 188$ .

<sup>c</sup> $p = 0.002$ .

<sup>d</sup> $p = 0.955$ .

**TABLE 4** Indices for evaluating the ability of naming test scores to discriminate participants with mild AD/MCI from true controls (see Table 1).

|                       | AUC <sup>a</sup> | Std. error   | 95% confidence interval | Cutoff value | Sensitivity  | Specificity  | Youden       |
|-----------------------|------------------|--------------|-------------------------|--------------|--------------|--------------|--------------|
| <b>Total correct</b>  | <b>0.911</b>     | <b>0.027</b> | <b>[0.858, 0.964]</b>   | <b>76.5</b>  | <b>0.867</b> | <b>0.857</b> | <b>0.724</b> |
| First pass correct    | 0.882            | 0.032        | [0.819, 0.946]          | 72.5         | 0.783        | 0.878        | 0.661        |
| Percent resolved (PR) | 0.830            | 0.039        | [0.753, 0.907]          | 58.1         | 0.733        | 0.796        | 0.529        |
| Efficiency first pass | 0.885            | 0.030        | [0.825, 0.945]          | 2.6          | 0.706        | 0.867        | 0.573        |
| 32-item MINT          | 0.849            | 0.038        | [0.774, 0.924]          | 30.5         | 0.817        | 0.735        | 0.552        |

Abbreviations: AD, Alzheimer's disease; AUC, area under the curve; MCI, mild cognitive impairment; MINT, Multilingual Naming Test.

Note: The row with the highest AUC value and Youden score is highlighted in bold.

<sup>a</sup>All AUC values were significant at  $p < 0.001$  level.

3 minutes and cognitively impaired participants in just under 6 minutes (see Table 1).

Correlations among mental status test scores, MINT scores, and various measures from the MINT Sprint 2.0 are shown in Table 3. Scores on the 32-item MINT were strongly correlated with MINT Sprint 2.0 total scores ( $r = 0.929^1$ ), and both naming measures were strongly correlated with mental status test scores. Contrary to expectations, the first pass efficiency scores (i.e., first pass completion time/first pass percent correct) were not more strongly correlated with mental status test scores ( $r = -0.617$ – $0.697$ ) than the first pass correct alone ( $r = 0.639$ – $0.715$ ). This could imply that including speed in

the measure does not improve sensitivity to cognitive status, or that time pressure and time measurement do not uniquely contribute to test sensitivity. A bit more in line with our expectations, MINT Sprint 2.0 total efficiency scores tended to be more strongly correlated with mental status test scores ( $r = -0.673$ – $0.747$ ) than were total correct scores ( $r = 0.575$ – $0.658$ ), but these differences were just marginally or not significant using a Fisher  $r$  to  $z$  transformation ( $ps \geq 0.09$ ). With few exceptions, all correlations shown in Table 3 were moderate to large and significant at the  $p < 0.001$  level. Notable exceptions were correlations between MINT Sprint 2.0 PR scores and other measures of naming (32-Item MINT:  $r = 0.457$ ) or mental status ( $r = 0.230$ – $0.289$ ). This suggests that PR scores might measure a relatively independent aspect of cognitive functioning that is not assessed by existing tests.

<sup>1</sup> The high correlation between tests is not surprising (given overlap in items between tests administered within the same session) but does suggest that the MINT Sprint 2.0 could supplant the 32-item MINT in the UDS without significantly disrupting longitudinal analyses.

### 3.1 | True controls versus cognitively impaired

Table 1 shows that true controls (cognitively normal, biomarker negative;  $n = 61$ ) and cognitively impaired (MCI/AD;  $n = 52$ ) groups did not differ in age, education, or sex distribution. ROC curves distinguishing cognitively impaired individuals from true controls for the four key MINT Sprint 2.0 measures and the 32-Item MINT are shown in Figure 1 and the AUC values are shown in Table 4. The AUC values for all MINT Sprint 2.0 and 32-item MINT measures were highly robust and in the good range ( $\geq 0.8$ ). The MINT Sprint 2.0 total correct provided the highest AUC value (AUC = 0.911), the best Youden score, and was the only measure with both sensitivity and specificity above 0.85.

Logistic regressions were performed to determine whether MINT Sprint 2.0 first pass correct and PR scores, in combination, could differentiate cognitively impaired from true control individuals better than the first pass correct score alone (a consideration of practical interest given that additional time is needed to administer the second pass). In this analysis, we used PR scores rather than second pass correct scores because the latter depend on the number of items missed on the first pass, and PR scores were relatively independent from measures of mental status (see Table 3). To facilitate interpretation, all predictor variables were  $z$  scored, and odds ratio and 95% confidence interval (CI) values were transformed (by dividing 1 by the outputted value). The full model correctly classified 85.0% of true controls and 78.4% of cognitively impaired individuals (AUC = 0.915, SE = 0.026, 95% CI [0.864, 0.966],  $p < 0.001$ ;  $\chi^2[2] = 10.1$ ;  $p = 0.006$ ; and Nagelkerke  $R^2 = 0.529$ ). The Hosmer and Lemeshow test indicated good model fit ( $p = 0.829$ ). First pass correct scores contributed the most in terms of predictive power (Wald  $\chi^2[1] = 15.844$ ; odds ratio = 12.195; 95% CI [3.559, 41.667];  $p < 0.001$ ), closely followed by PR scores (Wald  $\chi^2[1] = 10.912$ ; odds ratio = 2.933; 95% CI [1.550, 5.556];  $p < 0.001$ ). By contrast, when we repeated the logistic regression using minutes to complete the first pass instead of the PR score, the effect of time was only marginally significant (Wald  $\chi^2[1] = 2.945$ ; odds ratio = 0.466; 95% CI [0.195, 1.115];  $p = 0.086$ ), again suggesting that adding a measure of speed does not improve detection of cognitive status.

### 3.2 | Detection of AD risk in cognitively normal individuals (true controls vs. preclinical AD)

For all cognitively normal individuals with biomarker data ( $n = 87$ ), Table 1 shows that the true controls versus preclinical AD groups did not differ in education or sex distribution. Participants with preclinical AD were significantly older ( $p = 0.034$ ) and achieved lower MINT Sprint 2.0 total correct scores than true controls ( $p = 0.025$  before and  $p = 0.023$  after adjusting for age and years since LP; Table 1). PR scores were marginally different between groups after controlling for age and years since LP ( $p = 0.145$  before and  $p = 0.052$  after adjusting for age and years since LP). However, AUC values from ROC curve analyses were below 0.7 for total correct scores (AUC = 0.615, SE = 0.068, 95% CI [0.481, 0.749],  $p = 0.092$ ), PR scores (AUC = 0.600, SE = 0.068, 95% CI [0.466, 0.734],  $p = 0.144$ ), and age (AUC = 0.635, SE = 0.066, 95%

CI [0.505, 0.765],  $p = 0.048$ ). To further explore possible age effects on PR scores, we examined correlations between age and PR scores in each group. Figure 2 shows that PR scores increased with age in true controls ( $r = 0.289$ ,  $p = 0.025$ ), but not in preclinical AD ( $r = -0.085$ ,  $p = 0.680$ ), and not in cognitively impaired individuals ( $r = -0.011$ ,  $p = 0.941$ ). By contrast, DRS scores decreased with age in all three groups; in true controls ( $r = -0.360$ ,  $p = 0.004$ ), preclinical AD ( $r = -0.543$ ,  $p = 0.004$ ), and marginally in cognitively impaired individuals ( $r = -0.237$ ,  $p = 0.090$ ).

There were no significant differences between true controls and the preclinical AD group on first pass scores, first pass efficiency scores, or the 32-item MINT after covarying age and years since LP (all  $ps \geq 0.104$ ).

For case-control matched groups ( $n = 26$  in each group), Table 5 shows demographic characteristics, mental status test scores, CSF biomarkers, and naming test scores. The preclinical AD group scored significantly lower than the case-control matched true controls on total correct ( $p = 0.04$  or  $p = 0.03$  after adjusting for years since LP), and on PR scores ( $p = 0.04$  or  $p = 0.02$  after adjusting for years since LP).<sup>2</sup> However, ROC analysis showed AUC values below 0.7 for both MINT Sprint 2.0 measures (total correct AUC = 0.643, SE = 0.077, 95% CI [0.492, 0.795],  $p = 0.076$ ; PR score AUC = 0.645, SE = 0.078, 95% CI [0.493, 0.797],  $p = 0.078$ ). None of the other MINT Sprint 2.0 measures, 32-item MINT scores, or mental status scores were significantly different between groups (all  $ps \geq 0.15$ ).

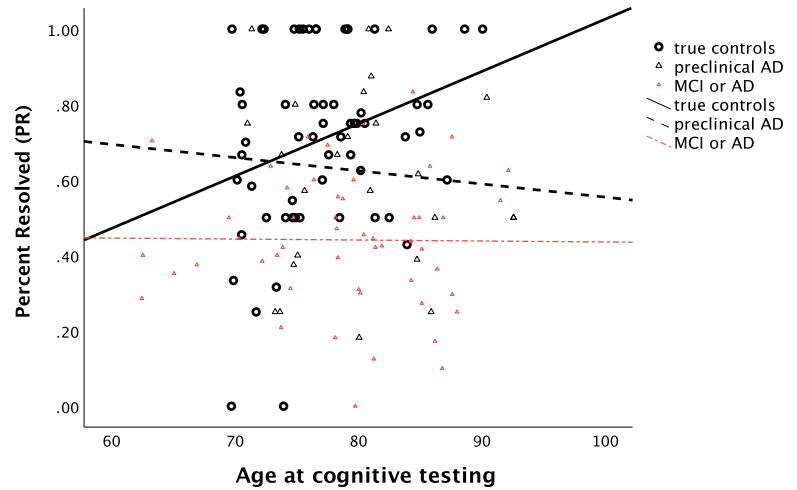
## 4 | DISCUSSION

These results suggest three main conclusions: (1) AD reduces picture-naming ability even preclinically, (2) retrieval failures are less likely to resolve spontaneously in cognitively impaired individuals relative to true controls (and to a lesser extent also in preclinical AD vs. in true controls), and (3) precise measurement of naming speed does not increase sensitivity to AD (but the effects of time pressure remain to be determined). Overall, the most robust MINT Sprint 2.0 measure for detecting AD at any stage was the total correct score, which returned the highest AUC value and Youden index (see Table 4, Figure 1), the only measure with sensitivity and specificity above 0.85, and the only measure that consistently showed significant differences between preclinical AD versus true controls (see Figure 3; Tables 1 and 5).

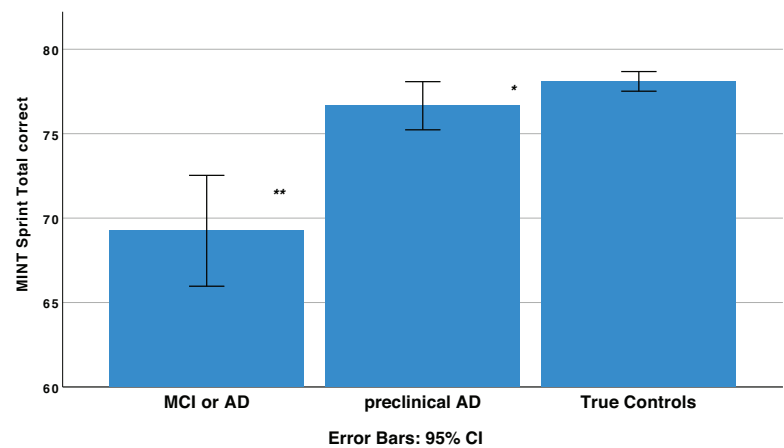
The PR measure exhibited promise as an independent marker of AD. PR scores were less correlated than other naming test measures with existing tests of cognitive status (see Table 3) and explained unique variance along with first pass correct scores in classifying participants as true controls versus cognitively impaired (i.e., AD/MCI). PR scores also differed significantly between true controls versus preclinical AD in case-control matched groups (this difference was just significant overall after controlling statistically for lag between testing and LP and

<sup>2</sup> These were independent  $t$  tests. Note that the difference in the PR scores was only marginally significant when tested with paired  $t$  tests (both  $ps = 0.06$ ). The paired  $t$  test in this instance assumes that matching removed all background variation between matched groups, which likely understates actual variation. The independent  $t$  tests, in contrast, likely overstate variation by ignoring the matching.

**FIGURE 2** Correlation between age at testing and the PR score in true controls (i.e., biomarker negative,  $n = 61$ ), preclinical AD (i.e., biomarker positive;  $n = 26$ ), and cognitively impaired participants (i.e., MCI or AD,  $n = 52$ ; see Table 1). AD, Alzheimer's disease; MCI, mild cognitive impairment; MINT, Multilingual Naming Test; PR, percent resolved



**FIGURE 3** MINT Sprint 2.0 total correct scores were highest in true controls (i.e., biomarker negative,  $n = 61$ ), significantly lower in preclinical AD (i.e., biomarker positive;  $n = 26$ ), and lower still in cognitively impaired participants (i.e., MCI or AD,  $n = 52$ ; see Table 1). AD, Alzheimer's disease; CI, confidence interval; MCI, mild cognitive impairment; MINT, Multilingual Naming Test; PR, percent resolved



between-group differences in age). Diagnostic sensitivity and specificity of PR scores for distinguishing between these two groups was not strong, but total scores also did not meet this standard. Finally, PR scores were *positively* correlated with age in true controls, but not in participants with preclinical AD or in cognitively impaired individuals (see Figure 2).

#### 4.1 | Interpreting the apparently null effects of time pressure

The finding that naming speed did not increase sensitivity to cognitive status seems inconsistent with a study that reported a negative correlation between AD pathology (i.e.,  $A\beta$  plaque burden measured with F-florbetapir PET) and scores on the Rapid Naming Test (RNT), which measured the number of pictures cognitively healthy older adults could name in 60 seconds.<sup>22</sup> The RNT was more strongly associated with level of AD pathology than a 15-item version of the Boston Naming Test (BNT), although diagnostic utility was not assessed. However, the 15-item BNT had fewer items than most participants named in 60 seconds on the RNT (older adults averaged 34 pictures, standard deviation = 7.5), and thus the RNT was better powered to detect deficits. In the

MINT Sprint 2.0 participants are instructed not to spend more than a few seconds trying to name objects that are not immediately accessible. By contrast, in the RNT participants must try to retrieve each failed name for a full 5 seconds. This costly interval during a testing period that lasts only 1 minute likely reduced naming scores for participants who experienced a large number of retrieval failures. Given that the RNT score conflates naming success and naming speed, it is not possible to determine whether the RNT was more sensitive than the 15-item BNT to preclinical AD because of differences in the number of items, time-pressured administration, or both.

The MINT Sprint 2.0 provides separate measures of naming success (total correct score) versus speed (time to complete the first pass). The speed measure did not improve diagnostic utility; however, speed and accuracy were highly correlated (e.g.,  $r = -0.629$  for first pass and total correct scores in Table 3; see also Gollan and Brown<sup>4</sup>). Additional research will be needed to determine whether time pressure and/or naming speed increases diagnostic power in some circumstances. Nonetheless, a compelling reason to include time pressure is simply to increase the number of items than can be administered in a short amount of time. Indeed, retrieval speed is clearly important given sensitivity of PR scores to AD, which reflected the ability to rapidly resolve temporary retrieval failures.



**TABLE 5** Characteristics and test scores for case-control matched cognitively normal participant true controls versus preclinical AD.

| Type of characteristic | Demographic/measure               | Case-control matched cognitively normal participants with CSF biomarker data |       |  |       |
|------------------------|-----------------------------------|--|-------|--|-------|
|                        |                                   | True controls (biomarker negative, <i>n</i> = 26)                            |       | Preclinical AD (biomarker positive, <i>n</i> = 26) |       |
|                        |                                   | M  | SD    | M  | SD    |
| Demographic            | Age                               | 78.7 <sup>a</sup>  | 4.8   | 80.0 <sup>a</sup>                                  | 6.2   |
|                        | Education                         | 17.8 <sup>a</sup>  | 2.1   | 17.1 <sup>a</sup>                                  | 1.6   |
|                        | Male/female                       | 13/13  |       | 13/13  |       |
| Cognitive status       | MMSE                              | 28.7 <sup>a</sup>  | 1.5   | 28.8 <sup>a</sup>                                  | 1.4   |
|                        | MoCA                              | 25.4 <sup>a</sup>  | 2.2   | 25.3 <sup>a</sup>                                  | 2.7   |
|                        | DRS                               | 138.7 <sup>a</sup>   | 3.7   | 138.0 <sup>a</sup>                                 | 4.8   |
| CSF biomarker          | Tau/Aβ42                          | .3 <sup>a</sup>  | 0.1   | 1.4 <sup>b</sup>                                   | 2.0   |
|                        | Tau                               | 300.9 <sup>a</sup>   | 150.6 | 576.0 <sup>b</sup>                                 | 235.4 |
|                        | Aβ42                              | 1006.5 <sup>a</sup>  | 473.2 | 534.2 <sup>b</sup>                                 | 182.2 |
|                        | CSF to cognitive test lag (years) | 2.5 <sup>a</sup>   | 1.5   | 3.4 <sup>a</sup>                                   | 3.1   |
| Existing naming test   | 32-item MINT                      | 31.2 <sup>a</sup>  | 1.0   | 30.6 <sup>a</sup>                                  | 1.8   |
| MINT Sprint scores     | First pass correct                | 74.3 <sup>a</sup>  | 3.8   | 72.7 <sup>a</sup>                                  | 5.4   |
|                        | Second pass correct               | 4.0 <sup>a</sup>   | 2.2   | 4.0 <sup>a</sup>                                   | 2.7   |
|                        | <b>Total correct</b>              | 78.3 <sup>a</sup>  | 2.0   | 76.7 <sup>b</sup>                                  | 3.5   |
|                        | <b>Percent resolved (PR)</b>      | 75.6 <sup>a</sup>  | 18.5  | 62.4 <sup>b</sup>                                  | 25.4  |
|                        | Efficiency first pass             | 2.0 <sup>a</sup>   | 0.6   | 2.1 <sup>a</sup>                                   | 0.4   |
|                        | Efficiency total                  | 3.1 <sup>a</sup>   | 1.4   | 3.6 <sup>a</sup>                                   | 1.2   |
|                        | Minutes first pass                | 2.2 <sup>a</sup>   | 0.8   | 2.3 <sup>a</sup>                                   | 0.6   |
|                        | Minutes total                     | 3.2 <sup>a</sup>   | 1.5   | 3.8 <sup>a</sup>                                   | 1.5   |

Note: Values in the same row that do not share the same superscript are significantly different at *p* < 0.05 (2-tailed). Significant differences between groups highlighted in bold. Efficiency = completion time/percent correct.

Abbreviations: Aβ, amyloid beta; AD, Alzheimer's disease; CSF, cerebrospinal fluid; DRS, Dementia Rating Scale; MCI, mild cognitive impairment; MINT, Multilingual Naming Test; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; SD, standard deviation.

## 4.2 | How does AD pathology versus healthy aging affect picture naming?

With 80 items, the MINT Sprint 2.0 is more powerful than briefer naming tests. The original BNT had 85 items,<sup>43</sup> but was shortened to 60 items.<sup>44</sup> More recently a 30-item version or a 32-item version of the MINT<sup>17,20</sup>—or even shorter 15- or 12-item versions of the BNT—have been used.<sup>19,45–47</sup> The procedure of stopping administration after six failed items (e.g., on the BNT) further limits power. Our results suggest that brief naming tests are underpowered to detect subtle changes in naming in the earliest stages of cognitive decline due to AD. For example, the 32-item MINT score was not sensitive to biomarker status (see Tables 1 and 5) and had specificity below 0.8 for distinguishing true controls from cognitively impaired individuals (see Table 4). The increased sensitivity of the MINT Sprint 2.0 does not come at the cost of administration time. Most cognitively normal participants completed the first pass in only 2 minutes and completed the entire test in approximately 3 minutes (see Table 1 and supporting information).

The unique two-pass administration procedure of the MINT Sprint provided added power. PR scores were lower in cognitively impaired than in cognitively normal participants (see Table 1 and supporting information). Additionally, PR scores increased with age but only in true controls (see Figure 2). Items gained on the second pass, as indexed by the PR score, might include names that were just short of successful retrieval on the first pass. Thus, lower PR scores presumably reflect an increase in the number of profoundly failed retrievals, whereas higher PR scores reflect an increase in almost successful retrievals. This interpretation fits with broader claims that many seemingly adverse cognitive aging effects may instead reflect an increasing knowledge base with age<sup>48,49</sup> rather than cognitive decline (although in the present study this was specific to naming as DRS scores declined with increasing age). This differs from the interpretation given in many studies that reported that TOTs increase in aging, a finding commonly viewed as evidence of cognitive decline in older age.<sup>1</sup> PR scores may have increased with age in true controls because increased age leads to greater semantic knowledge, which

increases the likelihood of knowing very low-frequency names and the chance that when retrieval fails, it fails only briefly (and resolves with just “try again” as we did not administer semantic or phonological cues).<sup>4,50</sup> In those with AD, by contrast, a greater proportion of failed retrievals reflects prelexical damage to semantic representations<sup>8</sup> that is unlikely to resolve quickly and spontaneously.

This interpretation assumes a single semantic locus for the negative effects of AD and the positive effects of normal aging on naming (differing only in the nature of semantic representations), and seems easier to accommodate within a theoretical framework that assumes mild versus moderate anomia can be explained with differences in semantic processing.<sup>51,52</sup> Note that age and first pass scores were not correlated in true controls ( $r = 0.055$ ,  $p = 0.672$ ), in preclinical AD ( $r = -0.300$ ,  $p = 0.137$ ), or in MCI/AD ( $r = 0.142$ ,  $p = 0.316$ ). However, age-related slowing<sup>53,54</sup> might have offset age-related gains in semantic knowledge and naming ability. Without time pressure older adults scored higher than young adults in picture naming tests in some studies.<sup>55,56</sup> Studies that revealed aging-related declines in naming scores might have included some participants with preclinical AD.<sup>57-59</sup>

It remains to be determined how the present findings can be accommodated within theoretical frameworks that assume aging effects on naming reflect difficulties with accessing phonology,<sup>60,61</sup> or executive control impairments.<sup>62</sup> While it seems reasonable to assume that increased age leads to richer semantic representations,<sup>63</sup> it is not clear why aging would increase the ability to access phonology only after temporarily failing to do so (in fact the opposite is typically assumed, especially for proper name retrieval<sup>5</sup>).

## 5 | CONCLUSIONS

The MINT Sprint 2.0 enables quick administration of an 80-item naming test that provides separate measures of naming success, naming speed, and ability to spontaneously resolve failed retrievals. More research is needed to replicate the test's apparent potential for detecting early AD; to evaluate conversion from biomarker-negative controls to AD; to test diagnostic power in other languages; and to better characterize how time pressure, aging, and AD jointly affect naming ability. The MINT Sprint 2.0 reveals that naming deficits emerge in the earliest preclinical stages of AD, and provides preliminary evidence for a semantic locus underlying cognitive changes in naming ability in both normal aging and AD.

## ACKNOWLEDGMENTS

We thank the participants, staff, and volunteers at the UCSD Shiley-Marcos ADRC for their ongoing commitment to the research program. This work was supported by NIH grants P30-AG062429, AG076415, NS119285, AG077915 and by an NSF BCS grant 1923065.

## CONFLICT OF INTEREST STATEMENT

Authors T.H.G., D.G., A.S., C.K., and M.M. have no conflicts of interest to declare. D.P.S. is a consultant for Aptinyx and Biogen. D.G. is a consultant for Biogen, Fujirebio, Cognition Therapeutics, Amprion, vTv

Pharmaceuticals, and GE Healthcare. Author disclosures are available in the [supporting information](#).

## CONSENT STATEMENT

The research protocol was approved by the UCSD Institutional Review Board in accordance with the Declaration of Helsinki. Informed consent was obtained at the point of entry into the longitudinal study from all patients or their caregivers consistent with California state law.

## REFERENCES

- Burke DM, MacKay DG, Worthley JS, Wade E. On the tip of the tongue: what causes word finding failures in young and older adults? *J Mem Lang*. 1991;30(5):542-579. doi:10.1016/0749-596X(91)90026-G
- Evrard M. Ageing and lexical access to common and proper names in picture naming. *Brain Lang*. 2002;81(1):174-179. doi:10.1006/brln.2001.2515
- Rastle KG, Burke DM. Priming the tip of the tongue: effects of prior processing on word retrieval in young and older adults. *J Mem Lang*. 1996;35(4):586-605. doi:10.1006/jmla.1996.0031
- Gollan TH, Brown AS. From tip-of-the-tongue (TOT) data to theoretical implications in two steps: when more TOTs means better retrieval. *J Exp Psychol Gen*. 2006;135(3):462-483. doi:10.1037/0096-3445.135.3.462
- Shafto MA, Burke DM, Stamatakis EA, Tam PP, Tyler LK. On the tip-of-the-tongue: neural correlates of increased word-finding failures in normal aging. *J Cogn Neurosci*. 2007;19(12):2060-2070. doi:10.1162/jocn.2007.19.12.2060
- White KK, Abrams L. Does priming specific syllables during tip-of-the-tongue states facilitate word retrieval in older adults? *Psychol Aging*. 2002;17(2):226-235. doi:10.1037/0882-7974.17.2.226
- Nebes RD. Semantic memory in Alzheimer's disease. *Psychol Bull*. 1989;106(3):377-394. doi:10.1037/0033-2909.106.3.377
- Hodges JR, Patterson K, Graham N, Dawson K. Naming and knowing in dementia of Alzheimer's type. *Brain Lang*. 1996;54(2):302-325. doi:10.1006/brln.1996.0077
- Delazer M, Semenza C, Reiner M, Hofer R, Benke T. Anomia for people names in DAT—evidence for semantic and post-semantic impairments. *Neuropsychologia*. 2003;41(12):1593-1598. doi:10.1016/s0028-3932
- Astell AJ, Harley TA. Tip-of-the-tongue states and lexical access in dementia. *Brain Lang*. 1996;54(2):196-215. doi:10.1006/brln.1996.0071
- Beeson PM, Holland AL, Murray LL. Naming famous people: an examination of tip-of-the-tongue phenomena in aphasia and Alzheimer's disease. *Aphasiology*. 1997;11(4-5):323-336.
- Bondi MW, Salmon DP, Galasko D, Thomas RG, Thal LJ. Neuropsychological function and apolipoprotein E genotype in the preclinical detection of Alzheimer's disease. *Psychol Aging*. 1999;14(2):295-303. doi:10.1037/0882-7974.14.2.295
- Mickes L, Wixted JT, Fennema-Notestine C, et al. Progressive impairment on neuropsychological tasks in a longitudinal study of preclinical Alzheimer's disease. *Neuropsychology*. 2007;21(6):696-705. doi:10.1037/0894-4105.21.6.696
- Belleville S, Fouquet C, Hudon C, Zomahoun HTV, Croteau J, Consortium for the Early Identification of Alzheimer's disease-Quebec. Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer's type dementia in older adults: a systematic review and meta-analysis. *Neuropsychol Rev*. 2017;27(4):328-353. doi:10.1007/s11065-017-9361-5
- Weintraub S, Wicklund AH, Salmon DP. The neuropsychological profile of Alzheimer disease. *Cold Spring Harb Perspect Med*. 2012;2(4):a006171. doi:10.1101/cshperspect.a006171
- Balthazar MLF, Yasuda CL, Pereira FRS, Bergo FPG, Cendes F, Damasceno BP. Coordinated and circumlocutory semantic naming

- errors are related to anterolateral temporal lobes in mild AD, amnesic mild cognitive impairment, and normal aging. *J Int Neuropsychol Soc.* 2010;16(6):1099-1107. doi:10.1017/S1355617710000998
17. Stasenko A, Jacobs DM, Salmon DP, Gollan TH. The multilingual naming test (MINT) as a measure of picture naming ability in Alzheimer's disease. *J Int Neuropsychol Soc.* 2019;25(8):821-833. doi:10.1017/S1355617719000560
  18. Eckerström C, Olsson E, Bjerke M, et al. A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to Dementia. *J Alzheimer's Dis.* 2013;36(3):421-431. doi:10.3233/JAD-122440
  19. Jacobs DM, Sano M, Dooneief G, Marder K, Bell KL, Stern Y. Neuropsychological detection and characterization of preclinical Alzheimer's disease. *Neurology.* 1995;45(5):957-962. doi:10.1212/wnl.45.5.957
  20. Weissberger GH, Salmon DP, Bondi MW, Gollan TH. Which neuropsychological tests predict progression to Alzheimer's disease in Hispanics? *Neuropsychology.* 2013;27(3):343-355. doi:10.1037/a0032399
  21. Garcia DL, Gollan TH. The MINT Sprint: exploring a fast administration procedure with an expanded multilingual naming test. *J Int Neuropsychol Soc.* 2022;28(8):845-861. doi:10.1017/S1355617721001004
  22. Stiver J, Staffaroni AM, Walters SM, et al. The rapid naming test: development and initial validation in typically aging adults. *Clin Neuropsychol.* 2022;36(7):1822-1843. doi:10.1080/13854046.2021.1900399
  23. Albert MS, DeKosky ST, Dickson D, et al. 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. *Alzheimers Dement.* 2011;7(3):270-279. doi:10.1016/j.jalz.2011.03.008
  24. McKhann GM, Knopman DS, Chertkow H, et al. 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. *Alzheimers Dement.* 2011;7(3):263-269. doi:10.1016/j.jalz.2011.03.005
  25. McKeith IG, Boeve BF, Dickson DW, et al. Diagnosis and management of dementia with Lewy bodies. *Neurology.* 2017;89(1):88. doi:10.1212/WNL.0000000000004058
  26. Rascofsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain.* 2011;134(Pt 9):2456-2477. doi:10.1093/brain/awr179
  27. Xiao MF, Xu D, Craig MT, et al. NPTX2 and cognitive dysfunction in Alzheimer's disease. *Elife.* 2017;6:e23798. doi:10.7554/eLife.23798
  28. Vanderstichele H, Bibl M, Engelborghs S, et al. Standardization of preanalytical aspects of cerebrospinal fluid biomarker testing for Alzheimer's disease diagnosis: a consensus paper from the Alzheimer's. *Alzheimers Dement.* 2012;8(1):65-73. doi:10.1016/j.jalz.2011.07.004
  29. Kaplow J, Vandijck M, Gray J, et al. Concordance of Lumipulse cerebrospinal fluid t-tau/A $\beta$ 42 ratio with amyloid PET status. *Alzheimers Dement.* 2020;16(1):144-152. doi:10.1002/alz.12000
  30. Lleó A, Alcolea D, Martínez-Lage P, et al. Longitudinal cerebrospinal fluid biomarker trajectories along the Alzheimer's disease continuum in the BIOMARKAPD study. *Alzheimers Dement.* 2019;15(6):742-753. doi:10.1016/j.jalz.2019.01.015
  31. Baayen RH, Piepenbrock R, Gulikers L. CELEX2 LDC96L14. Web Download. Philadelphia: Linguistic Data Consortium, 1995. doi:10.35111/GS6S-GM48
  32. Marian V, Bartolotti J, Chabal S, Shook A. CLEARPOND: cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS One.* 2012;7(8):e43230.
  33. Stasenko A, Gollan TH. Tip of the tongue after any language: reintroducing the notion of blocked retrieval. *Cognition.* 2019;193:104027. doi:10.1016/j.cognition.2019.104027
  34. Gollan TH, Weissberger GH, Runnqvist E, Montoya RI, Cera CM. Self-ratings of spoken language dominance: a multi-lingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Biling (Camb Engl).* 2012;15(3):594-615. doi:10.1017/S1366728911000332
  35. Gollan TH, Acenas LAR. What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and tagalog-English bilinguals. *J Exp Psychol Learn Mem Cogn.* 2004;30(1):246-269. doi:10.1037/0278-7393.30.1.246
  36. Gollan TH, Fennema-Notestine C, Montoya RI, Jernigan TL. The bilingual effect on Boston naming test performance. *J Int Neuropsychol Soc.* 2007;13(2):197-208. doi:10.1017/S1355617707070038
  37. Bruyer R, Brysbaert M. Combining speed and accuracy in cognitive psychology: is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica.* 2011;5(1):5-13.
  38. Morris JC, Weintraub S, Chui HC, et al. The uniform data set (UDS): clinical and cognitive variables and descriptive data from Alzheimer disease centers. *Alzheimer Dis Assoc Disord.* 2006;20(4):210-216. [https://journals.lww.com/alzheimerjournal/Fulltext/2006/10000/The\\_Uniform\\_Data\\_Set\\_UDS\\_\\_Clinical\\_and\\_Cognitive.7.aspx](https://journals.lww.com/alzheimerjournal/Fulltext/2006/10000/The_Uniform_Data_Set_UDS__Clinical_and_Cognitive.7.aspx)
  39. Weintraub S, Besser L, Dodge HH, et al. Version 3 of the Alzheimer disease centers' neuropsychological test battery in the uniform data set (UDS). *Alzheimer Dis Assoc Disord.* 2018;32(1):10-17. Accessed January-March, 2018. [https://journals.lww.com/alzheimerjournal/Fulltext/2018/01000/Version\\_3\\_of\\_the\\_Alzheimer\\_Disease\\_Centers\\_2.aspx](https://journals.lww.com/alzheimerjournal/Fulltext/2018/01000/Version_3_of_the_Alzheimer_Disease_Centers_2.aspx)
  40. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12(3):189-198.
  41. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* 2005;53(4):695-699. doi:10.1111/j.1532-5415.2005.53221.x
  42. Mattis S. Dementia rating scale: professional manual. Published online 1988.
  43. Kaplan E, Goodglass H, Weintraub S. The Boston naming test: experimental edition. Unpublished manuscript. Published online 1976.
  44. Kaplan E, Goodglass H, Weintraub S. The Boston naming test. Published online 1983.
  45. Kaplan E, Goodglass H, Weintraub S. Boston naming test. Pro-Ed. Published online 2001.
  46. Serrano C, Allegri RF, Drake M, et al. A shortened form of the Spanish Boston naming test: a useful tool for the diagnosis of Alzheimer's disease. *Rev Neurol.* 2001;33(7):624-627.
  47. Graviotto HG, Sorbara MG, Rodriguez CMT, Serrano C. 12-item version of Boston naming test: usefulness in the diagnosis of primary progressive aphasia, frontotemporal dementia, and Alzheimer's disease. *Dement Neuropsychol.* 2022;16(2):181-186. doi:10.1590/1980-5764-DN-2021-0043
  48. Ramsar M, Hendrix P, Shaoul C, Milin P, Baayen H. The myth of cognitive decline: non-linear dynamics of lifelong learning. *Top Cogn Sci.* 2014;6(1):5-42. doi:10.1111/tops.12078
  49. Verhaeghen P. Aging and vocabulary score: a meta-analysis. *Psychol Aging.* 2003;18(2):332-339. doi:10.1037/0882-7974.18.2.332
  50. Dahlgren DJ. Impact of knowledge and age on tip-of-the-tongue rates. *Exp Aging Res.* 1998;24(2):139-153. doi:10.1080/036107398244283
  51. Farrell MT, Abrams L. Tip-of-the-tongue states reveal age differences in the syllable frequency effect. *J Exp Psychol: Learn Mem Cogn.* 2011;37(1):277-285. doi:10.1037/a0021328
  52. Ralph MAL, McClelland JL, Patterson K, Galton CJ, Hodges JR. No right to speak? The relationship between object naming and semantic impairment: neuropsychological evidence and a computational model. *J Cogn Neurosci.* 2001;13(3):341-356. doi:10.1162/08989290151137395
  53. Salthouse TA. The processing-speed theory of adult age differences in cognition. *Psychol Rev.* 1996;103(3):403-428.

54. Salthouse TA. Structural models of the relations between age and measures of cognitive functioning. *Intelligence*. 2001;29(2):93-115. doi:[10.1016/S0160-2896\(00\)00040-4](https://doi.org/10.1016/S0160-2896(00)00040-4)
55. Gollan TH, Goldrick M. Aging deficits in naturalistic speech production and monitoring revealed through reading aloud. *Psychol Aging*. 2019;34:25-42.
56. Stasenko A, Kleinman D, Gollan TH. Older bilinguals reverse language dominance less than younger bilinguals: evidence for the inhibitory deficit hypothesis. *Psychol Aging*. 2021;36(7):806-821. doi:[10.1037/pag0000618](https://doi.org/10.1037/pag0000618)
57. Connor LT, Spiro A III, Obler LK, Albert ML. Change in object naming ability during adulthood. *J Gerontol Ser B*. 2004;59(5):P203-P209. doi:[10.1093/geronb/59.5.P203](https://doi.org/10.1093/geronb/59.5.P203)
58. Wen H, Dong Y. The effect of ageing on confrontation naming in healthy older adults: a three-level meta-analysis. *J Cogn Psychol*. 2023;35(4):480-508. doi:[10.1080/20445911.2023.2184745](https://doi.org/10.1080/20445911.2023.2184745). Published online March 9.
59. Verhaegen C, Poncelet M. Changes in naming and semantic abilities with aging from 50 to 90 years. *J Int Neuropsychol Soc*. 2013;19(2):119-126. doi:[10.1017/S1355617712001178](https://doi.org/10.1017/S1355617712001178)
60. Levelt WJM. *Speaking: From Intention to Articulation*. The MIT Press; 1989.
61. Levelt WJ, Roelofs A, Meyer AS. A theory of lexical access in speech production. *Behav Brain Sci*. 1999;22(1):1-38.
62. Higby E, Cahana-Amitay D, Vogel-Eyny A, Spiro A, Albert ML, Obler LK. The role of executive functions in object- and action-naming among older adults. *Exp Aging Res*. 2019;45(4):306-330. doi:[10.1080/0361073X.2019.1627492](https://doi.org/10.1080/0361073X.2019.1627492)
63. Taylor JK, Burke DM. Asymmetric aging effects on semantic and phonological processes: naming in the picture-word interference task. *Psychol Aging*. 2002;17(4):662-676. doi:[10.1037//0882-7974.17.4.662](https://doi.org/10.1037//0882-7974.17.4.662)

#### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Gollan TH, Garcia DL, Stasenko A, et al. The MINT Sprint 2.0: A picture naming test for detection of naming impairments in Alzheimer's disease and in preclinical AD. *Alzheimer's Dement*. 2024;20:112-123.  
<https://doi.org/10.1002/alz.13381>