Interindividual and intraindividual variability in amnestic mild cognitive impairment (aMCI) measured with an online cognitive assessment

Annalise A. LaPlume, PhD.[1], Theone S. E. Paterson[2, 3], PhD., Sandra Gardner, PhD.[1, 4], Kathryn A. Stokes, PhD., C.Psych[3], Morris Freedman, M.D. [1, 5, 6, 7], Brian Levine, PhD., C.Psych[1, 7, 8], Angela K. Troyer, PhD., C.Psych[3, 8], and Nicole D. Anderson, PhD, C.Psych.[1, 8, 9*]

[1] Rotman Research Institute, Baycrest (fully affiliated with the University of Toronto), Toronto, Canada; [2] Department of Psychology, University of Victoria, Victoria, Canada; [3] Neuropsychology and Cognitive Health Program, Baycrest, Toronto, Canada; [4]Biostatistics Division, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada; [5] Division of Neurology, Baycrest, Toronto, Canada; [6] Mt. Sinai Hospital, Toronto, Canada; [7] Department of Medicine (Neurology), University of Toronto, Toronto, Canada; [8] Department of Psychology, University of Toronto, Toronto, Canada; [9] Department of Psychiatry, University of Toronto, Toronto, Canada

[*] Corresponding author: Dr. Nicole D. Anderson, nanderson@research.baycrest.org

ACCEPTED VERSION of PEER-REVIEWED MANUSCRIPT:

This preprint has been peer-reviewed and accepted for publication in the Journal of Clinical and Experimental Neuropsychology, for publication in late 2021.

A link to the published article will be provided once available.

Abstract

Introduction: Mean cognitive performance is worse in amnestic mild cognitive impairment (aMCI) compared to control groups. However, studies on variability of cognitive performance in aMCI have yielded inconclusive results, with many differences in variability measures and samples from one study to another.

Methods: We examined variability in aMCI using an existing older adult sample (n=91; 51 with aMCI, 40 with normal cognition for age), measured with an online self-administered computerized cognitive assessment (Cogniciti's Brain Health Assessment). Our methodology extended past findings by using pure measures of variability (controlling for confounding effects of group performance or practice), and a clinically representative aMCI sample (reflecting the continuum of cognitive performance between normal cognition and aMCI).

Results: Between-group *t*-tests showed significantly greater between-person variability (*interindividual variability* or *diversity*) in overall cognitive performance in aMCI than controls, with a small to moderate effect size, d = 0.44. No significant group differences were found in within-person variability (*intraindividual variability*) across cognitive tasks (*dispersion*) or across trials of a response time task (*inconsistency*), which may be because we used a sample measuring the continuum of cognitive performance. Exploratory correlation analyses showed that a worse overall score was associated with greater inter- and intraindividual variability, and that variability measures were correlated with each other, indicating people with worse cognitive performance were more variable.

Discussion: The current study demonstrates that self-administered online tests can be used to remotely assess different types of variability in people at risk of Alzheimer`s. Our findings show small but significantly more interindividual differences in people with aMCI. This diversity is considered as ‘noise’ in standard assessments of mean performance, but offers an interesting and cognitively informative ‘signal’ in itself.

Keywords

Mild cognitive impairment (MCI), cognition, episodic memory, online tests, intraindividual variability, individual differences

Neurodegenerative changes from Alzheimer's disease begin many years before a formal diagnosis, often in the absence of marked cognitive or functional symptoms (Jack et al., 2013; Serrano-Pozo et al., 2011; Villemagne et al., 2013), making reversal of the disease highly unlikely (Petersen, 2003). The prodromal or pre-clinical stage of Alzheimer's disease has therefore received growing interest, with the hope of developing disease-modifying interventions. Amnestic mild cognitive impairment (aMCI) reflects a transitional phase between normal cognitive aging and dementia, most commonly due to Alzheimer's disease (Damian et al., 2014; Petersen et al., 1999). It is characterized by memory impairment and preservation of daily functioning (N. D. Anderson, 2019; Petersen, 2004; Petersen et al., 1999, 2001; Winblad et al., 2004).

People diagnosed with aMCI perform more poorly, on average, than controls with normal cognition for age on event (episodic) memory (particularly tasks of associative recognition), interference control, set shifting, and working memory (N. D. Anderson et al., 2008; Belleville et al., 2007; Bennett et al., 2006; Brandt et al., 2009; Dudas et al., 2005; Gorus et al., 2008; Irish et al., 2011; Koen & Yonelinas, 2014; Rabi et al., 2020; Saunders & Summers, 2011; Serra et al., 2010; Traykov et al., 2007; Troyer et al. 2008, 2012; Westerberg et al., 2006; Zhang et al., 2007). However, a limitation of studying mean performance is that aggregating data across people to calculate the group mean loses potentially useful information about variability between participants. Studies typically also aggregate data within each person, either across trials to create a single task score or across tasks to create a single overall/composite score, thus losing potentially useful information about variability within participants.

Variability in aMCI

In the current investigation, we look beyond the standard measure of mean group performance to examine variability within groups and individuals with aMCI. We examine variability in three ways, (1) between people (interindividual variability or *diversity*), (2) within people (intraindividual variability) across tasks (*dispersion* or cross-domain variability), and (3) within people across trials (*inconsistency* or per-trial variability; Hultsch et al., 2002; Stuss et al., 2003). Diversity is calculated by how far each person's performance on a cognitive task deviates from the group average. Dispersion is calculated by the average deviation of each person's performance across cognitive tasks. Inconsistency is calculated by the average deviation of each person's performance across trials of one cognitive task. Both dispersion and inconsistency measure intraindividual performance fluctuations, but the former does so using one measure per task (typically both response time and accuracy tasks), while the latter using trials of a single task (typically a response time task).

Diversity (Interindividual Variability)

Studies that directly measured diversity have yielded inconclusive findings. Increased diversity was found in aMCI groups compared to control groups on some measures of processing speed and accuracy, but not on other measures of the same abilities (Gorus et al., 2008; Ramratan et al., 2012).

Dispersion (Intraindividual Variability Across Tasks)

Studies have yielded inconclusive findings on group differences in dispersion between control groups and aMCI groups, and whether dispersion was associated with a higher likelihood

of being classified as having aMCI (Costa et al., 2019; Kälin et al., 2014; c.f., Halliday et al., 2018; Roalf et al., 2016). However, greater dispersion is associated with greater odds of progression from typical aging to MCI, and from MCI to Alzheimer's dementia (E. D. Anderson et al., 2016; Brewster et al., 2002; Gleason et al., 2018; Holtzer et al., 2008; Koscik et al., 2016; MacDonald et al., 2012; Roalf et al., 2016; Vaughn et al, 2013; Watermeyer et al., 2020). Differences between group studies and longitudinal progression studies on dispersion may be explained by the finding that not all people diagnosed with aMCI progress to Alzheimer's disease (Clark et al., 2012; Cloutier et al., 2015; Ganguli et al., 2011; Jak et al., 2009; Mitchell et al., 2009; Petersen et al., 2001; Roberts et al., 2014; Tabert et al., 2006; Twamley et al., 2006). A meta-analysis of studies also showed a moderate positive association across studies of elevated dispersion in mixed MCI types compared to controls (Aita, 2020).

Inconsistency (Intraindividual Variability Across Trials)

Similar to dispersion, studies have yielded inconclusive findings on group differences and diagnostic accuracy of inconsistency between control groups and aMCI groups (Chow et al., 2021; Christensen et al., 2005; Costa et al., 2019; Kay et al., 2017; Lu & Lam, 2017; McLaughlin et al., 2010; Ramratan et al., 2012; Strauss et al., 2007; Troyer et al., 2016; c.f., Gorus et al., 2008; Kay, 2017; Phillips et al., 2013; Tales et al., 2013; Tarnanas et al., 2015). As with dispersion, greater inconsistency is associated with greater odds of progression from normal aging to MCI, and from MCI to Alzheimer's dementia (Bielak et al., 2010; Haynes et al., 2017; Kochan et al., 2016; Ramratan, 2016; Tales et al., 2012), and a meta-analysis also supported a small positive association across studies between elevated inconsistency in mixed MCI types compared to controls (Aita, 2020).

Differences Between Past Studies on Variability in aMCI

In addition to differences in results between studies, reviews suggest many differences in methodology between studies on variability in aMCI (Aita, 2020; Costa et al., 2019). In particular, notable differences were observed in the measurement of variability and the type of sample used, both of which will influence the results obtained.

Measurement of Variability

Group mean performance is a major confounding variable for diversity and dispersion calculations. Greater-variance is associated with more extreme means (e.g., slower reaction times or lower accuracy), and mean-level performance is expected to differ between groups, with more extreme scores for the aMCI group. Most past studies did not account for mean performance, leaving it unclear whether existing findings of greater diversity and dispersion in aMCI are an artifact of group mean differences. For example, reaction times on the Stroop task may be more spread out in aMCI groups compared to control groups, which suggests more diversity in aMCI, but this greater diversity may be simply due to aMCI groups having slower response times rather than true group differences in variability.

Meanwhile, trial order is a major confounding variable for inconsistency calculations. Trials of the same difficulty level will show a greater task effect for trials that occur earlier than trials that occur later (as participants learn the task and adapt their strategies over time), and fatigue may influence performance on later trials. It is important to account for when a trial occurs, relative to others, when calculating variability across trials.

We use a purified residuals technique designed to control for confounding effects prior to calculating variability (Hultsch et al., 2000, 2002). Substantial individual differences in interindividual or intraindividual variability have been found in normal aging using this procedure, even after controlling for group effects of age or cognitive impairment status (e.g., Dixon et al., 2007; Hultsch et al., 2000, 2002), including on the BHA (LaPlume et al., in press). Researchers using this procedure found elevated inconsistency in aMCI (Chow et al., 2021; Troyer et al., 2016), but not dispersion (Halliday et al., 2018).

The Importance of Sampling in Understanding aMCI

Another factor contributing to prior mixed results in that studies differ in how they sample groups of aMCI and matched controls. Group difference studies treat aMCI as a relatively homogenous stage (an assumption built into the study of group means), but it actually reflects a continuum between normal cognitive aging and dementia (Freitas et al., 2013). Classification into aMCI or normal aging thus requires a binary division of a continuum of ability. Individuals with aMCI are characterized by memory impairments, but there is some overlap in cognitive performance between aMCI and normal aging groups. Non-impaired individuals who perform below average on cognitive tasks are more similar in cognitive performance to aMCI groups than to non-impaired individuals who perform at or above average on cognitive tasks (Ylikoski et al., 1999). Similarly, individuals with aMCI who do not progress to dementia show similar cognitive performance to non-impaired older adults, and show better cognitive performance than those who do progress, on episodic memory and executive function measures (Cloutier et al., 2015). This overlap in cognitive ability creates a ‘gray’ area of

performance between normal aging and aMCI (Freitas et al., 2013), and poses a difficulty for between-group comparison studies.

The difficulty of differentiating between normal aging and aMCI is also seen in the heterogeneity of outcomes in aMCI: Some people classified with aMCI at a point in time will revert to normal cognition, while others classified as having normal cognition may later be diagnosed with dementia without ever having being classified as aMCI (Clark et al., 2012; Cloutier et al., 2015; Ganguli et al., 2011; Jak et al., 2009; Mitchell et al., 2009; Petersen et al., 2001; Roberts et al., 2014; Tabert et al., 2006; Twamley et al., 2006). Heterogeneity occurs in aMCI even for groups screened to be as homogeneous as possible (Palmer et al., 2002; Halliday et al., 2018). As there are many potential factors impacting cognitive function, and whether/how frequently an individual's cognition is assessed, a diagnosis into one group or another is therefore difficult for individuals who fall into the gray area.

Understanding the gray area is crucial for improving the early detection of prodromal AD, and its differentiation from normal cognitive aging. Therefore, our sample includes older adults who were easily classified with aMCI (e.g., those with lower than expected performance on multiple memory tests) or normal cognition (e.g., those with scores in the expected range on all cognitive tests), as well as those in between who were more difficult to classify (e.g., those with less obvious memory impairments, but with scores somewhat below expectation on some tasks). We did this by accepting all individuals into the study that were functionally unimpaired and met other inclusion criteria, and by using a consensus approach to applying research-based diagnostic criteria (using gold standard neuropsychological assessment with consensus diagnosis by three clinical neuropsychologists). Borderline cases were not excluded. As a result,

recruitment was not restricted to those with extreme neuropsychological profiles (i.e., 'clean' groups of people who were easily classifiable as aMCI or controls using a neuropsychological battery), but rather was tailored to include a continuous sample that was representative of the spectrum of performance, as might be seen in a typical clinical setting. This sampling technique is not always used in aMCI research, which tends to select 'cleaner' aMCI and control groups, but is typical of clinical practice, in which all individuals are accepted.

A Self-Administered Online Cognitive Assessment

In the current study, we examined variability in cognitive performance using a self-administered computerized online test, Cogniciti's Brain Health Assessment (BHA, Troyer et al., 2014). The BHA includes online versions of four neuropsychological tasks. Tasks were specifically selected and designed to detect early cognitive impairment and to be suitable for older adults and for people with basic computer skills (Troyer et al., 2014).

Although not tested as extensively as some standard neuropsychological tasks, the BHA tasks show adequate internal consistency, test-retest reliability, alternate version reliability, and construct validity (Troyer et al., 2014), as well as adequate convergent validity when compared to clinician-administered neuropsychological tests of the same constructs (Paterson et al., 2021). The BHA also demonstrates similar diagnostic accuracy to the widely-used Montreal Cognitive Assessment (Nasreddine et al., 2005) and has good diagnostic accuracy for aMCI identified using a gold-standard neuropsychological assessment (Paterson et al., 2021).

Remote digital technology is becoming more common in clinical and experimental neuropsychology. Validation studies have shown that internet testing can have comparable

reliability to lab testing on cognitive tasks of memory and attention (Chetverikov & Upravitelev, 2016; Crump et al., 2013; de Leeuw & Motz, 2016; Enochson & Culbertson, 2015; Hilbig, 2016; Reimers & Stewart, 2007, 2015; Slote & Strand, 2015). Computerized cognitive measures have also been used to facilitate detection of subtle and early signs of Alzheimer's disease (Tarnanas et al., 2015; Ramratan et al., 2012), and the current test is one of few tests shown to be suitable for aMCI (Paterson et al., 2021). On the other hand, concerns have been expressed on the lack of a controlled testing environment, reliability of response time data due to software/hardware/device variability, accuracy of self-reported information, internet connectivity, a person's familiarity with computers, etc. (Bauer et al., 2012; Feenstra et al., 2017; Germine et al., 2019 ; Iverson et al., 2009; Miller & Barr, 2017; Parsons et al., 2018). Overall, the benefits of online testing may outweigh the concerns, but attention should be paid to maximize reliability and accuracy of the collected data (Germine et al., 2019; Miller & Barr, 2017; Reimers & Stewart, 2015). To minimize issues, the current data involves a previously validated test that was appropriate for the study sample, subjects completed the test at a memory clinic (which offered a more controlled environment, minimized distractions, removed software/hardware/device variability, and ensured internet connectivity), data were examined and cleaned for outliers (n = 2), ability to use a computer was an inclusion criterion (and one individuals who reported familiarity with computers but was unable to use a mouse/laptop was removed, n = 1), and all reported background data (e.g., age, sex) were verified with data collected in an interview with a neuropsychologist (Paterson et al., 2021).

The Current Study

In the current study, we extend the findings on mean group differences to study variability across people, tasks, and trials in aMCI. We do this by (a) using a pure measure of variability (controlling for confounding effects), and (b) using a clinically representative MCI group (including people who are difficult to classify), as discussed above. We use an existing dataset, in which individuals with aMCI had significantly lower scores on the BHA than a normal aging group, with a large effect size, $d = 0.86$ (Paterson et al., 2021). However, variability was not previously measured in this sample.

The inconclusive findings from past studies mean that our hypotheses are exploratory. Meta-analytic and longitudinal past findings indicate that we may expect greater variability on all three measures in the aMCI group compared to the control group. However, no studies have replicated these differences using both pure measures of variability and a clinically relevant sample.

Materials and Methods

Participants and Procedure

Existing data were used from a previous investigation (Paterson et al., 2021). The sample consists of 91 individuals, 51 diagnosed with aMCI and 40 age-matched controls with no cognitive impairment. A control group with no cognitive impairment was needed to ensure the internal validity of the research, by providing a baseline to ensure that findings associated with aMCI were not due to other factors. The sample was recruited from referrals for clinical neuropsychological assessment to a memory clinic or from a participant database at a hospital for

older adults (Baycrest). Participants were approached for study participation if they matched the inclusion/exclusion criteria.

Demographic characteristics and test scores for each group are reported in Table 1. Participant groups were on average in their mid-seventies, included both men and women, and were well educated. There were no significant differences between the aMCI and control group on age, education, or sex. The dependent variables were normally distributed for both groups, as indicated by normality plots and published criteria to assess normality (skew<2, kurtosis<7; West et al., 1995).

As outlined in the original report (Paterson et al., 2021), all participants received a standard clinical interview as a part of their neuropsychological assessment, which included collection of information about demographics, sensory ability, physical symptoms, cognitive concerns, medical history, medications, mood, and daily functioning. This information was considered in diagnosis, as in standard clinical neuropsychological practice. Participants were excluded if they lacked abilities needed to complete the cognitive tests adequately (visual acuity, hearing, English language proficiency, reading ability, and inability to use a computer). Participants were also excluded if they reported functional impairments, or had observable clinical features of neurodegenerative disease other than AD, a history of brain tumour, stroke, seizures, traumatic brain injury, more than two lacunar infarcts on brain imaging, current cancer, untreated sleep apnea, substance abuse history in the past six months, and other neurological or psychiatric conditions that could influence cognitive performance (such as moderate to severe depressive or anxiety symptoms). For complete details on recruitment and eligibility criteria refer to Paterson and colleagues (2021).

All participants completed a full neuropsychological assessment, the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), and Cogniciti's Brain Health Assessment (BHA) on the same day. Administration of the neuropsychological tests, MoCA, and BHA was counterbalanced across participants to control for practice or exposure effects from transfer of skill between similar tasks. The study protocol was approved by the Research Ethics Board at Baycrest (#REB# 09-02).

Neuropsychological Assessment and aMCI Diagnosis

The neuropsychological assessment was conducted by a licensed clinical neuropsychologist (TP), and the tests were administered by the neuropsychologist or by a research assistant or graduate student working under the direct supervision of the neuropsychologist. The assessment included a semi-structured clinical interview and standardized cognitive tasks. The interview was used to collect demographic information, determine presence or absence of a memory concern, and evaluate functional ability. The test battery assessed intellectual ability (Vocabulary, Matrix Reasoning; Wechsler, 1999), processing speed (Digit-Symbol Coding; Wechsler, 1997), language (Verbal Fluency, Boston Naming Test; Leach et al., 2000; Kaplan et al., 2001), visuospatial abilities (Spatial Location, Clocks; Leach et al., 2000), memory (Word Lists, Complex Figure, Logical Memory, Digit-Symbol Incidental Recall; Leach et al., 2000; Wechsler, 1987; Wechsler, 1997), and executive function (Trail Making Test, Color-Word Interference Test; Army Individual Test Battery, 1944; Delis et al., 2001). Information on functional status was obtained using the Instrumental Activities of Daily Living Questionnaire (IADL; Lawton & Brody, 1969) from all individuals, and from a

significant other where possible (52% of the total sample). The assessment lasted for three to four hours.

Three neuropsychologists (co-authors AT, KS, and BL) used the neuropsychological assessments to classify participants with aMCI or normal cognition for age by consensus. All three neuropsychologists were blinded to participants' performance on the MoCA and BHA. Diagnosis of aMCI was done using National Institute on Aging-Alzheimer's Association criteria (Albert et al., 2011), and was generally considered when objective memory impairment (defined as deficits relative to expected performance for the participant's age, education, and intellectual status) was present on at least two out of four memory tests. Participants with impairments on fewer than two tests within any cognitive domain were classified with normal cognition. The aMCI group includes individuals diagnosed with single-domain aMCI (n = 36), multiple-domain aMCI (n = 14), or who could not be distinguished between single and multiple-domain aMCI (n = 1). Those with impairments only in non-memory cognitive domains were excluded from the study as the focus was on aMCI as a precursor to Alzheimer's disease specifically.

Brain Health Assessment

Participants completed the Brain Health Assessment, a free self-administered online computerized test designed for older adults concerned about their memory (www.cognitici.com). Full details of the tasks and their development can be accessed in the original paper (Troyer et al., 2014). Participants individually completed the test in a quiet room at the testing location using a laptop and a mouse. They were instructed to follow the instructions on screen, and were unsupervised while doing the test.

The BHA involves four cognitive tests, presented in the following order: (a) a Spatial Working Memory task, measuring working memory for shapes and locations, in which participants match pairs of shapes by remembering their locations on a 4 x 3 grid over three trials, by clicking on grid tiles to uncover shapes with the goal of finding all pairs in as few clicks as possible; (b) a Face-Name Association task, measuring associative recognition, in which participants complete a recognition test for learned pairs of faces and names; (c) a number-word Stroop task, measuring interference control, in which participants count the number of words for neutral (e.g., "call call"), congruent (e.g., "two two"), and incongruent (e.g., "three three") stimuli; (d) a Letter-Number Alternation task (an online version of the Trails B task), measuring set shifting and attention control, in which participants clicked alternating numbers and letters in ascending order. Practice trials with feedback were provided for the Stroop and Letter-Number Alternation tasks. Visual examples were provided for all tasks.

A single representative measure per task was used: the total number of clicks for the Spatial Working Memory task (combining Trials one, two, and three), the median response time on the incongruent condition for the Stroop task (because it combines baseline speed and interference effects, and is more appropriate for the non-normal response time distribution), the accuracy rate for the Face-Name Association task, and the completion time for the Letter-Number Alternation task. Scores per task were converted to age-corrected z-scores based on normative data (Troyer et al., 2014). The overall BHA score was calculated as an average of the individual task z-scores.

Variability Measures

A purified residuals procedure was used following previously published guidelines, in which effects of confounding factors (group or trial position) were removed before calculating variability measures (Hultsch et al., 2002). Purified residuals scores were obtained by linear regression models (in which confounding factors were the predictor and the target variable was the dependent measure), separately for each participant, and then used to calculate variability.

Diversity across people was measured using the overall BHA score. Linear regression was run with the overall score as the dependent variable, and the group as the predictor variable, to create residual scores uncontaminated by group differences in performance. Absolute scores were calculated using the purified residuals to remove negative values, and used as the index of diversity. Low values reflect similar performance (i.e., less variability) amongst different individuals, while larger values reflect a greater range in performance across people.

Dispersion across tasks was measured using individual task scores, with a single representative score per task. Linear regressions were run with each task score as the dependent variable, and the group as predictor, to create residual scores uncontaminated by group differences in performance. Intraindividual standard deviations were calculated for each individual across the purified residuals from each task, and used as the index of dispersion. Low values reflect relatively similar performance (i.e., less variability) across tasks for an individual, while large values reflect uneven performance across tasks.

Inconsistency across trials was measured using Stroop task trials (with response times per trial). The Stroop task was used because it has the highest number of trials (90 trials in total, 30

per condition: Neutral, Incongruent, Congruent), and is a response time task (which aligns with the past literature measuring inconsistency using response time tasks, and is ideal as response time based measures are more sensitive than accuracy to variability across trials). Linear regressions were run with each trial response score as the dependent variable, and the trial number as the predictor variable, to create residual scores uncontaminated by mean performance, or practice or fatigue effects from trial order. Intraindividual standard deviations were calculated for each individual across the purified residuals from each trial, and used as the index of inconsistency. Low values reflect relatively similar performance (i.e., less variability) across trials for an individual, while large values reflect uneven performance across trials.

Analyses

Analyses and data processing were conducted with the R language and environment for statistical computing (R Core Team, 2020). Calculation of intraindividual variability was conducted with the SAS/STAT software version 15.2 and the SAS System for Windows version 9.4. Copyright © 2016 SAS Institute. Raincloud distribution plots were selected for data visualization as they provide maximal information about the data, by displaying raw jittered data (to show individual performance) as well as statistical values of mean performance, confidence intervals, and a split-half violin to display summaries of the data (Allen et al., 2021).

We did not correct for multiple comparisons because each analysis used different data (the overall score for diversity, one measure per task for dispersion, and Stroop trials for inconsistency), and tested a separate hypothesis (a group effect on diversity across people, on dispersion across tasks, and on inconsistency across trials). The separate hypotheses were selected prior to analyses. Further, multiple comparisons lead to inconsistency across studies (as

the number of tests conducted varies across studies), and the number of analyses conducted should not influence the conclusion for each analysis provided each analysis has an independent hypothesis (Cribbie, 2017).

Results

Group Differences

Welch independent-samples *t*-tests (Welch, 1947) were conducted to compare cognitive performance between age-matched control and aMCI groups. The Welch *t*-test was selected as variances were not expected to be equal between groups.

Diversity Across People

There was a significant difference in diversity on the overall score between groups, $t(89) = -2.13$, $p = .04$, and a small to moderate effect size (Table 1; Cohen, 1988), indicating that there was more variability among people in the aMCI group than the control group (Figure 1).

Dispersion Across Tasks

There was no significant difference for dispersion across the four task scores, $t(87) = -1.29$, $p = .20$, with a small effect size (Table 1), indicating that variability within a person across tasks was similar for both groups (Figure 1). Individual performance profiles showed that different people showed deviations from their typical performance (the drops in lines) on different tasks (Figure 2).

Inconsistency Across Trials

There was no significant difference for inconsistency in responses on the Stroop task, $t(83) = -0.81$, $p = .42$, with a negligible effect size (Table 1), indicating that variability within a person across trials was similar for both groups (Figure 1).

Correlation Analyses

The significant group effects on diversity (Figure 1), along with visual examination of the intraindividual variability scores (Figure 2), indicated that there were significant inter-individual differences in the current sample, with some individuals in the aMCI group more variable than others. We ran exploratory correlational analyses using Pearson correlations, to examine whether interindividual differences in overall cognitive performance were associated with variability measures, and whether variability measures were associated with each other.

Correlations with Overall Performance.

Overall Performance and Diversity Across People. Diversity and overall cognitive performance were negatively correlated in the control and aMCI groups, with large and medium effect sizes, $r(38) = -.62$, $p < .0001$, $r^2 = .38$, and $r(49) = -.43$, $p = .002$, $r^2 = .18$, respectively (Gravetter & Wallnau, 2009).

Overall Performance and Dispersion Across Tasks. Dispersion and overall cognitive performance were negatively correlated in the control and aMCI groups, with large effect sizes, $r(38) = -.60$, $p < .0001$, $r^2 = .36$, and $r(49) = -.66$, $p < .001$, $r^2 = .44$, respectively.

Overall Performance and Inconsistency Across Trials. Inconsistency and overall cognitive performance were negatively correlated in the control and aMCI groups, with large effect sizes, $r(38) = -.50$, $p < .0001$, $r^2 = .25$ and $r(49) = -.70$, $p < .0001$, $r^2 = .49$, respectively.

Correlations Among Variability Measures

Diversity Across People and Dispersion Across Tasks. Diversity and dispersion were positively correlated in the control and aMCI groups, with medium effect sizes, $r(38) = .42$, $p = .01$, $r^2 = .18$, and $r(49) = .46$, $p = .001$, $r^2 = .21$, respectively.

Diversity Across People and Inconsistency Across Trials. Diversity and inconsistency were positively correlated in the control and aMCI groups, with medium effect sizes, $r(38) = .47$, $p = .002$, $r^2 = .22$, and $r(49) = .44$, $p = .001$, $r^2 = .19$, respectively.

Dispersion Across Tasks and Inconsistency Across Trials. Dispersion and inconsistency were not correlated in the control group, $r(38) = .20$, $p = .22$, but were positively correlated in the aMCI group with a large effect size $r(49) = .57$, $p < .0001$, $r^2 = .32$.

Discussion

We examined variability in cognitive performance across people, tasks, and trials (diversity, dispersion, and inconsistency) in an existing sample measured with a digital web-based cognitive assessment designed for early detection of cognitive impairment (Paterson et al., 2021; Troyer et al., 2014). At an aggregate level, the group diagnosed with aMCI had more diversity than a group with normal cognition for their age, with a small to moderate effect size.

At an individual level, people with more ~~higher~~ interindividual variability also had more intraindividual variability across tasks and trials in both groups.

Our results suggest that clinicians should keep in mind that there is slightly more diversity in cognitive performance across people with aMCI than among those with normal cognition for their age. Knowing there is greater diversity in aMCI than in normal aging likely would make clinicians more cautious rather than more confident about applying a clinical diagnosis of aMCI. This variability highlights the continued importance of appropriate referral for more extensive neuropsychological assessment, especially in less clear-cut cases. The elevated diversity indicates that individuals with aMCI do not perform similarly as a group, which emphasizes the importance of focusing on the individual during the clinical decision process within neuropsychological practice (e.g., how far an individual's memory score differs from a baseline measure such as IQ). Some individuals may have better scores than the group average in aMCI, but may still have poorer performance on some domains relative to their individual performance on non-impaired domains.

From a research perspective, our results suggest that it is useful to look beyond the mean to examine interindividual differences when studying aMCI. The majority of the aMCI literature compares the average cognitive performance between groups. Although most studies report diversity to some extent (at the very least by displaying group standard deviations in addition to group average), it is used to control for group variability when comparing group means rather than being a variable of interest in itself. In the current study we instead compared group diversity after controlling for average group performance. While the standard approach of focusing on group means treats individual differences as error variance or 'noise', our findings

suggest that individual differences reflect a 'signal' by offering additional useful information to mean performance.

In addition to group variability, we studied intraindividual variability across tasks and trials (i.e., dispersion and inconsistency). We did not find any group differences on intraindividual variability. Instead, worse cognitive performance was associated with greater intraindividual variability, regardless of group status. Thus, intraindividual variability did not provide additional information beyond mean performance for the current study, which may be because we accounted for confounding effects on variability measures (group performance or practice), and used a clinically representative sample that measured the continuum of cognitive performance. It is also possible that elevated dispersion and inconsistency are more easily demonstrable in later stages of neurocognitive decline —group differences in dispersion and inconsistency are reliably found between normal aging or aMCI and Alzheimer's disease (E. D. Anderson et al., 2016; Bielak et al., 2010; Brewster et al., 2002; Gleason et al., 2018; Haynes et al., 2017; Holtzer et al., 2008; Kochan et al., 2016; Koscik et al., 2016; MacDonald et al., 2012; Ramratan, 2016; Roalf et al., 2016; Tales et al., 2012; Vaughn et al., 2013; Watermeyer et al., 2020), but are less reliably found between normal aging and aMCI (e.g., Halliday et al., 2018; Kälin et al., 2014; Roalf et al., 2016).

Overall, we demonstrate that online assessments can be used to measure variability in cognitive performance in normal aging and individuals with aMCI, in addition to standard measurements of mean performance. The BHA is computerized and online, making it more practical than traditional in-person neuropsychological assessments. Moreover, the BHA is self-administered, while other popular digital assessments are not (e.g., Cambridge

Neuropsychological Test Automated Battery, CANTAB, or the National Institutes of Health, NIH, toolbox; Fray et al., 1996; Gershon et al., 2010).

Connection to Past Studies on Variability in aMCI and Insights for Future Work

Long-term evidence shows heterogeneity (i.e., between-person differences) in the outcomes (and thus underlying aetiology) of people with aMCI; some go on to develop AD, others develop another form of dementia, while others revert back to normal cognition (Ganguli et al., 2011; Jak et al., 2009; Libon et al., 2010; Mitchell et al., 2009; Petersen et al., 2001; Roberts et al., 2014; Sachdev et al., 2013; Schneider et al., 2009; Tabert et al., 2006). However, while many studies have examined heterogeneity in outcomes for aMCI groups, few studies have examined heterogeneity in cognitive performance for aMCI groups. Our finding of more interindividual variability in the aMCI group than the normal cognitive aging group parallels evidence of high interindividual variability in aMCI outcomes.

The number and type of tasks used to measure variability may explain the differences in past findings. We followed current practice of measuring inconsistency using trials of a single response time task, and dispersion or diversity using different tasks (Hultsch et al., 2002). The tasks used to measure inconsistency are similar across studies, but the tasks used to measure dispersion vary across studies, with dispersion measurements ranging from two to 15 tasks, and from speeded tasks only, accuracy tasks only, or both types (Aita, 2020; Costa et al., 2019; Duchek et al., 2009; Gorus et al., 2008; Gleason et al., 2018; Kalin et al., 2014; McLaughlin et al., 2010; Vaughn et al., 2013). We followed recommendations on task selection for dispersion, namely to include a broad cognitive battery, with tasks measuring cognitive abilities shown to

differ between aMCI and control groups, a combination of speed and accuracy measures (Hilborn et al., 2009; MacDonald et al., 2012; Ownby et al., 2004).

Our findings are specific to the cognitive tasks in the online assessment used in the current study, although these were selected as they collectively screen for cognitive impairment, and are also similar to typical tasks used in neuropsychology research and clinical practice. A representative measure was used per task (Troyer et al., 2014), but we did not expect that results would vary considerably for other measures, since a meta-analysis showed no major differences on predictive accuracy from MCI to Alzheimer's disease between different measures of the same ability (e.g., delayed versus immediate recall, cued versus free recall for episodic memory; Belleville et al., 2017).

Future work can examine whether variability in cognitive performance differs between amnestic and non-amnestic MCI, single-domain and multiple-domain aMCI (e.g., see Chow et al., 2021). In the current study, although both types were present in our sample, there was not enough power to separate single- and multiple-domain aMCI.

It could be argued that the aMCI group, by definition, would be expected to be more variable over tasks. Therefore one may expect larger deviations between memory and non-memory tasks (i.e., more variability) in the aMCI group (who were primarily defined by an amnestic impairment) than the control group. However, observed mean group differences for the current sample, and findings from past studies, show that the aMCI group had poorer performance than the control group on all cognitive abilities measured, including non-memory

abilities (Belleville et al., 2007; Brandt et al., 2009; Paterson et al., 2021; Rabi et al., 2020; Saunders & Summers, 2011; Traykov et al., 2007; Zhang et al., 2007).

Strengths and Limitations

Although the BHA can be completed at home, in the current study it was completed by participants in a quiet room at the testing location, which could limit the generalizability of findings. However, experimental studies have also shown no differences in online cognitive task performance between people randomly assigned to a home or research testing location (Hilbig, 2016). Further, the obtained cognitive scores from the control group are nearly identical to very large samples of individuals in the same age range who did complete the test in their own homes (LaPlume et al., in press; Troyer et al., 2014).

The sampling methodology in the current study (of measuring the spectrum of cognitive performance between normal aging and aMCI) is representative of the challenges of diagnosis and of what aMCI looks like in clinical practice. Differences between aMCI and control groups are not always clear-cut (Cloutier et al., 2015; Freitas et al., 2013; Ylikoski et al., 1999). While it may seem 'cleaner' to measure groups at ends of this continuum in empirical investigations, or to drop people who are difficult to classify, sampling wide distributions was done to understand borderline cases (people in the 'gray area' between aMCI and normal cognition).

We could have included a third group of individuals to directly study the gray area of borderline cases, but a third group would have lowered the sample size and therefore the power to detect an effect, because the number of borderline cases were low, and power is based on the group with the smallest sample size. More importantly, we wanted to reflect the clinical reality

that not all cases are clear, and yet borderline individuals are classified into one group or another. Further, group assignment was made using holistic clinical diagnosis by professional neuropsychologists and published criteria (Albert et al., 2011), as would be the case in standard clinical practice, with consensus agreement of three neuropsychologists offering additional certainty. Although impairments on the neuropsychological tests were more subtle for borderline individuals than is often the case in empirical studies, there were still signs from the neuropsychological profile that indicated aMCI.

The operationalization of aMCI can differ between studies. As a result, there are significant differences between study samples, greater than those that would be expected from random sampling from one study to another (Aita, 2020). In the current study, published criteria were used (Albert et al., 2011), along with consensus diagnosis by three clinical neuropsychologists to increase reliability.

Effect sizes for group differences were negligible to medium. The small effect sizes mean that our sample sizes would have had limited power to detect true effects. Larger samples may have enabled us to detect these effects, but they may still be small. However, a sampling approach of ‘cleaner’ groups rather than a clinical representative sample may produce larger effect sizes from more pronounced group differences.

Measurement of variability is dependent on whether the measures used allow spread in the data. Not all tasks are designed to allow for spread in performance; they may yield skewed distributions, or show floor or ceiling effects. A strength of the diversity measurement in the current study is that all tasks had adequate normative distributions, allowing comparison of

variation across individuals and tasks (Troyer et al., 2014). Further, tasks used to calculate variability (the BHA online assessment) were separate to those used to diagnose participants (a neuropsychological assessment).

The early detection of Alzheimer's disease has received growing interest, with aMCI proposed as an ideal time-point for intervening to prevent impending decline. In the current study, we demonstrate the value of accounting for interindividual variability when interpreting cognitive tests. In particular, there is greater interindividual variability in MCI than normal cognitive aging. Digital neuropsychology, especially with self-administered assessments such as the BHA, can increase efficiency and outreach by enabling virtual assessment of populations with neuropsychiatric disorders. Our study shows that online self-administered cognitive tests can be used to study variability in people at risk of Alzheimer's disease.

Funding

This work was supported by research grants from the Centre for Aging and Brain Health Innovation (CABHI) Researcher-Clinician Partnership Program to BL and AT, Natural Sciences and Engineering Council (NSERC) of Canada (grant number RGPIN-2017-06057) to NA, and an Alzheimer Society of Canada PDF (fellowship number 20-16) to AL.

Acknowledgements

We would like to acknowledge the following individuals' contributions: Michael Meagher assisted with preparation of the CABHI grant application and provided project management. Caitlin Johnston provided project management for data collection. Brintha

Sivajohan, Komal Shaikh, Rebecca Trossman, and Rachel Downey conducted participant assessments.

Data Availability Statement

Data, code for statistical analyses, and study materials can be made available to researchers upon request. The analysis plan was pre-registered in advance (https://osf.io/k6gf2/?view_only=0820ced3b1524cd8be7c175b63d66b87).

Disclosure Statement

The authors declare no conflicts of interest.

Author Contributions CRediT Statement

AL: Conceptualization, Methodology, Data curation, Formal Analysis, Writing – Original Draft, Writing – Review and Editing, Funding Acquisition; TP: Conceptualization, Methodology, Investigation (data acquisition), Data curation, Project administration, Writing – Review and Editing; SG: Data curation, Formal Analysis, Writing – Review and Editing; KS: Investigation (data acquisition), Writing – Review and Editing; MF: Investigation, Writing – Review and Editing; BL: Investigation (data acquisition), Writing – Review and Editing, Supervision, Funding Acquisition; AT: Conceptualization, Methodology, Investigation (data acquisition), Writing – Review and Editing, Supervision, Funding Acquisition; NA: Conceptualization, Methodology, Writing – Review and Editing, Supervision, Funding Acquisition

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*(3), 270–279. https://doi.org/10.1016/j.jalz.2011.03.008

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., van Langen, J., & Kievit, R. A. (2021). Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Res*earch, *4*, 63. https://doi.org/10.12688/wellcomeopenres.15191.2

Aita, S. L. (2020). *Neuropsychological Intra-Individual Variability: Review and Meta-Analysis in Clinical Adult Samples* (Doctoral dissertation, University of South Alabama).

Anderson, E. D., Wahoske, M., Huber, M., Norton, D., Li, Z., Koscik, R. L., Umucu, E., Johnson, S. C., Jones, J., Asthana, S., Gleason, C. E., & Alzheimer's Disease Neuroimaging Initiative. (2016). Cognitive variability—A marker for incident MCI and AD: An analysis for the Alzheimer's Disease Neuroimaging Initiative. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *4*(1), 47–55. https://doi.org/10.1016/j.dadm.2016.05.003

Anderson, N. D. (2019). State of the science on mild cognitive impairment (MCI). *CNS Spectrums*, *24*(1), 78–87. https://doi.org/10.1017/S1092852918001347

Anderson, N. D., Ebert, P. L., Jennings, J. M., Grady, C. L., Cabeza, R., & Graham, S. J. (2008). Recollection- and familiarity-based memory in healthy aging and amnestic mild cognitive

impairment. *Neuropsychology*, *22*(2), 177–187. https://doi.org/10.1037/0894-4105.22.2.177

Anstey, K. J., Mack, H. A., Christensen, H., Li, S. C., Reglade-Meslin, C., Maller, J., ... & Sachdev, P. (2007). Corpus callosum size, reaction time speed and variability in mild cognitive disorders and in a normative sample. *Neuropsychologia*, *45*(8), 1911-1920. https://doi.org/10.1016/j.neuropsychologia.2006.11.020

Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized neuropsychological assessment devices: Joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Archives of Clinical Neuropsychology*, *27*(3), 362–373. https://doi.org/10.1093/arclin/acs027

Belleville, S., Chertkow, H., & Gauthier, S. (2007). Working memory and control of attention in persons with Alzheimer's disease and mild cognitive impairment. *Neuropsychology*, *21*(4), 458–469. https://doi.org/10.1037/0894-4105.21.4.458

Bennett, I. J., Golob, E. J., Parker, E. S., & Starr, A. (2006). Memory evaluation in mild cognitive impairment using recall and recognition tests. *Journal of Clinical and Experimental Neuropsychology*, *28*(8), 1408–1422. https://doi.org/10.1080/13803390500409583

Benton, A., and Hamsher, K. de. (1976). *Multilingual Aphasia Examination: Manual of instruction*. University of Iowa, Iowa City.

Bielak, A. A. M., Hultsch, D. F., Strauss, E., MacDonald, S. W. S., & Hunter, M. A. (2010). Intraindividual variability in reaction time predicts cognitive outcomes 5 years later. *Neuropsychology, 24*(6), 731–741. https://doi.org/10.1037/a0019802

Brandt, J., Aretouli, E., Neijstrom, E., Samek, J., Manning, K., Albert, M. S., & Bandeen-Roche, K. (2009). Selectivity of executive function deficits in mild cognitive impairment. *Neuropsychology*, *23*(5), 607–618. https://doi.org/10.1037/a0015851

Brewster, P., Tuokko, H., & MacDonald, S. (2012). Inter-test variability contributes independently to the five-year prediction of Alzheimer's disease in nondemented older adults. *Alzheimer's & Dementia*, *4*(8), P369.

Chetverikov, A., & Upravitelev, P. (2016). Online versus offline: The Web as a medium for response time data collection. *Behavior Research Methods, 48,* 1086–1099. https://doi.org/10.3758/s13428-015-0632-x

Chow, R., Rabi, R., Paracha, S., Vasquez, B. P., Hasher, L., Alain, C., & Anderson, N. D. (2021). Reaction time intra-individual variability reveals inhibitory deficits in single- and multiple-domain amnestic mild cognitive impairment. *Journal of Gerontology: Psychological Sciences.* Epub ahead of print. https://doi.org/10.1093/geronb/gbab051.

Christensen, Helen, Dear, K. B. G., Anstey, K. J., Parslow, R. A., Sachdev, P., & Jorm, A. F. (2005). Within-occasion intraindividual variability and preclinical diagnostic status: Is intraindividual variability an indicator of mild cognitive impairment? *Neuropsychology*, *19*(3), 309–317. https://doi.org/10.1037/0894-4105.19.3.309

Clark, L. R., Schiehser, D. M., Weissberger, G. H., Salmon, D. P., Delis, D. C., & Bondi, M. W. (2012). Specific measures of executive function predict cognitive decline in older adults. *Journal of the International Neuropsychological Society*, *18*(1), 118–127. https://doi.org/10.1017/S1355617711001524

Cloutier, S., Chertkow, H., Kergoat, M.-J., Gauthier, S., & Belleville, S. (2015). Patterns of cognitive decline prior to dementia in persons with mild cognitive impairment. *Journal of Alzheimer's Disease*, *47*(4), 901–913. https://doi.org/10.3233/JAD-142910

Cohen. J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Costa, A. S., Dogan, I., Schulz, J. B., & Reetz, K. (2019). Going beyond the mean: Intraindividual variability of cognitive performance in prodromal and early neurodegenerative disorders. *The Clinical Neuropsychologist*, *33*(2), 369–389. https://doi.org/10.1080/13854046.2018.1533587

Cribbie, R. A. (2017). Multiplicity control, school uniforms, and other perplexing debates. *Canadian Journal of Behavioural Science / Revue Canadienne des Sciences du Comportement*, *49*(3), 159–165. https://doi.org/10.1037/cbs0000075

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE*, *8*(3), e57410. https://doi.org/10.1371/journal.pone.0057410

Damian, M., Hausner, L., Jekel, K., Richter, M., Froelich, L., Almkvist, O., Boada, M., Bullock, R., De Deyn, P.P., Frisoni, G.B., Hampel, H., Jones, R.W., Kehoe, P., Lenoir, H., Minthon, L., Olde Rikkert, M.G.M., Rodriguez, G., Scheltens, P., Soininen, H., Spiru, L., Touchon, J., Tsolaki, M., Vellas, B., Verhey, F.R.J., Winblad, B., Wahlund, L.-O., Wilcock, G., & Visser, P. J. (2013). Single-domain amnestic mild cognitive impairment identified by cluster analysis predicts Alzheimer's disease in the European prospective DESCRIPA study. *Dementia and Geriatric Cognitive Disorders*, *36*(1-2), 1-19. https://doi.org/10.1159/000348354

de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods, 48,* 1–12. https://doi.org/10.3758/s13428-015-0567-2

Delis, D., Kaplan, E., and Kramer, J. (2001). *Delis-Kaplan Executive Function System: Technical Manual*. San Antonio, TX: Harcourt Assessment Company.

Dixon, R. A., Garrett, D. D., Lentz, T. L., MacDonald, S. W. S., Strauss, E., & Hultsch, D. F. (2007). Neurocognitive markers of cognitive impairment: Exploring the roles of speed and inconsistency. *Neuropsychology, 21*(3), 381–399. https://doi.org/10.1037/0894-4105.21.3.381

Dudas, R. B., Clague, F., Thompson, S. A., Graham, K. S., & Hodges, J. R. (2005). Episodic and semantic memory in mild cognitive impairment. *Neuropsychologia*, *43*(9), 1266–1276. https://doi.org/10.1016/j.neuropsychologia.2004.12.005

Feenstra, H. E. M., Vermeulen, I. E., Murre, J. M. J., & Schagen, S. B. (2017). Online cognition: Factors facilitating reliable online neuropsychological test results. The Clinical *Neuropsychologist*, *31*(1), 59–84. https://doi.org/10.1080/13854046.2016.1190405

Fray, P. J., Robbins, T. W., & Sahakian, B. J. (1996). Neuropsychiatric applications of CANTAB. *International Journal of Geriatric Psychiatry*, *11*(4), 329–336. https://doi.org/10.1002/(SICI)1099-1166(199604)11:4<329::AID-GPS453>3.0.CO;2-6

Freitas, S., Simões, M. R., Alves, L., and Santana, I. (2013). Montreal Cognitive Assessment. *Alzheimer's Disease and Associated Disorders*, *27*, 37–43. https://doi.org/doi:10.1097/WAD.0b013e3182420bfe

Ganguli, Mary, Snitz, B. E., Saxton, J. A., Chang, C.-C. H., Lee, C.-W., Vander Bilt, J., Hughes, T. F., Loewenstein, D. A., Unverzagt, F. W., & Petersen, R. C. (2011). Outcomes of mild cognitive impairment by definition: A population study. *Archives of Neurology*, *68*(6), 761–767. https://doi.org/10.1001/archneurol.2011.101

Germine, L., Reinecke, K., & Chaytor, N. S. (2019). Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *The Clinical Neuropsychologist*, *33*(2), 271–286. https://doi.org/10.1080/13854046.2018.1535662

Gershon, R. C., Cella, D., Fox, N. A., Havlik, R. J., Hendrie, H. C., & Wagster, M. V. (2010). Assessment of neurological and behavioural function: The NIH Toolbox. *The Lancet Neurology*, *80*, S2–S6. https://doi.org/10.1212/WNL.0b013e3182872e5f

Gleason, C. E., Norton, D., Anderson, E. D., Wahoske, M., Washington, D. T., Umucu, E., Koscik, R. L., Dowling, N. M., Johnson, S. C., Carlsson, C. M., & Asthana, S. (2017).

Cognitive variability predicts incident Alzheimer's disease and mild cognitive impairment comparable to a cerebrospinal fluid biomarker. *Journal of Alzheimer's Disease*, *61*(1), 79–89. https://doi.org/10.3233/JAD-170498

Gorus, E., De Raedt, R., Lambert, M., Lemper, J.-C., & Mets, T. (2008). Reaction times and performance variability in normal aging, mild cognitive impairment, and Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology*, *21*(3), 204–218. https://doi.org/10.1177/0891988708320973

Gravetter, F.J., & Wallnau, L.B. (2009). Statistics for the Behavioural Sciences (8th ed.). Wadsworth CENGAGE Learning: Belmont, CA, USA.

Halliday, D. W. R., Stawski, R. S., Cerino, E. S., DeCarlo, C. A., Grewal, K., & MacDonald, S. W. S. (2018). Intraindividual variability across neuropsychological tests: Dispersion and disengaged lifestyle increase risk for Alzheimer's disease. *Journal of Intelligence*, *6*, E12. https://doi.org/10.3390/jintelligence6010012

Haynes, B. I., Bauermeister, S., & Bunce, D. (2017). A systematic review of longitudinal associations between reaction time intraindividual variability and age-related cognitive decline or impairment, dementia, and mortality. *Journal of the International Neuropsychological Society*, *23*(5), 431-445. https://doi.org/10.1017/S1355617717000236

Hilbig, B. E. (2016). Reaction time effects in lab-versus web-based research: Experimental evidence. *Behavior Research Methods, 48,* 1718–1724. https://doi.org/10.3758/s13428-015-0678-9

Holtzer, R. (2008). Within-person across-neuropsychological test variability and incident dementia. *JAMA, 300*(7), 823–830. https://doi.org/10.1001/jama.300.7.823

Hultsch, D. F., MacDonald, S. W. S., Hunter, M. A., Levy-Bencheton, J., & Strauss, E. (2000). Intraindividual variability in cognitive performance in older adults: Comparison of adults with mild dementia, adults with arthritis, and healthy adults. *Neuropsychology, 14*(4), 588–598. https://doi-org.myaccess.library.utoronto.ca/10.1037/0894-4105.14.4.588

Hultsch, D. F., MacDonald, S. W. S., & Dixon, R. A. (2002). Variability in reaction time performance of younger and older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 57*(2), P101–P115. https://doi.org/10.1093/geronb/57.2.P101

Irish, M., Lawlor, B. A., Coen, R. F., & O'Mara, S. M. (2011). Everyday episodic memory in amnestic mild cognitive impairment: A preliminary investigation. *BMC Neuroscience, 12*(1), 80. https://doi.org/10.1186/1471-2202-12-80

Iverson, G. L., Brooks, B. L., Ashton, V. L., Johnson, L. G., & Gualtieri, C. T. (2009). Does familiarity with computers affect computerized neuropsychological test performance? *Journal of Clinical and Experimental Neuropsychology, 31*(5), 594–604. https://doi.org/10.1080/13803390802372125

Jack Jr, C. R., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., Shaw, L. M., Vemuri, P., Wiste, H. J., Weigand, S. D., Lesnick, T. G., Pankratz, V. S., Donohue, M. C., & Trojanowski, J. Q. (2013). Tracking pathophysiological processes in

Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *The Lancet Neurology*, *12*(2), 207-216. https://doi.org/10.1016/S1474-4422(12)70291-0

Jak, A. J., Bondi, M. W., Delano-Wood, L., Wierenga, C., Corey-Bloom, J., Salmon, D. P., & Delis, D. C. (2009). Quantification of five neuropsychological approaches to defining mild cognitive impairment. *The American Journal of Geriatric Psychiatry*, *17*(5), 368–375. https://doi.org/10.1097/JGP.0b013e31819431d5

Kälin, A. M., Pflüger, M., Gietl, A. F., Riese, F., Jäncke, L., Nitsch, R. M., & Hock, C. (2014). Intraindividual variability across cognitive tasks as a potential marker for prodromal Alzheimer's disease. *Frontiers in Aging Neuroscience*, *6*, 147. https://doi.org/10.3389/fnagi.2014.00147

Kaplan, E., Goodglas, H., and Weintraub, S. (2001). *The Boston Naming Test-2*. 2nd Edition. Austin, TX, USA: Pro-Ed.

Kay, C. D., Seidenberg, M., Durgerian, S., Nielson, K. A., Smith, J. C., Woodard, J. L., & Rao, S. M. (2017). Motor timing intraindividual variability in amnestic mild cognitive impairment and cognitively intact elders at genetic risk for Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, *39*(9), 866-875. https://doi.org/10.1080/13803395.2016.1273321

Kochan, N. A., Bunce, D., Pont, S., Crawford, J. D., Brodaty, H., & Sachdev, P. S. (2016). Reaction time measures predict incident dementia in community-living older adults: The Sydney Memory and Ageing Study. *The American Journal of Geriatric Psychiatry*, *24*(3), 221-231. https://doi.org/10.1016/j.jagp.2015.12.005

Koen, J. D., & Yonelinas, A. P. (2014). The effects of healthy aging, amnestic mild cognitive impairment, and Alzheimer's disease on recollection and familiarity: A meta-analytic review. *Neuropsychology Review*, *24*(3), 332–354. https://doi.org/10.1007/s11065-014-9266-5

Koscik, R. L., Berman, S. E., Clark, L. R., Mueller, K. D., Okonkwo, O. C., Gleason, C. E., Hermann, B. P., Sager, M. A., & Johnson, S. C. (2016). Intraindividual cognitive variability in middle age predicts cognitive impairment 8–10 years later: Results from the Wisconsin registry for Alzheimer's prevention. *Journal of the International Neuropsychological Society*, *22*(10), 1016–1025. https://doi.org/10.1017/S135561771600093X

LaPlume, A. A., Anderson, N. D., McKetton, L., Levine, B., & Troyer, A. K. (in press). When I'm 64: Age -related variability in over 40,000 online cognitive test takers. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences.*

Lawton, M. P., and Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist*, *9*, 179–86. https://doi.org/10.1093/geront/9.3_Part_1.179.

Leach, L., Kaplan, E., Rewilak, D., Richards, B., and Proulx, G. (2000). *Kaplan Baycrest Neurocognitive Assessment Manual.* San Antonio, TX: The Psychological Corporation.

Libon, D. J., Xie, S. X., Eppig, J., Wicas, G., Lamar, M., Lippa, C., Bettcher, B. M., Price, C. C., Giovannetti, T., Swenson, R., & Wambach, D. M. (2010). The heterogeneity of mild

cognitive impairment: A neuropsychological analysis. *Journal of the International Neuropsychological Society, 16*(1), 84–93. https://doi.org/10.1017/S1355617709990993

Lu, H., & Lam, L. C. W. (2017). Impacts of 'two-level' variability on the differential power for Montreal Cognitive Assessment (MoCA) in prodromal dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, *88*(2), 186–187. https://doi.org/10.1136/jnnp-2016-314435

MacDonald, S., Brewster, P., Laukka, E., Fratiglioni, L., & Bäckman, L. (2012). Intraindividual variability across neuropsychological tasks is associated with risk of Alzheimer's disease. *Alzheimer's & Dementia*, *4*(8), P370.

Malek-Ahmadi, M., Lu, S., Chan, Y., Perez, S. E., Chen, K., & Mufson, E. J. (2017). Cognitive domain dispersion association with Alzheimer's disease pathology. *Journal of Alzheimer's Disease*, *58*(2), 575-583. https://doi.org/10.3233/JAD-161233

McLaughlin, P. M., Borrie, M. J., & Murtha, S. J. E. (2010). Shifting efficacy, distribution of attention and controlled processing in two subtypes of mild cognitive impairment: Response time performance and intra-individual variability on a visual search task. *Neurocase*, *16*(5), 408–417. https://doi.org/10.1080/13554791003620306

Miller, J. B., & Barr, W. B. (2017). The technology crisis in neuropsychology. *Archives of Clinical Neuropsychology*, *32*(5), 541–554. https://doi.org/10.1093/arclin/acx050

Mitchell, J., Arnold, R., Dawson, K., Nestor, P. J., & Hodges, J. R. (2009). Outcome in subgroups of mild cognitive impairment (MCI) is highly predictable using a simple algorithm. *Journal of Neurology, 256*(9), 1500–1509. https://doi.org/10.1007/s00415-009-5152-0

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Parsons, T. D., McMahan, T., & Kane, R. (2018). Practice parameters facilitating adoption of advanced technologies for enhancing neuropsychological assessment paradigms. *The Clinical Neuropsychologist*, *32*(1), 16–41. https://doi.org/10.1080/13854046.2017.1337932

Paterson, T., Sivajohan, B., Gardner, S., Binns, M. A., Stokes, K. A., Freedman, M., Levine, B., & Troyer, A. K. (2021). Accuracy of a self-administered online cognitive assessment in detecting amnestic mild cognitive impairment. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, gbab097. https://doi.org/10.1093/geronb/gbab097

Petersen, R. C. (Ed.). (2003). *Mild cognitive impairment: aging to Alzheimer's disease*. Oxford University Press.

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183-194.

Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., Ritchie, K., Rossor, M., Thal, L., & Winblad, B. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology*, *58*(12), 1985–1992. https://doi.org/10.1001/archneur.58.12.1985

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, *56*(3), 303–308. https://doi.org/10.1001/archneur.56.3.303

Phillips, M., Rogers, P., Haworth, J., Bayer, A., & Tales, A. (2013). Intra-individual reaction time variability in mild cognitive impairment and Alzheimer's disease: Gender, processing load and speed factors. *PloS One*, *8*(6), e65712. https://doi.org/10.1371/journal.pone.0065712

Rabi, R., Vasquez, B. P., Alain, C., Hasher, L., Belleville, S., & Anderson, N. D. (2020). Inhibitory control deficits in individuals with amnestic mild cognitive impairment: A meta-analysis. *Neuropsychology Review*, *30*(1), 97–125. https://doi.org/10.1007/s11065-020-09428-6

Ramratan, W. S., Rabin, L. A., Wang, C., Zimmerman, M. E., Katz, M. J., Lipton, R. B., & Buschke, H. (2012). Level of recall, retrieval speed, and variability on the Cued-Recall Retrieval Speed Task (CRRST) in individuals with amnestic mild cognitive impairment. *Journal of the International Neuropsychological Society: JINS*, *18*(2), 260-268. https://doi.org/10.1017/S1355617711001664

Reimers, S., & Stewart, N. (2007). Adobe Flash as a medium for online experimentation: A test of reaction time measurement capabilities. *Behavior Research Methods*, *39*, 365–370. https://doi.org/10.3758/BF03193004

Reimers, S., & Stewart, N. (2015). Presentation and response timing accuracy in Adobe Flash and HTML5/JavaScript Web experiments. *Behavior Research Methods, 47,* 309–327. https://doi.org/10.3758/s13428-014-0471-1

Roalf, D. R., Quarmley, M., Mechanic-Hamilton, D., Wolk, D. A., Arnold, S. E., & Moberg, P. J. (2016). Within-individual variability: An index for subtle change in neurocognition in mild cognitive impairment. *Journal of Alzheimer's Disease*, *54*(1), 325–335. https://doi.org/10.3233/JAD-160259

Roberts, R. O., Knopman, D. S., Mielke, M. M., Cha, R. H., Pankratz, V. S., Christianson, T. J. H., Geda, Y. E., Boeve, B. F., Ivnik, R. J., Tangalos, E. G., Rocca, W. A., & Petersen, R. C. (2014). Higher risk of progression to dementia in mild cognitive impairment cases who revert to normal. *Neurology*, *82*(4), 317–325. https://doi.org/10.1212/WNL.0000000000000055

Rozzini, L., Chilovi, B. V., Conti, M., Bertoletti, E., Delrio, I., Trabucchi, M., & Padovani, A. (2007). Conversion of amnestic mild cognitive impairment to dementia of Alzheimer type is independent to memory deterioration. *International Journal of Geriatric Psychiatry*, *22*(12), 1217-1222. https://doi.org/10.1002/gps.1816

Sachdev, P. S., Lipnicki, D. M., Crawford, J., Reppermund, S., Kochan, N. A., Trollor, J. N., ... & Ageing Study Team. (2013). Factors predicting reversion from mild cognitive impairment to normal cognitive functioning: A population-based study. *PloS One*, *8*(3), e59649. https://doi.org/10.1371/journal.pone.0059649

Saunders, N. L. J., & Summers, M. J. (2011). Longitudinal deficits to attention, executive, and working memory in subtypes of mild cognitive impairment. *Neuropsychology*, *25*(2), 237–248. https://doi.org/10.1037/a0021134

Schneider, J. A., Arvanitakis, Z., Leurgans, S. E., & Bennett, D. A. (2009). The neuropathology of probable Alzheimer disease and mild cognitive impairment. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *66*(2), 200-208. https://doi.org/10.1002/ana.21706

Serrano-Pozo, A., Frosch, M. P., Masliah, E., & Hyman, B. T. (2011). Neuropathological alterations in Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*, *1*(1), a006189. https://doi.org/10.1101/cshperspect.a006189

Serra, L., Bozzali, M., Cercignani, M., Perri, R., Fadda, L., Caltagirone, C., & Carlesimo, G. A. (2010). Recollection and familiarity in amnesic mild cognitive impairment. *Neuropsychology*, *24*(3), 316–326. https://doi.org/10.1037/a0017654

Slote, J., & Strand, J. F. (2016). Conducting spoken word recognition research online: Validation and a new timing method. *Behavior Research Methods*, *48,* 553–566. https://doi.org/10.3758/s13428-015-0599-7

Strauss, E., MacDonald, S. W. S., Hunter, M. A., Moll, A., & Hultsch, D. F. (2002). Intraindividual variability in cognitive performance in three groups of older adults: Cross-domain links to physical status and self-perceived affect and beliefs. *Journal of the International Neurological Society*, *8*, 893–906.\

Stuss, D. T., Murphy, K. J., Binns, M. A., & Alexander, M. P. (2003). Staying on the job: the frontal lobes control individual performance variability, *Brain*, *126*(11), 2363–2380. https://doi.org/10.1093/brain/awg237

Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., Zamora, D., Goodkind, M., Bell, K., Stern, Y., & Devanand, D. P. (2006). Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Archives of General Psychiatry*, *63*, 916–924. https://doi.org/10.1001/archpsyc.63.8.916

Tales, A., Leonards, U., Bompas, A., Snowden, R. J., Philips, M., Porter, G., Haworth, J., Wilcock, G., & Bayer, A. (2012). Intra-individual reaction time variability in amnestic mild cognitive impairment: A precursor to dementia? *Journal of Alzheimer's Disease*, *32*(2), 457–466. https://doi.org/10.3233/JAD-2012-120505

Tarnanas, I., Papagiannopoulos, S., Kazis, D., Wiederhold, M., Widerhold, B., & Tsolaki, M. (2015). Reliability of a novel serious game using dual-task gait profiles to early characterize aMCI. *Frontiers in Aging Neuroscience*, *7*, 1–15. https://doi.org/doi:10.3389/fnagi.2015.00050

Traykov, L., Raoux, N., Latour, F., Gallo, L., Hanon, O., Baudic, S., Bayle, C., Wenisch, E., Remy, P., & Rigaud, A.-S. (2007). Executive functions deficit in mild cognitive impairment. *Cognitive and Behavioral Neurology*, *20*(4), 219–224. https://doi.org/10.1097/WNN.0b013e31815e6254

Troyer, A. K., Murphy, K. J., Anderson, N. D., Craik, F. I. M., Moscovitch, M., Maione, A., & Gao, F. (2012). Associative recognition in mild cognitive impairment: Relationship to

hippocampal volume and apolipoprotein E. *Neuropsychologia*, *50*(14), 3721–3728. https://doi.org/10.1016/j.neuropsychologia.2012.10.018

Troyer, A. K., Murphy, K. J., Anderson, N. D., Hayman-Abello, B. A., Craik, F. I. M., & Moscovitch, M. (2008). Item and associative memory in amnestic mild cognitive impairment: Performance on standardized memory tests. *Neuropsychology*, *22*(1), 10–16. https://doi.org/10.1037/0894-4105.22.1.10

Troyer, A. K., Rowe, G., Murphy, K. J., Levine, B., Leach, L., & Hasher, L. (2014). Development and evaluation of a self-administered on-line test of memory and attention for middle-aged and older adults. *Frontiers in Aging Neuroscience*, *6*. https://doi.org/10.3389/fnagi.2014.00335

Troyer, A. K., Vandermorris, S., & Murphy, K. J. (2016). Intraindividual variability in performance on associative memory tasks is elevated in amnestic mild cognitive impairment. *Neuropsychologia*, *90*, 110–116. https://doi.org/10.1016/j.neuropsychologia.2016.06.011

Twamley, E. W., Ropacki, S. A. L., & Bondi, M. W. (2006). Neuropsychological and neuroimaging changes in preclinical Alzheimer's disease. *Journal of the International Neuropsychological Society*, *12*(5), 707–735. https://doi.org/10.1017/S1355617706060863

Vaughan, L., Leng, I., Dagenbach, D., Resnick, S. M., Rapp, S. R., Jennings, J. M., Brunner, R. L., Simpson, S. L., Beavers, D. P., Coker, L. H., Gaussoin, S. A., Sink, K. M., & Espeland, M. A. (2013). Intraindividual variability in domain-specific cognition and risk

of mild cognitive impairment and dementia. *Current Gerontology and Geriatrics Research, 2013*, 1–10. https://doi.org/10.1155/2013/495793

Villemagne, V. L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K. A., Salvado, O., Szoeke, C., Macaulay, S. L., Martins, R., Maruff, P., Ames, D., Rowe, C. C., & Masters, C. L. (2013). Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *The Lancet Neurology*, *12*(4), 357-367. https://doi.org/10.1016/S1474-4422(13)70044-9

Watermeyer, T., Goerdten, J., Johansson, B., & Muniz-Terrera, G. (2020). Cognitive dispersion and ApoEe4 genotype predict dementia diagnosis in 8-year follow-up of the oldest-old. *Age and Ageing*, afaa232. https://doi-org /10.1093/ageing/afaa232

Wechsler, D. (1987). *Manual for the Wechsler Memory Scale - Revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale*. 3rd Edition. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.

Welch, B. L. (1947). The generalization of Student's problem when several different population variances are involved. *Biometrika* 34(1/2): 28–35. https://doi.org/10.2307/2332510

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Sage Publications, Inc.

Westerberg, C. E., Paller, K. A., Weintraub, S., Mesulam, M.-M., Holdstock, J. S., Mayes, A. R., & Reber, P. J. (2006). When memory does not fail: Familiarity-based recognition in mild cognitive impairment and Alzheimer's disease. *Neuropsychology*, *20*(2), 193–205. https://doi.org/10.1037/0894-4105.20.2.193

Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org.

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.-O., Nordberg, A., Backman, L., Albert, M., Almkvist, O., Arai, H., Basun, H., Blennow, K., de Leon, M., DeCarli, C., Erkinjuntti, T., Giacobini, E., Graff, C., Hardy, J., … Petersen, R. C. (2004). Mild cognitive impairment - beyond controversies, towards a consensus: Report of the International Working Group on mild cognitive impairment. *Journal of Internal Medicine*, *256*(3), 240–246. https://doi.org/10.1111/j.1365-2796.2004.01380.x

Ylikoski, R., Ylikoski, A., Keskivaara, P., Tilvis, R., Sulkava, R., & Erkinjuntti, T. (1999). Heterogeneity of cognitive profiles in aging: Successful aging, normal aging, and individuals at risk for cognitive decline. *European Journal of Neurology*, *6*, 645–652.

Zhang, Y., Han, B., Verhaeghen, P., & Nilsson, L.-G. (2007). Executive functioning in older adults with mild cognitive impairment: MCI has effects on planning, but not on

inhibition. *Aging, Neuropsychology, and Cognition*, *14*(6), 557–570. https://doi.org/10.1080/13825580600788118

Table 1. Demographic characteristics and test scores for each group.

Significant group differences (p<.05) are shown in bold.

Participant characteristics	**Controls** (n=40)	**aMCI** (n=51)	**Group differences** (d)
Age	74 (7.0)	75 (5.7)	-0.18
Sex (n, Female)	21 (53%)	25 (49%)	0.01[a]
Education (Years of Education)	16 (2.5)	16 (2.8)	0.14
MoCA Total score	**25 (2.3)**	**24 (2.6)**	**0.61***
BHA tests			
Face Name Association (z-score)	**0.40 (1.0)**	**-0.63 (1.0)**	**1.01***
Letter-Number Alternation (z-score)	0.06 (1.2)	-0.03 (1.0)	0.09
Spatial Working Memory (z-score)	**-0.02 (1.1)**	**-1.06 (1.6)**	**0.77***
Stroop (z-score)	**0.12 (1.2)**	**-0.72 (2.1)**	**0.48***
BHA Overall score (z-score)	**0.14 (0.7)**	**-0.61 (1.0)**	**0.86***
BHA variability			
Diversity across people (absolute residual z-scores)	**0.53 (0.5)**	**0.78 (0.6)**	**0.44***
Dispersion across tasks (ISD of residual z-scores)	0.86 (0.5)	1.03 (0.7)	0.27
Inconsistency across trials (residual response times, ms)	239 (160)	266 (156)	0.17

***** *p < .05*

Note. [a] All group differences were calculate with a between-groups t-test and Cohen's d to measure effect size, except for sex, which was a categorical variable and so was calculated using a chi-square test and Cramer's V to measure effect size.

Note. Displayed values show means (SDs), unless otherwise specified. ISD=intraindividual standard deviation.

Figure 1. Raincloud plots of mean performance, diversity, dispersion, and inconsistency per group.

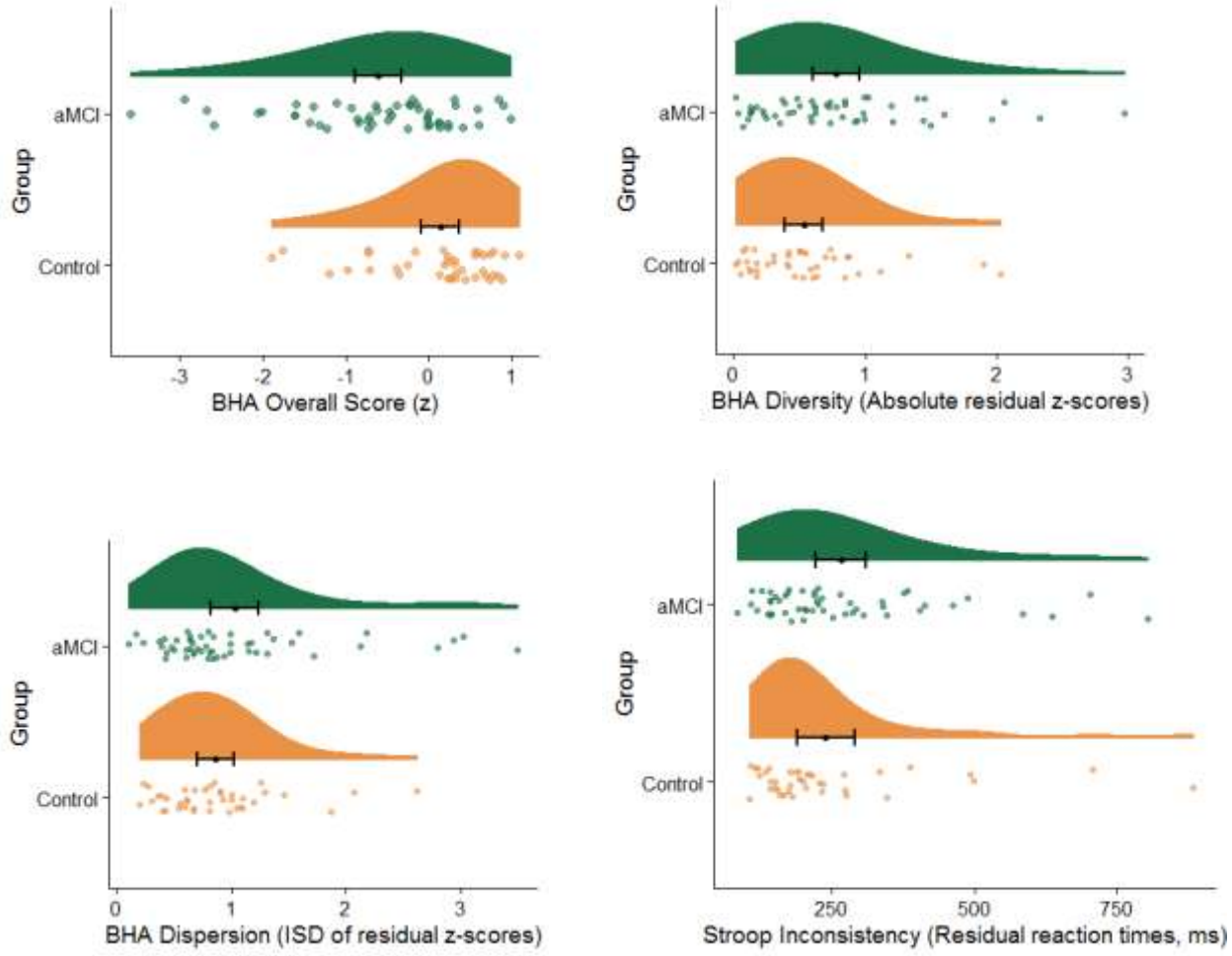

Figure 2. Line plots of individual performance across tasks.

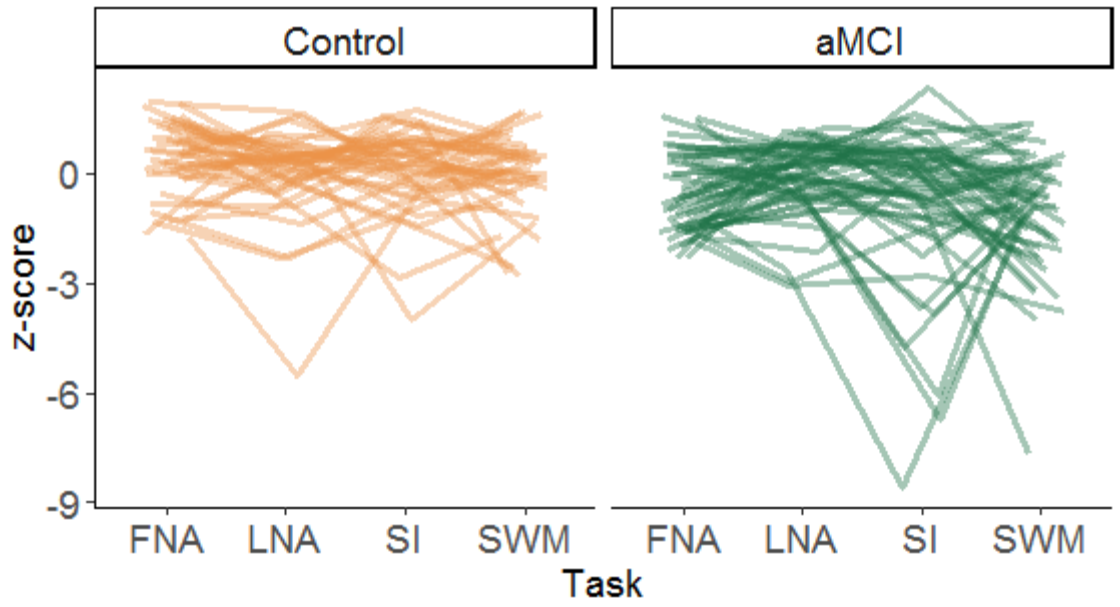

Figure 3. Scatterplots comparing groups on mean performance, diversity, dispersion, and inconsistency.

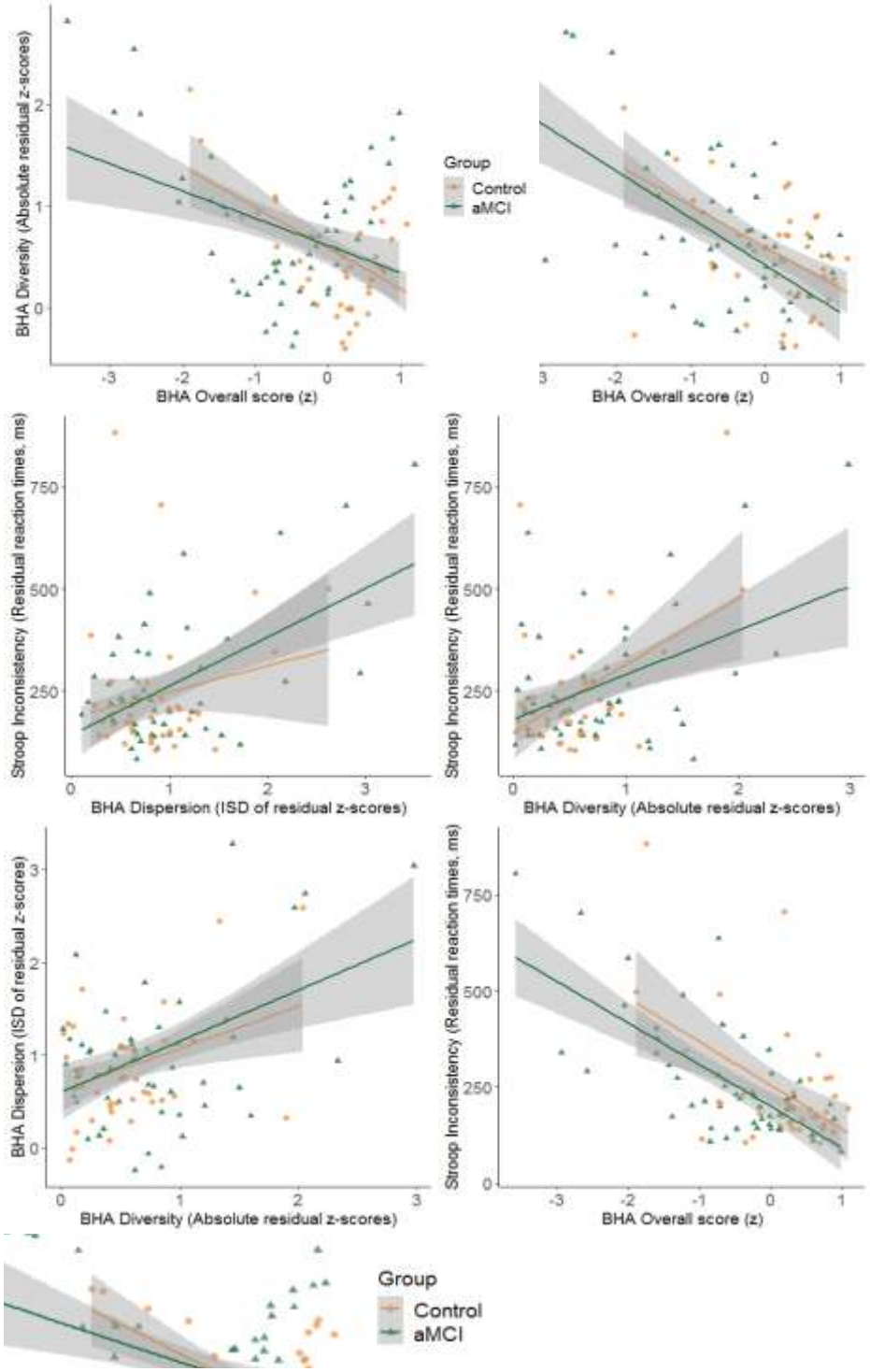

Figure 1. Raincloud plots of mean performance, diversity, dispersion, and inconsistency per group.

Note. Plots show group means (black dot), 95% confidence intervals (black error bars), and distributions (split-half violin), plus individual performance (dotplots).

Figure 2. Line plots of individual performance across tasks.

Note. Plots show a line per individual, connecting their performance on each task: Face Name Association (FNA), Letter Number Alternation (LNA), Stroop Interference (SI), and Spatial Working Memory (SWM).

Figure 3. Scatterplots comparing groups on mean performance, diversity, dispersion, and inconsistency.

Note. Plots show individual performance (points) and line of best fit.