Types of Intrusions in Verbal Fluency Tasks in Mild Cognitive Impairment and Dementia: A Longitudinal Analysis

by

Layaly Shihadeh

A Thesis Submitted to the Faculty of

Charles E. Schmidt College of Science

In Partial Fulfillment of the Requirements for the Degree of

Master of Arts

Florida Atlantic University

Boca Raton, FL

August 2022

Types of Intrusions in Verbal Fluency Tasks in Mild Cognitive Impairment and Dementia: A Longitudinal Analysis

by

Layaly Shihadeh

This thesis was prepared under the direction of the candidate's thesis advisor, Dr. Monica Rosselli, Department of Psychology, and has been approved by all members of the supervisory committee. It was submitted to the faculty of the Charles E. Schmidt College of Science and was accepted in partial fulfillment of the requirements for the degree of Master of Arts.

SUPERVISORY COMMITTEE:

MONICA ROSSELLI
MONICA ROSSELLI (Aug 11, 2022 17:09 EDT)

Monica Rosselli, Ph.D.
Thesis Advisor

John D. Morris
John D. Morris (Aug 14, 2022 20:24 EDT)

John Morris, Ph.D.

Alan Kersten, Ph.D.

Robin Vallacher

Robin Vallacher, Ph.D.
Interim Chair, Department of Psychology

Valery Forbes
Valery Forbes (Aug 15, 2022 12:21 EDT)

Valery E. Forbes, Ph.D.
Dean, Charles E. Schmidt College of Science

Robert W. Stackman Jr., Ph.D.
Dean, Graduate College

August 15, 2022
Date

Acknowledgements

I would like to express my utmost gratitude to my mentor and advisor, Dr. Monica Rosselli, for her continued support, patience, and knowledge offered throughout my graduate journey thus far. I would like to credit the rest of my committee, Dr. Alan Kersten and Dr. John Morris, for their encouragement, insightful comments, and challenging questions as well.

Furthermore, thank you to my Neuropsychology Lab team for the stimulating discussions and sleepless nights working together on deadlines. I did not just gain colleagues but lasting friendships as well.

To my family and friends, thank you for your endless supply of love and support. You all know who you are and I am truly grateful for you.

Lastly, words cannot properly express my gratitude to my mother, who inspired me to take a leap into my graduate journey and taught me with knowledge there is growth.

Abstract

Author: Layaly Shihadeh

Title: Types of Intrusions in Verbal Fluency Tasks in Mild Cognitive Impairment and Dementia: A Longitudinal Analysis

Institution: Florida Atlantic University

Thesis Advisor: Dr. Monica Rosselli

Degree: Master of Arts

Year: 2022

Research shows intrusions in memory tests can predict cognitive impairment in abnormal aging. However, there still is a need for additional research regarding the association of intrusions in verbal fluency tasks and clinical diagnosis of mild cognitive impairment, and dementia. The aim of my master's thesis research is to determine if there is an association between intrusion totals in verbal fluency tasks and diagnosis, longitudinally (across 3 years), if there are significant differences between category and phonemic fluency tasks in intrusion total scores, and if progression from CN to MCI or dementia and from MCI to dementia can be indicated through differences in intrusion scores. Results indicated that intrusions are significantly associated with diagnosis in Phonemic fluency tasks, however this was not the case in progressors versus non-progressors.

Keywords: verbal fluency, intrusions, progressors, mild cognitive impairment, dementia

Types of Intrusions in Verbal Fluency Tasks in Mild Cognitive Impairment and Dementia: A Longitudinal Analysis

List of Tables

Introduction

Cognitive Diagnoses

While memory can broaden one's overall knowledge, it can also be largely deceiving, unreliable, and flawed. Although information is retrieved and stored, unintended interruptions, or intrusions, in cognitive processes may occur. Intrusions are commonly characterized as the retrieval of words or images that are not among the original subject matter presented in an exercise or event, such as saying 'cup' when listing words in the category of 'fruits' (Haberlandt, 2011; APA). As we age, cognitive functions (e.g., inhibition of prepotent responses or executive functions) begin to slowly decline making someone more susceptible to making an intrusion (Borella et al., 2017). Intrusion errors on verbal cognitive tasks reflect subtle cognitive impairment in preclinical stages prior to the progression to dementia and have been identified as early cognitive markers of prodromal Alzheimer's disease (AD) and Dementia (Loewenstein et al., 2018; Torres et al., 2019).

The DSM-IV defines dementia as the appearance of deficits in cognition among one or more cognitive domains, such as executive functioning, learning and memory, language, etc.

Dementia is also characterized by a significant reduction in the independent participation in routine-activities, as well as possible behavioral issues, such as paranoia, hallucinations, and communication interferences. Dementia includes many etiological subtypes such as vascular dementia, Lewy body dementia, and AD. Dementia is

diagnosed using diagnostic criteria using the DSM-IV, as well as clinical neuropsychological assessments that test varying cognitive domains that may be affected by the disease (DSM-5; American Psychiatric Association (APA), 2013). Neurodegenerative diseases, like dementia, abnormally impact the aging brain heightening the rate of cognitive decline. As of 2021, AD affects approximately 6.2 million people in the United States, and nearly 24 million people worldwide (Mayeux & Stern, 2012; Maytas et al., 2019; 2021 Alzheimer's disease facts and figures, 2021). Average AD survival rate post-diagnosis is typically 5 to 8 years (Paradise et al., 2009), making early detection of dementia vital in maximizing the survivability of the disease. AD is progressive and takes effect typically through two stages prior to clinical AD diagnosis. The pre-clinical stage of AD, often regarded as the 'silent' stage, is when AD symptomology has not yet manifested (Dubois et al., 2019), but is accumulating in the brain in terms of β-amyloid AD biomarkers (Buchhave et al., 2012). Then there is the mild cognitive impairment (MCI) stage of dementia, which is characterized as the bridge between cognitive function and the more severe behavioral and cognitive effects of dementia (Roberts & Knopman, 2014).

Individuals with MCI may experience significant impairment, however daily routine and activities are typically maintained (Roberts & Knopman, 2014). MCI subtypes (e.g., non-amnestic versus amnestic MCI) differ in that amnestic MCI is considered to present clinical manifestations with memory impairment, whereas non-amnestic MCI presents with significant impairments without any presence of memory deterioration (Csukly et al., 2016). MCI can be unpredictable because of the large between- and within-person variations in disease improvement and stability. It is thought that prevention of further AD progression could be maximized during these pre-clinical

AD stages. Prior research on intrusion errors, using verbal recall and verbal list learning tasks, across cognitive groups has suggested that intrusions are valuable as measures to investigate early indications of cognitive decline (Borella et al., 2017; Thomas et al., 2019; Torres et al., 2019).

Thomas et al. (2019) explored the detection of preclinical AD through intrusion error scores using the Auditory Verbal Learning Test (AVLT). While the effects were subtle, significant differences in intrusion errors among cognitively normal and MCI groups emerged. This finding implied that intrusions can provide an additional measure for pre-clinical AD detection at its earlier stages (Thomas et al., 2019). Additionally, inhibitory deficits have been explored between cognitively normal participants and individuals who fit MCI criteria (Borella et al., 2017). Inhibition were assessed with the Color-Word Stroop task, Proactive Interference task, and a distractor task. Intrusion errors on the Proactive Interference task, which included three lists to recall, were quantified by subtracting the words correctly recalled in list 2 from the number of words correctly recalled in list 1 and dividing that number by the words correctly recalled in list 1. This quantification method were done again for list 3, as well. The MCI group produced significantly more intrusion errors than the cognitively normal group (Borella et al., 2017). Other error types (e.g., competing list intrusions, semantically related intrusions, and unrelated intrusions) have been investigated using the Loewenstein-Acevedo Scales of Semantic Interference and Learning (LASSI-L) task assessed across cognitively normal, amnestic MCI, and dementia groups (Torres et al., 2019). Significant differences emerged among the cognitively normal and the aMCI groups regarding number of competing list and semantically related intrusions, with the aMCI group producing significantly more intrusions in both error types (Torres et al., 2019). These

studies provided insight into the influence of cognitive decline on intrusion output, however it is important to note that researchers used tasks that did not measure verbal fluency directly, rather they assessed verbal recall. Due to the difference in assessments used, intrusions will be categorized differently within the verbal fluency tasks in this study compared to previous studies that have also categorized intrusions (Torres et al., 2019).

Verbal Fluency Tasks

Verbal fluency tasks are commonly used to evaluate verbal ability and executive functioning (Diamond et al., 2013; Shao et al., 2014). Verbal fluency performance has been positively associated with frontal and temporal lobe functioning (Fama et al., 2000; Kitabayashi et al., 2001). Indications of cognitive impairment on verbal fluency tasks include total of words correct, number rule violations, and percentage of perseverations. Data collected from the Wisconsin Registry for Alzheimer's Prevention (WRAP) includes differences in verbal fluency in adults that met criteria for amnestic MCI, as well as cognitively healthy adults revealing significantly lower verbal fluency scores in amnestic MCI participants than those classified as cognitively healthy (Mueller et al., 2015). Additionally, a meta-analysis of 153 studies investigating decline in verbal fluency (i.e., semantic and phonemic verbal fluency scores) among individuals experiencing cognitive impairment found patients with AD displayed significantly lower scores in semantic and phonemic fluency tasks compared to the healthy controls (Henry et al., 2004). McDonnell et al. (2020) evaluated the validity of semantic and phonemic fluency tasks as tools to distinguish cognitively normal, aMCI, and AD groups and found semantic and phonemic fluency task scores significantly differed among groups, with

aMCI and AD groups producing fewer correct words. Interestingly, semantic task scores produced more significant differences between groups than phonemic fluency task scores overall. These results add to the validity of verbal fluency tasks as an effective tool for measuring cognitive differences across cognitively normal and AD and Dementia related groups. Intrusions have not been heavily explored, especially with regard to verbal fluency tasks. Prior studies investigating the validity of both semantic and phonemic fluency tasks among individuals with AD and normal controls found that category fluency, an item of verbal fluency, had the highest rate of specificity in distinguishing patients with AD and healthy controls (Henry et al. 2004; Monsch et al., 1992). Kitabayashi et al. (2001) measured cerebral blood flow in individuals with AD using single photon emission computed tomography scans (SPECT) and administered category and phonemic fluency tests and found phonemic fluency scores were significantly correlated with cerebral blood flow in the left prefrontal lobe region and category fluency scores were significantly correlated with cerebral blood flow in the left temporal lobe region. Brain structure using structural Magnetic Resonance Imaging (MRI) and verbal fluency tasks in AD patients has revealed that semantic fluency were significantly correlated with frontal and temporal gray matter volumes (Fama et al., 2000).

Intrusion errors

Due to the minimal research conducted regarding intrusion errors and intrusion error types within verbal fluency tasks, there were not much guidance provided in terms of what classifications may be best used to measure intrusions within category and phonemic verbal fluency tasks. During the data transcribing process, two category fluency intrusion error types and three phonemic fluency intrusion error types were

commonly observed. The category fluency intrusion error types include unrelated intrusions, which are words that do not belong in the given category. The three phonemic fluency tasks include the following intrusion errors: unrelated intrusions, as well as phonemic intrusions, which are words that may have similar sounding phonemes to the letter given. These intrusion error types will be measured in this study.

While there is evidence suggesting intrusions may be a key indicator for cognitive decline (Borella et al., 2017; Thomas et al., 2019; Torres et al., 2019), minimal research has been conducted examining associations between intrusions in verbal fluency tasks and cognitive diagnosis. Binetti et al. (1995) evaluated semantic fluency performance, measured using an Animal category task, across normal elderly controls, mild AD, and moderate to severe AD patient groups. Intrusions were classified as words that did not belong in the presented category and were quantified as the sum of intrusions produced; no control variables were indicated (Binetti et al., 1995). Groups did not significantly differ in intrusions produced. Semantic verbal fluency intrusions have also been investigated longitudinally among nondemented participants and AD diagnosed patients, and intrusions were as any extra category words produced (e.g., if the word "rose" were produced in the "Animals" category) and used percentages of intrusion totals out of number of words produced to measure rate of intrusions (Auriacombe et al., 2006). Longitudinal data were separated into 5 time points: baseline, follow up at 3 years, 5 years, 8 years, and 10 years, and significant differences emerged in intrusions between nondemented and AD groups only at the 8-year timepoint (Auriacombe et al., 2006). Additionally, cognitively normal controls and MCI patients have been compared on verbal fluency tests across a 4-year time frame (Clark et al., 2016). This work tested novel verbal fluency scores, including intrusions, as a measure for predicting the

conversion of MCI to AD (Clark et al., 2016). Intrusions were scored by the count of nonvalid items and results revealed no significant differences between cognitive diagnostic groups (Clark et al., 2016).

Prior research shows intrusions in memory tests can predict cognitive impairment in abnormal aging (Borella et al., 2017; Thomas et al., 2019; Torres et al., 2019; Lowenstein et al, 2018). However, there still is a need for additional research regarding the association of intrusions in verbal fluency tasks and the clinical diagnosis of mild cognitive impairment and dementia. The aim of this study is to examine differences in intrusions on verbal fluency tasks across cognitive diagnostic groups and observe if these differences could provide early indications of cognitive decline and AD. It is predicted that (1) participants that are diagnosed with MCI or dementia at year 1 will produce a higher rate of intrusion error totals and percentages of all types on both phonemic and category verbal fluency tasks at year 1 but particularly higher on the category tasks compared to CN participants. (2) Participants classified as cognitively normal (CN) at year 1 that progress to MCI or dementia at years 2 or 3, are predicted to experience a significantly higher rate of intrusion totals and percentages on verbal fluency tasks compared to participants who remained cognitively stable. (3) Regression models using intrusion type totals and percentages at year 1 as independent variables will predict progression from CN to MCI and from MCI to dementia after controlling for demographic variables. This study could contribute to research on early detection and treatment of AD through verbal fluency intrusions, can indicate how they may relate to cognitive diagnosis, and can highlight the role intrusions play in the manifestation of AD progression.

Methods

Participants

Participants were recruited from the Memory Disorders Center at Wein Center for Alzheimer's Disease and Memory Disorders at Mount Sinai Medical Center in Miami, Florida to take part in the ongoing 1Florida Alzheimer's Disease Research Center (ADRC) project. The ADRC is a study that assesses the cognitive and neurological effects of mild cognitive impairment and dementia.

As shown in Table 1, a total of 375 participants were diagnosed every visit for three years using Global Clinical Dementia Rating Scale (CDR-Global) score criteria and grouped as cognitively normal (CN), with mild cognitive impairment (MCI), or with dementia. At baseline, participants had an average of 15.00 years of education (M = 15.00, SD = 3.65), were an average of 72.24 years old (M = 72.24, SD = 7.99), and were 62.93% female. At baseline (Visit 1), there were 88 CN, 229 MCI, and 58 dementia participants. The diagnostic groups variable were significantly associated with sex, age, and education, therefore sex, age, and education will be used as a covariates in the analyses including diagnostic groups. Females were more likely to be classified into the MCI or dementia groups compared to their male counterparts. The higher the age or the participant, the more likely to be classified as with MCI or dementia. The higher the years of education of a participant, the less likely to be classified as with MCI or dementia.

As shown in Table 2, a total of 269 participants were classified as progressors (n = 66) or non-progressors (n = 203). Progressors were considered to be participants that progressed from cognitively normal at Visit 1 (baseline) to MCI or dementia at Visits 2 or 3. Participants were also considered progressors if they progressed from MCI at Visit 1 to Dementia at Visits 2 or 3. Participants with diagnoses that remained the same across all three visits were classified as non-progressors. The progressor versus non-progressor variable were significantly associated with education, therefore education will be used as a covariate in the progressor versus non-progressor analyses.

The CN group were identified as having no memory complaints, cognitive or functional decline, and a CDR-Global Score (CDR-GS) of 0. Participants received a diagnosis of mild cognitive impairment (MCI) at the baseline assessment if they met Petersen's criteria for MCI (Petersen et al., 2014) and demonstrated all of the following: a) subjective cognitive complaints by the participant and/or collateral informant; b) evidence by clinical evaluation or history of memory or other cognitive decline; c) a CDR-GS of 0.5, and d) task performance of 1.5 SD below the means of matched norms for participants of the same age, education, and language on one or more cognitive measures. Finally, criteria for dementia included neuropsychological task performance of at least 2 SD below the mean in two cognitive domains (one memory and one non-memory task) using age, education, and language-matched normative data, as well as a CDR-GS score of 1.0 or greater. These participants also met DSM-5 criteria for Major Neurocognitive Disorder, including functional impairment sufficient to interfere with the performance of instrumental activities of daily living (IADLs).

Analyses of demographic variables (i.e., age, sex, and education) across diagnostic groups were performed to determine the use of covariates in further analyses.

Materials and Procedures

The Verbal Fluency Test consists of a categorical and a phonemic fluency section. During the categorical fluency task, participants were given a category (e.g., animals, fruits, vegetables) and were asked to name as many words that belong to the specified category in a 60 second time period. During the phonemic fluency task, also referred to as the Controlled Oral Word Association Task (COWAT), participants were given a letter (e.g., F, L, A, S) and were asked to give as many words as they can that begin with the specified letter within 60 seconds (Benton & Hamsher, 1976). Verbal responses were unlimited for each sub-section of each task, so long as the responses were produced within the given 60 seconds.

Verbal fluency tasks were either evaluated in Spanish, English, or both languages depending on the subject's primary language. The ADRC verbal fluency data available will include the number of correct responses produced for each section and sub-section, however the words produced weren't transcribed into a dataset, classified by intrusion type, or put into totals. Responses were transcribed into electronic text files and all responses were included, regardless of error. Recorded responses for each participant were organized under the category and phonemic fluency tasks' respective subsections.

Due to evidence of neuropsychological differences (Henry et al., 2004; Monsch et al., 1992) and neurobiological variations (Fama et al., 2000; Kitabayashi et al., 2001) in the brain across category and phonemic fluency tasks, intrusions in category and phonemic fluency tasks were classified, scored, and analyzed separately. Intrusion totals were quantified as the sum of intrusions within each subsection of the tasks (i.e., fruits, vegetables, and animals for category; F, L, A, S for phonemic). Intrusion percentages

were quantified as the sum of intrusions within each subsection of the tasks divided by the total number of words produced in the corresponding subsections.

Data Analyses

Double-blind data entry were used among research assistants during training and data entry. English verbal fluency task responses were transcribed by two trained research assistants fluent in the English language while Spanish verbal fluency task responses were transcribed by two trained research assistants fluent in the Spanish language. After successful validity checks across transcription data sets, recorded responses were then classified into rule violations, perseverations, or intrusion errors by two trained research assistants. Category fluency task intrusions were either classified as unrelated intrusions (e.g., 'pencil' in the animals category or 'salad' in the vegetables category). Responses classified as rule violations were quantified as the addition of suffixes to the same root words (e.g., fast, fasting, faster). Perseverations were quantified as when a word were produced more than once in a given subsection.

Phonemic fluency task intrusions were classified as unrelated intrusions (e.g., 'learn' for letter 'F') or phonemically related intrusions (e.g., 'phone' for the letter 'F' or 'circle' for the letter 'S'). Phonemic fluency task rule violations included the addition of suffixes to the same root words, names of people, names of places, and numbers. Perseverations were determined based on if a word were produced more than once in a given section.

Reliability: All participants' responses were transcribed and scored by one judge. A second judge independently verified the transcriptions and scored all responses. Inter-judge agreement for total number of intrusions were established. Two judges also verified

the scoring for the total number of intrusion type, rule violations and perseverations within each subsection of the phonemic and category fluency tasks. Inter-judge reliability coefficients were obtained for each dependent measure. After intrusion totals and percentages for English and Spanish verbal fluency tests were quantified, they were recoded and merged into variables based on the participants' primary language. So intrusion score variables included intrusion totals and percentages that were derived from tests that were conducted in the participant primary language.

Statistical Analysis

The first prediction is that participants in Visit 1 that are diagnosed with MCI or dementia will produce a higher rate of intrusion errors of all types on both phonemic and category verbal fluency tasks but particularly higher on the category tasks compared to CN participants. To test this, six repeated measures ANCOVAs were conducted to analyze differences in category fluency and phonemic intrusion scores [totals and percentages] (within group factor) across the three diagnostic groups (between groups factor) using demographic variables, including sex, age, and education as covariates. An interaction between group and verbal fluency task intrusion scores were expected with dementia and MCI presenting more intrusions in the category fluency tasks than CN.

The second prediction expects those classified as CN at year 1 that progress to MCI or dementia at years 2 or 3 would have experienced a significantly higher rate of intrusion errors on verbal fluency tasks compared to participants who did not progress to MCI or dementia. To examine this, six one-way analyses of covariance (ANCOVAs) were used to analyze differences in category fluency and phonemic intrusion scores [totals and percentages] between progressors versus non-progressors among the CN and

the MCI groups. A Bonferroni Post-Hoc analysis was conducted to determine between-groups differences on intrusion types.

The third prediction is that intrusion types as independent variables will predict progression from CN at year 1 to MCI at years 2 or 3 and from MCI at year 1 to dementia at years 2 or 3. To test the third prediction, two binomial logistic regressions were conducted using progressors vs. non-progressors as dependent measures and intrusion types [percentages and totals] as independent predictors.

Results

Hypothesis 1: Intrusion Totals and Percentages at years 1 and 2 across diagnoses: Category Fluency

A 2 x 3 repeated measures ANCOVA were conducted to evaluate if participants that are diagnosed with MCI or dementia will produce a higher rate of intrusion errors of all types on category verbal fluency tasks compared to CN participants. Mauchly's Test of Sphericity was not violated. We observed differences in category fluency intrusion totals in visit 1 and 2 (within group factor) according to diagnosis (between groups factor containing 3 levels: CN, MCI, Dementia) using demographic variables, including sex, age, and education as covariates (means and standard deviations shown in Table 3). The results of the ANCOVA indicated no significant main effects or interactions.

A 2 x 3 repeated measures ANCOVA repeated measures ANCOVA observed differences in category fluency intrusion percentages in visit 1 and 2 (within group factor) according to diagnosis (between groups factor containing 3 levels: CN, MCI, Dementia) using demographic variables, including sex, age, and education as covariates (means and standard deviations shown in Table 3). Mauchley's Test of Sphericity was not violated. The results of the ANCOVA indicated that there were no significant main effects or interactions.

A 2 x 3 repeated measures ANCOVA (refer to Table 4) was conducted to evaluate differences in category fluency correct scores in visits 1 and 2 (within group factor) according to diagnosis (between groups factor containing 3 levels: CN, MCI, dementia)

using demographic variables, including sex, age, and education as covariates. Mauchly's Test of Sphericity was violated, p < .001, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. Using Greenhouse-Geisser estimates, it was found that while there were no significant main effects or interaction effects between category fluency intrusion correct scores and diagnoses. There was a significant main effect found between diagnostic groups on category fluency correct scores, $F(2, 156) = 24.94$, $p < .001$, $\eta_p^2 = .186$. This result indicated that lower the correct scores on the category verbal fluency task, the more severe the diagnosis. Two paired samples t-tests were used to made post hoc comparisons across visits. A first samples t-test indicated that there was a significant difference between the category correct scores at year 1 between CN and Dementia groups ($t(158) = 6.88$, $p = .001$), CN and MCI groups ($t(158) = 4.72$, $p < .001$), as well as MCI and Dementia groups ($t(158) = 4.32$, $p < .001$). A second paired samples t-test indicated that there was a significant difference in category correct scores between the dementia groups at year 1 and year 2 ($t(158) = 3.33$, $p = .001$).

Hypothesis 1: Intrusion Totals and Percentages at years 1 and 2 across diagnoses: Phonemic Fluency

A 2 x 3 repeated measures ANCOVA (refer to Table 5) were conducted to evaluate if participants that are diagnosed with MCI or dementia produce a higher rate of intrusion errors of all types on phonemic verbal fluency tasks compared to CN participants. We analyzed differences in phonemic fluency intrusion totals in visit 1 and 2 (within group factor) by diagnosis (between groups factor containing 3 levels: CN, MCI, Dementia) using demographic variables, including sex, age, and education as

covariates. Mauchly's Test of Sphericity were violated, $p < .001$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. Using Greenhouse-Geisser estimates, there were no significant main effects, however there is an interaction effect between phonemic fluency intrusion totals and diagnoses, $F(2, 152) = 6.89$, $p < .001$, $\eta_p^2 = .083$, with higher FLAS intrusion percentages produced in different diagnostic groups depending on the year of evaluation. As shown in Table 4, in year 1, the dementia group had a higher intrusion total ($M = 1.39$, $SD = 2.70$) than the CN ($M = .22$, $SD = .57$) and MCI groups ($M = .30$, $SD = .75$), although in years 1 and 2, MCI had a higher phonemic intrusion percentage than CN and dementia. There was a significant main effect found between diagnostic groups on phonemic intrusion total, $F(2, 152) = 3.27$, $p = .041$, $\eta_p^2 = .041$. This result indicated that the higher the intrusion percentage on the phonemic verbal fluency task, the more severe the diagnosis. Three paired samples t-tests were used to made post hoc comparisons across visits. A first samples t-test indicated that there were significant differences between the FLAS intrusion totals at year 1 between CN and Dementia groups ($t(158) = 3.39$, $p < .001$), as well as MCI and Dementia groups ($t(158) = 3.52$, $p < .001$). A second paired samples t-test indicated that there was a significant difference in FLAS intrusion total between the dementia groups at year 1 and year 2 ($t(158) = 3.55$, $p < .001$). A third samples t-test indicated that a significant difference in FLAS intrusion total between CN and Dementia groups ($t(158) = 2.55$, $p = .012$) as well as between MCI and Dementia groups ($t(158) = 2.05$, $p = .042$).

A 2 x 3 repeated measures ANCOVA (refer to Table 4) was conducted to evaluate differences in phonemic fluency intrusion percentages in visit 1 and 2 (within group factor) according to diagnosis (between groups factor containing 3 levels: CN, MCI, dementia) using demographic variables, including sex, age, and education as

covariates. Mauchly's Test of Sphericity was violated, $p < .001$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. Using Greenhouse-Geisser estimates, it was found that while there were no significant main effects, there was an interaction effect between phonemic fluency intrusion percentage and diagnoses, $F(2, 156) = 6.60$, $p = .002$, $\eta_p^2 = .080$. As shown in Table 4, in year 1, the dementia group had a higher intrusion percentages (M = .04, SD =.09) than the CN (M = .01, SD = .02) and MCI groups (M = .01, SD = .03), but then in years 1 and 2, MCI had a higher intrusion percentage than CN and dementia. There was a significant main effect found between diagnostic groups on phonemic intrusion percentages, $F(2, 156) = 3.90$, $p = .022$, $\eta_p^2 = .049$. This result indicated that higher the intrusion percentage on the phonemic verbal fluency task, the more severe the diagnosis. Three paired samples t-tests were used to made post hoc comparisons across visits. A first samples t-test indicated that there was a significant difference between the FLAS intrusion percent at year 1 between CN and Dementia groups ($t(158) = 3.27$, $p = .001$), as well as MCI and Dementia groups ($t(158) = 3.40$, $p = .001$). A second paired samples t-test indicated that there was a significant difference in FLAS intrusion percent between the dementia groups at year 1 and year 2 ($t(158) = 3.67$, $p < .001$). A third samples t-test indicated that a significant difference in FLAS intrusion percent between CN and Dementia groups ($t(158) = 2.63$, $p = .006$) as well as MCI and Dementia groups ($t(158) = 2.29$, $p = .020$).

A 2 x 3 repeated measures ANCOVA (refer to Table 4) was conducted to evaluate differences in phonemic fluency correct scores in visits 1 and 2 (within group factor) according to diagnosis (between groups factor containing 3 levels: CN, MCI, dementia) using demographic variables, including sex, age, and education as covariates. Mauchly's Test of Sphericity was violated, $p < .001$, therefore degrees of freedom were corrected

using Greenhouse-Geisser estimates of sphericity. Using Greenhouse-Geisser estimates, it was found that while there were no significant main effects or interaction effects between category fluency intrusion correct scores and diagnoses. There was a significant main effect found between diagnostic groups on phonemic correct scores, $F(2, 156) = 5.44$, $p = .005$, $\eta_p^2 = .052$. This result indicated that lower the correct scores on the phonemic verbal fluency task, the more severe the diagnosis. Two paired samples t-tests were used to made post hoc comparisons across visits. A first samples t-test indicated that there was a significant difference between the phonemic correct scores at year 1 between CN and Dementia groups ($t(158) = 2.77$, $p = .018$), CN and MCI groups ($t(158) = 2.88$, $p = .013$). A second paired samples t-test indicated that there was no significant difference in phonemic correct scores between the dementia groups at year 1 and year 2.

Hypothesis 2: Progressors versus Non-progressors on Intrusion totals and percentages at Year 1

According to the second hypothesis, those classified as CN or MCI at year 1 and progress to MCI or dementia at years 2 or 3 are predicted to have a significantly higher rate of intrusion errors on verbal fluency tasks compared to participants who did not progress to MCI or dementia. To examine this, ANCOVAs were used to compare the total number of intrusions between progressors versus non-progressors for category and phonemic verbal fluency tasks, while controlling for years of education. A one-way ANCOVA (refer to Table 7) was conducted to determine whether intrusion error totals in year in Category tasks conducted at year 1 differed based on whether cognitively normal participants at year 1 progressed to MCI or dementia at years 2 or 3, while controlling for years of education. Levene's test of homogeneity of variance and normality checks were

carried out and the assumptions met. There were no significant differences in intrusion totals between progressors and non-progressors.

A one-way ANCOVA (refer to Table 7) was conducted to determine whether intrusion error percentages in Category tasks conducted at year 1 differed based on whether cognitively normal participants at year 1 progressed to MCI or dementia at years 2 or 3, while controlling for years of education. Levene's test of homogeneity of variance and normality checks were carried out and the assumptions met. No significant differences were found in intrusion percentages between progressors and non-progressors.

A one-way ANCOVA (refer to Table 7) was conducted to determine whether correct scores in Category tasks conducted at year 1 differed based on whether cognitively normal participants at year 1 progressed to MCI or dementia at years 2 or 3, while controlling for years of education. Levene's test of homogeneity of variance and normality checks were carried out and the assumptions met. There were significant differences in correct scores between progressors and non-progressors, $F(1, 242) = 27.64$, $p < .001$, $\eta_p^2 = .102$, with non-progressors having significantly higher correct scores than progressors.

A one-way ANCOVA was also conducted to determine whether intrusion error totals in phonemic fluency conducted at year 1 differed based on whether cognitively normal participants at year 1 progressed to MCI or dementia at years 2 or 3, while controlling for years of education (refer to Table 8). Levene's test of homogeneity of variances and normality checks were carried out and the assumptions met. There were no significant differences in intrusion totals between progressors and non-progressors.

A one-way ANCOVA (refer to Table 8) was conducted to determine whether intrusion error percentages in Phonemic fluency tasks conducted at year 1 differed based on whether cognitively normal participants at year 1 progressed to MCI or dementia at years 2 or 3. Levene's test of homogeneity of variances and normality checks were carried out and the assumptions met. There was a significant difference in intrusion percentages between progressors and non-progressors, $F(1, 187) = 4.24$, $p = .041$, $\eta_p^2 = .022$. Due to a violation of Levene's test of homogeneity of variances, a one-way ANOVA was conducted using a Welsch's test for unequal variances to control for this violation. However, the test was not significant.

A one-way ANCOVA (refer to Table 7) was conducted to determine whether correct scores in Phonemic (FLAS) tasks conducted at year 1 differed based on whether cognitively normal participants at year 1 progressed to MCI or dementia at years 2 or 3, while controlling for years of education. Levene's test of homogeneity of variance and normality checks were carried out and the assumptions met. There were no significant differences in phonemic fluency correct scores between progressors and non-progressors.

Hypothesis 3: Intrusion types as predictors of progressors and non-progressors

Intrusion types, as well as unrelated and phonemic intrusions, are expected to predict progression from CN or MCI to dementia. To test the third hypothesis, two binomial logistic regressions were conducted using the progressor/non-progressor (coded: 0 for non-progressor, 1 for progressor) classification as the dependent measure, intrusion type as the independent predictor, and education as a covariate. A binary logistic regression was performed to investigate if category fluency unrelated intrusion type totals at year 1 would predict progression at years 2 or 3 (Refer to Table 9). For Block 1 of the model,

education was a significant predictor for progression, x^2 (1, 189) = 6.36, p = .012. Block 1 correctly classified 73.5 percent of the participants as either progressors or non-progressors, however the years of education correctly predicted 100 percent of non-progressors, while only correctly predicting 3.8 percent of progressors. Block 2 of the model included unrelated intrusion totals. The logistic regression for Block 2 was not statistically significant. The overall model remained significant, x^2 (2, 188) = 6.506, p = .039. No values in the analyses had a Cook's distance greater than 1, the largest was 0.55, indicating that none of the cases were determined to a to be a legitimate part of the sample and were retained for the logistic regression analyses.

A binary logistic regression evaluated if category fluency unrelated intrusion type percentages at year 1 would predict progression at years 2 or 3. The logistic regression model was not statistically significant. As shown in Table 9, Block 1 of the model, education was a significant predictor for progression, x^2 (1, 189) = 6.36, p = .012. Block 1 correctly classified 73.5 percent of the participants as either progressors or non-progressors, however years of education correctly predicted 100 percent of non-progressors, while correctly predicting none of the progressors. Block 2 of the model included unrelated intrusion percentages. The logistic regression for Block 2 was not statistically significant. The overall model remained significant, x^2 (2, 188) = 7.036, p = .030. No values in the analyses had a Cook's distance greater than 1, the largest was 0.55, indicating that none of the cases were determined to a to be a legitimate part of the sample and were retained for the logistic regression analyses.

A binary logistic regression evaluated if category fluency correct scores at year 1 would predict progression at years 2 or 3. The logistic regression model was statistically significant. As shown in Table 9, Block 1 of the model, education was a significant

predictor for progression, x^2 (1, 243) = 6.03, p = .014. Block 1 correctly classified 75.1 percent of the participants as either progressors or non-progressors, however years of education correctly predicted 100 percent of non-progressors, while correctly predicting none of the progressors. Block 2 of the model included correct scores. The logistic regression for Block 2 was statistically significant, x^2 (1, 243) = 26.26, p < .001, Block 2 correctly classified 78.1 percent of the progressors and non-progressors, however 96.2 percent of non-progressors were predicted correctly and 23 percent of progressors were correctly predicted. The overall model remained significant, x^2 (2, 242) = 32.26, p < .001. No values in the analyses had a Cook's distance greater than 1, the largest was 0.78, indicating that none of the cases were determined to a to be a legitimate part of the sample and were retained for the logistic regression analyses.

A binary logistic regression was also conducted to investigate if phonemic fluency unrelated intrusion type totals at year 1 predicted disease progression at years 2 or 3. As shown in Table 10, for Block 1 of the model, education was a significant predicator for progression, x^2 (1, 179) = 6.06, p = .014. Block 1 correctly classified 72 percent of the participants as either progressors or non-progressors however years of education correctly predicted 98.5 percent of the non-progressors while only predicting 5.8% of the progressors. Block 2 of the model included unrelated intrusion totals. The logistic regression for Block 2 was not statistically significant. The overall model remained significant, x^2 (2, 179) = 6.096, p = .047. No values in the analyses had a Cook's distance greater than 1, the largest was 0.53, indicating that none of the cases were determined to a to be a legitimate part of the sample and were retained for the logistic regression analyses.

A binary logistic regression was also performed to investigate if phonemic fluency unrelated intrusion type percentages at year 1 predicted progression at years 2 or 3. As shown in Table 10, Block 1 of the model, education was a significant predictor for progression, x^2 (1, 178) = 4.80, p = .028, correctly classifying 71.9 percent of the participants as either progressors or non-progressors. Years of education correctly predicted 100% of non-progressors, while only correctly predicting 2% of the progressors. Block 2 of the model included unrelated intrusion percentage, and was not statistically significant. The overall model did not remain significant, x^2 (2, 179) = 5.326, p = .070. No values in the analyses had a Cook's distance greater than 1, the largest was 0.53, indicating that none of the cases were determined to a to be a legitimate part of the sample and were retained for the logistic regression analyses.

A binary logistic regression investigated if phonemic fluency intrusion type totals at year 1 predicted progression at years 2 or 3. As shown in Table 10, Block 1 of the model, education was a significant predictor for progression, x^2 (1, 178) = 6.06, p = .014. Block 1 correctly classified 72 percent of the participants as either progressors or non-progressors, however years of education correctly predicted 100 percent of non-progressors, while correctly predicting none of the progressors.. Block 2 of the model, however, included phonemic intrusion totals. and was not statistically significant. The overall model remained significant, x^2 (2, 188) = 6.136, p = .047. No values in the analyses had a Cook's distance greater than 1, the largest was 0.55, indicating that none of the cases were determined to a to be a legitimate part of the sample and were retained for the logistic regression analyses.

A binary logistic regression was performed to investigate if phonemic intrusion type percentages at year 1 predicted progression at years 2 or 3. As shown in Table 10,

Block 1 of the model, education was a significant predictor for progression, x^2 (1, 181) = 4.84, p = .028, and correctly classified 72.4 percent of the participants as either progressors or non-progressors, however years of education correctly predicted 100 percent of non-progressors, while correctly predicting none of the progressors. Block 2 of the model included phonemic intrusion percentage. The logistic regression for model 2 was not statistically significant. The overall model remained significant, x^2 (2, 180) = 5.075, p = .079. No values in the analyses had a Cook's distance greater than 1, the largest was 0.52, indicating that none of the cases were determined to a to be a legitimate part of the sample and were retained for the logistic regression analyses.

A binary logistic regression evaluated if phonemic (FLAS) fluency correct scores at year 1 would predict progression at years 2 or 3. The logistic regression model was statistically significant. As shown in Table 9, Block 1 of the model, education was a significant predictor for progression, x^2 (1, 224) = 6.40, p = .011. Block 1 correctly classified 75.3 percent of the participants as either progressors or non-progressors, however years of education correctly predicted 100 percent of non-progressors, while correctly predicting none of the progressors. Block 2 of the model included correct scores. The logistic regression for Block 2 was not statistically significant. The overall model remained significant, x^2 (2, 223) = 9.34, =.009. No values in the analyses had a Cook's distance greater than 1, the largest was 0.53, indicating that none of the cases were determined to a to be a legitimate part of the sample and were retained for the logistic regression analyses.

Discussion

Hypothesis 1

The first hypothesis for this study aimed to test whether participants that are diagnosed with MCI or dementia at year 1 will produce a higher rate of intrusion error totals and percentages on both phonemic and category verbal fluency tasks at years 1 and 2 but particularly higher on the category tasks compared to CN participants.

Our results found that there were no interaction effects between category intrusion totals across both time points (year 1 and year 2) and diagnoses. There were no significant differences found between diagnostic groups on category intrusion totals. These results parallel the finding of no interaction effect between category intrusion percentages across both time points and diagnoses. There were also no significant differences found between diagnostic groups on category intrusion percentages. The category fluency task results do not support the hypothesis that participants diagnosed with MCI or dementia at year 1 will produce a higher rate of intrusion totals and percentages compared to the FLAS task at years 1 and 2. However, this pattern of results is consistent with finding from previous research investigating intrusion scores within verbal fluency tasks, where category and FLAS intrusions were not significantly related to cognitive diagnosis longitudinally (Auriacombe et al., 2006; Clark et al., 2016). The category fluency intrusion results were also disputed the prediction that category fluency task intrusion totals and percentages would be higher compared to on phonemic fluency tasks. These findings oppose research conducted by Henry et al. (2004) and Monsch et al. (1992), in which verbal fluency task

correct scores were investigated. It was found that when reviewing the validity of semantic and phonemic (FLAS) fluency tasks among those with CN or dementia, category fluency had a higher rate of reliability in differentiating between CN and dementia participants (Henry et al., 2004; Monsch et al., 1992).

Category correct scores at years 1 and 2 were significantly different between diagnostic groups, indicating that the lower the category correct scores, the more severe the diagnosis. This finding coincided with previous literature in which category correct scores were found to be significantly associated with cognitive diagnosis (Henry et al., 2004; Monsch et al., 1992).

Although the category fluency task intrusion totals and percentages did not present any significance in this analysis, the FLAS task presented significance. Our results found that there was a significant interaction effect between FLAS intrusion totals across both time points diagnoses. There were also significant differences found between diagnostic groups on FLAS intrusion totals, indicating that the higher the FLAS intrusion totals, the more severe the diagnosis. Just as FLAS intrusion totals held significance, there was a significant interaction effect between FLAS intrusion percentages across both time points and diagnoses, as well as significant differences found between diagnostic groups on FLAS intrusion percentages. These results for FLAS intrusion percentages also indicate that the higher the FLAS intrusion percentage, the more severe the diagnosis. The FLAS results of for this particular hypothesis provides supporting evidence that participants diagnosed with MCI or dementia at year 1 will produce a higher rate of intrusion totals and percentages on the FLAS task at years 1 and 2. This pattern of results is consistent with research conducted by Auriacombe et al., (2006) in which verbal fluency intrusions at the 4th (8 years post-baseline) timepoint of a 5 timepoint study

(baseline, 3 years, 5 years, 8 years, and 10 years post-baseline), were found to show significant differences between CN and Alzheimer's disease groups. However, it is inconsistent with previous research that presented no significant differences in verbal fluency intrusions across diagnostic groups (Binetti et al., 1995; Clark et al., 2016).

Phonemic fluency (FLAS) correct scores at years 1 and 2 were significantly different between diagnostic groups, indicating that the lower the phonemic correct scores, the more severe the diagnosis. This finding coincided with previous literature in which phonemic correct scores were found to be significantly associated with cognitive diagnosis (Henry et al., 2004; Monsch et al., 1992). This also aligned with previous literature in that it was less significantly related to diagnosis than category fluency correct scores, being that significant differences emerged between CN and Dementia groups as well as MCI and Dementia groups, while category fluency correct scores displayed differences between all groups, CN and MCI, MCI and Dementia, and CN and Dementia.

Hypothesis 2

The second hypothesis examined if participants classified as CN or MCI at year 1 that progress to MCI or dementia at years 2 or 3 will have experienced a significantly higher rate of intrusion errors on verbal fluency tasks compared to participants who did not progress to MCI or dementia.

No significant differences emerged in category intrusion totals or percentages between progressors and non-progressors, indicating that progressor versus non-progressor classification is not associated with category intrusion totals or percentages. Another investigation of intrusion totals on the Animal category of the category verbal fluency task, also found no significant differences in intrusion totals across participants

with CN, mild Alzheimer's disease, and severe Alzheimer's disease groups (Binetti et al., 1995).

In contrast to the category fluency intrusion totals and percentage analyses, there were significant associations found in category fluency correct scores between progressors and non-progressors.

These results mirrored the FLAS intrusion totals and percentages, in that no significances were found, therefore no association between FLAS intrusion totals or percentages and progressor versus non-progressor classification could be concluded.

Similarly to the FLAS fluency intrusion totals and percentage analyses, there were no significant associations found in FLAS fluency correct scores between progressors and non-progressors.

Hypothesis 3

Intrusion types were expected to predict progression from CN or MCI to MCI or dementia.

There were no significant results for this analysis. Unrelated intrusion totals and percentages for both Category fluency and FLAS tasks did not predict progressor versus non-progressor classification. Phonemic intrusion totals and percentages for the FLAS task did not significantly predict progressor versus non-progressor classification. Whereas past research has found that higher rates of intrusion types may demonstrate semantically related intrusion types were significantly associated with amnestic MCI and dementia (Lowenstein et al., 2018; Torres et al., 2019, Thomas et al., 2019), the present study has shown that intrusion types, unrelated and phonemic, in verbal fluency tasks are not significant predictors of cognitive diagnosis progression.

These discrepancies in the association of intrusions across diagnosis may be due to the difference in the neuropsychological examination used. Lowenstein et al. (2018) and Torres et al. (2019) used a verbal learning task called the Lowenstein-Acevedo Scales of Semantic Interference and Learning (LASSI-L). The tasks used in the present study included verbal fluency tasks such as the COWAT (FLAS) and Category Fluency tasks. Verbal fluency tasks are typically a measure of verbal fluency and lexical retrieval, while verbal learning tasks are associated with encoding, storage, and retrieval. The pattern of insignificant results is consistent with the previous literature in which intrusions in verbal fluency tasks showed no significant differences across diagnostic groups, participants with CN and Alzheimer's disease (Auriacombe et al., 2006). This is also consistent with findings from Clark et al., (2016), in which verbal fluency task intrusions were scored and revealed to have no significant across cognitive diagnostic groups. Although intrusion types in both category and FLAS tasks did not significantly predict progressors versus non-progressor classification, correct scores for category and FLAS tasks did significantly predict progressor versus non-progressor classification.

Future Directions & Limitations

The sample for this research consists of participants with a mean age of 72.24 years at baseline (SD = 7.99). Increased age of the sample population makes participants more vulnerable to health complications. Additionally, participant attrition rate across the 2-year time span was impactful to the sample size of longitudinal data due to the longitudinal design of this study (i.e., there is a general increase in susceptibility to losing participants due to varying reasons over time).

As stated previously, there is minimal research exploring intrusion errors, verbal fluency tasks, and cognitive diagnosis (Auriacombe et al., 2006; Clark et al., 2016). Due to the unlimited or continuous response design of verbal fluency tasks, formulas are unable to be properly created without potential skewing of results. For example, if percentages were used, where intrusion total would be divided by the total number of words, then someone who may have produce 10 words but 2 intrusions and someone else who may have produced 20 words but 4 intrusions would then be measured on the same level, although they had differing word totals. Another limitation that could impact proper measurement of intrusion error is the differences in intrusions and how unrelated it is to the sub-category being given. These may misrepresent the intrusion data from participants.

Due to these limitations, such as the impact of verbal fluency tasks not having a fixed number of correct and incorrect words said, it may be useful to control for this type of hindrance by creating a verbal fluency task in which only a set number of words are

used. For example, if exactly 30 words are used per sub-category in each verbal fluency task. Another limitation that may need to be examined is the lack of accounting for how irrelevant an intrusion is compared to others. If properly weighted, these intrusions may show more of a significant association with progressor versus mon-progressor classifications. These could be further investigated through future research.

Appendices

Appendix A: Demographic Descriptive Information

Table 1

Baseline Demographic for Whole Sample and by Diagnostic Group

	Whole Sample	CN	MCI	Dementia			
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*F*	*p*	η_p^2
	n = 375	*n* = 88	*n* = 229	*n* = 58			
Age	72.24 (7.99)	70.27 (6.54)	72.98 (7.72)	72.33 (10.36)	3.694	0.026	0.019
Education	15.00 (3.65)	16.01 (3.22)	15.06 (3.58)	13.26 (3.97)	10.523	< 0.001	0.054
CDR	.65 (.10).	.00 (0.00)	.5 (0.00)	1 (0.00)			
Length of follow-up (months)	17.60 (7.01)	18.76 (6.24)	17.22 (6.50)	17.34 (7.25)	.321	.576	.001
					x^2	*p*	
Female Frequency	236 (62.93%)	68 (77.30%)	134 (58.50%)	34 (56.70%)	10.135	0.006	

Table 2

Baseline Demographic for Whole Sample and in Progressors vs. Non-Progressors

	Whole Sample	Progressors	Non-Progressors			
	M (SD)	*M (SD)*	*M (SD)*	*F*	*p*	η_p^2
	n = 269	*n* = 66	*n* = 203			
Age	72.13 (7.80)	73.53 (9.00)	71.68 (7.34)	2.82	0.094	0.010
Education	14.93 (3.58)	14.12 (3.64)	15.19 (3.53)	4.49	0.035	0.0172
CDR	2.28 (3.15)	2.43 (3.60)	2.23 (2.99)			
Length of follow-up (months)	18.77 (6.93)	19.12 (6.18)	18.66 (7.17)	.225	.636	.001
				x^2	*p*	
Female Frequency	173 (64.31%)	45 (68.18%)	128 (63.05%)	.571	0.450	

Appendix B: Hypothesis 1

Table 3

Years 1 and 2 Category intrusion percentages and totals Means(SD)

	CN *n* = 45	MCI *n* = 112	Dementia *n* = 15	Total *n* = 172
Year 1 Category Intrusion Percent	.02 (.03)	.02 (.04)	.01 (.03)	.02 (.04)
Year 2 Category Intrusion Percent	.01 (.03)	.02 (.04)	.03 (.06)	.02 (.04)
Year 1 Category Intrusion Total	.73 (1.72)	.61 (2.02)	.27 (.59)	.62 (1.86)
Year 2 Category Intrusion Total	.42 (1.54)	.71 (1.57)	.60 (1.06)	.62 (1.52)
Year 1 Category Correct Score	47.49 (10.99)	38.90 (11.02)	29.29 (10.81)	40.11 (12.22)
Year 2 Category Correct Score	47.29 (12.06)	37.74 (11.73)	24.58 (11.21)	38.81 (13.45)

Appendix B: Hypothesis 1 (*continued*..)

Table 4

Years 1 and 2 FLAS intrusion percentages and totals Means(SD)

	CN *n* = 41	MCI *n* = 104	Dementia *n* = 13	Total *n* = 158
Year 1 FLAS Intrusion Percent	.01(.01)	.01(.03)	.04(.09)	.01(.03)
Year 2 FLAS Intrusion Percent	.00(.01)	.01(.02)	.01(.03)	.01(.02)
Year 1 FLAS Intrusion Total	.22(.57)	.30(.75)	1.38(2.69)	.37(1.05)
Year 2 FLAS Intrusion Total	.10(.37)	.40(.96)	.31(.85)	.32(.85)
Year 1 FLAS Correct Score	51.96 (15.60)	44.78 (15.07)	38.19 (16.60)	45.98 (15.83)
Year 2 FLAS Correct Score	55.08 (16.61)	44.57 (15.39)	37.71 (17.20)	46.60 (16.74)

Appendix B: Hypothesis 1 (*continued*..)

Table 5

2 x 3 Repeated Measures ANCOVA results for FLAS intrusion total

	SS	df	MS	F	*p*-value	ηp2
Between groups						
Intercept	0.171	1	0.171	0.151	0.698	0.001
Sex	0.511	1	0.511	0.449	0.504	0.003
Age	0.454	1	0.454	0.399	0.529	0.003
Education	1.125	1	1.125	0.989	0.322	0.006
Diagnosis	7.434	2	3.717	3.902	0.022	0.041
Within-groups						
Year	0.093	1	0.093	0.155	0.694	0.001
Year * Sex	0.453	1	0.453	0.756	0.386	0.005
Year * Age	0.015	1	0.015	0.025	0.873	0.000
Year * Education	0.000	1	0.000	0.000	0.996	0.000
Year * Diagnosis	8.263	2	4.131	6.894	0.001	0.083
Error		152				
Total		158				

Table 6

2 x 3 Repeated Measures ANCOVA results for FLAS intrusion percentage

	SS	df	MS	F	*p*-value	ηp2
Between groups						
Intercept	0.000	1	0.000	0.176	0.676	0.001
Sex	0.001	1	0.001	0.801	0.372	0.005
Age	0.000	1	0.000	0.406	0.525	0.003
Education	0.002	1	0.002	1.453	0.23	0.009
Diagnosis	0.008	2	0.004	3.902	0.022	0.049
Within-groups						
Year	0.000	1	0.000	0.375	0.541	0.002
Year * Sex	0.001	1	0.001	1.027	0.313	0.007
Year * Age	0.000	1	0.000	0.094	0.759	0.001
Year * Education	0.000	1	0.000	0.038	0.846	0.000
Year * Diagnosis	0.007	2	0.003	6.603	0.002	0.080
Error		152				
Total		158				

Appendix C: Hypothesis 2

Table 7

Category intrusion totals & percentages at year 1 in progressors versus non-progressors

	Non-Progressor *n* = 148	Progressor *n* = 50	Total *n* = 198
Year 1 Category Intrusion Total	.60(1.95)	.62(1.09)	.61(1.77)
Year 1 Category Intrusion Percent	.014(.04)	.029(.061)	.018(.047)
Year 1 Category Correct Score	41.92 (12.16)	31.43 (13.13)	39.31 (13.13)

Table 8

FLAS intrusion totals & percentages at year 1 in progressors versus non-progressors

	Non-progressor *n* = 143	Progressor *n* = 46	Total *n* = 189
Year 1FLAS Intrusion Total	.22(.57)	.30(.75)	1.38(2.69)
Year 1 FLAS Intrusion Percent	.10(.37)	.40(.96)	.31(.85)
Year 1 FLAS Correct Score	45.92 (16.79)	39.55 (17.95)	44.35 (17.26)

Appendix D: Hypothesis 3

Table 9

Summary of Hierarchical Logistic Analyses of Demographic and Year 1 Category Intrusion Type Totals & Percentages

Progressor group *n* = 52; Non-progressor group *n* = 137	*B*	*SE B*	Wald χ^2	*p*	Exp(B)
Block 1: Education	-0.114	0.046	6.019	0.014	0.892
Block 2: Category Unrelated Intrusion Total	-0.037	0.101	0.134	0.714	0.964
Block 3: Category Unrelated Intrusion Percentage	-3.310	4.37	0.574	0.449	0.37

Table 10

Summary of Hierarchical Logistic Analyses of Demographic and Year 1 FLAS Intrusion Type Totals & Percentages

Progressor group *n* = 52; Non-Progressor group *n* = 130	*B*	*SE B*	Wald χ^2	*p*	Exp(B)
Block 1: Education	-0.114	0.048	5.746	0.017	0.892
Block 2: FLAS Unrelated Intrusion Total	-0.014	0.079	0.029	0.864	0.987
Block 3: FLAS Unrelated Intrusion Percentage	-1.632	2.723	0.359	0.549	0.195
Block 4: FLAS Phonemic Intrusion Total	-0.089	0.333	0.071	0.79	0.915
Block 5: FLAS Phonemic Intrusion Percentage	-6.401	13.405	0.228	0.633	0.002

References

American Psychological Association (APA). (n.d.). *Apa Dictionary of Psychology*. American Psychological Association. Retrieved November 15, 2021, from

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Ardila, A., & Ramos, E. (2008). Normal and abnormal aging in bilinguals. *Dementia & Neuropsychologia*, *2*(4), 242–247.

Auriacombe, S., Lechevallier, N., Amieva, H., Harston, S., Raoux, N., & Dartigues, J. F. (2006). A longitudinal study of quantitative and qualitative features of category verbal fluency in incident Alzheimer's disease subjects: results from the PAQUID study. *Dementia and geriatric cognitive disorders*, *21*(4), 260–266.

Beekly, D. L., Ramos, E. M., Lee, W. W., Deitrich, W. D., Jacka, M. E., Wu, J., & Kukull, W. (2007). The National Alzheimer's coordinating Center (NACC) database: The uniform data set. *Alzheimer Disease and Associated Disorders, 21*(3), 249–258.

Benedict, R. H. B., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins verbal learning test – Revised: Normative data and analysis of inter-form and test-retest reliability. *Clinical Neuropsychologist, 12*(1), 43–55.

Benton, A. L., & Hamsher, K. (1976). Multilingual aphasia examination (2nd ed.). Iowa City, IA: AJA Associates.

Binetti, G., Magni, E., Cappa, S. F., Padovani, A., Bianchetti, A., & Trabucchi, M. (1995). Semantic memory in Alzheimer's disease: an analysis of category fluency. *Journal of clinical and experimental neuropsychology*, *17*(1), 82–89.

Borella, E., Carretti, B., Mitolo, M., Zavagnin, M., Caffarra, P., Mammarella, N., Fairfield, B., Gamboz, N., & Piras, F. (2017). Characterizing cognitive inhibitory deficits in mild cognitive impairment. *Psychiatry Research*, *251*, 342–348.

Buchhave, P., Minthon, L., Zetterberg, H., Wallin, A. K., Blennow, K., & Hansson, O. (2012). Cerebrospinal fluid levels of β-amyloid 1-42, but not of tau, are fully changed already 5 to 10 years before the onset of Alzheimer dementia. *Archives of general psychiatry*, *69*(1), 98–106.

Clark, D. G., McLaughlin, P. M., Woo, E., Hwang, K., Hurtz, S., Ramirez, L., Eastman, J., Dukes, R. M., Kapur, P., DeRamus, T. P., & Apostolova, L. G. (2016). Novel verbal fluency scores and structural brain imaging for prediction of cognitive outcome in mild cognitive impairment. *Alzheimer's & dementia (Amsterdam, Netherlands)*, *2*, 113–122.

Corrigan, J. D., & Hinkeldey, N. S. (1987). Relationships between Parts A and B of the Trail Making Test. *Journal of Clinical Psychology, 43*(4), 402–409.

Csukly, G., Sirály, E., Fodor, Z., Horváth, A., Salacz, P., Hidasi, Z., Csibri, É., Rudas, G., & Szabó, Á. (2016). The Differentiation of Amnestic Type MCI from the Non-Amnestic Types by Structural MRI. Frontiers in aging neuroscience, 8, 52.

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168.

Dubois, B., Hampel, H., Feldman, H. H., Scheltens, P., Aisen, P., Andrieu, S., Bakardjian, H., Benali, H., Bertram, L., Blennow, K., Broich, K., Cavedo, E., Crutch, S., Dartigues, J. F., Duyckaerts, C., Epelbaum, S., Frisoni, G. B., Gauthier, S., Genthon, R., Gouw, A. A., … Proceedings of the Meeting of the International Working Group (IWG) and the American Alzheimer's Association on "The Preclinical State of AD"; July 23, 2015; Washington DC, USA (2016). Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. Alzheimer's & dementia : the journal of the Alzheimer's Association, 12(3), 292–323.

Fama, R., Sullivan, E. V., Shear, P. K., Cahn-Weiner, D. A., Marsh, L., Lim, K. O., Yesavage, J. A., Tinklenberg, J. R., & Pfefferbaum, A. (2000). Structural brain correlates of verbal and nonverbal fluency measures in Alzheimer's disease. *Neuropsychology*, *14*(1), 29–40.

Golden, C. J., & Freshwater, S. M. (2002). *Stroop color and word test: Adult version. A manual for clinical and experimental uses* (2a ed.). Wood Dale, Illinois: Stoelting.

Gollan, T., Weissberger, G., Runnqvist, E., Montoya, R., & Cera, C. (2012). Self-ratings of spoken language dominance: A multi-lingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Bilingualism: Language and Cognition, 15*(3), 594–615.

Haberlandt K. (2011) Intrusion Errors. In: Kreutzer J.S., DeLuca J., Caplan B. (eds) Encyclopedia of Clinical Neuropsychology. Springer, New York, NY.

Kitabayashi, Y., Ueda, H., Tsuchida, H., Iizumi, H., Narumoto, J., Nakamura, K., Kita, H., & Fukui, K. (2001). Relationship between regional cerebral blood flow and

verbal fluency in Alzheimer's disease. Psychiatry and Clinical Neurosciences, 55(5), 459–463.

Loewenstein, D. A., Curiel, R. E., DeKosky, S., Bauer, R. M., Rosselli, M., Guinjoan, S. M., Adjouadi, M., Peñate, A., Barker, W. W., Goenaga, S., Golde, T., Greig-Custo, M. T., Hanson, K. S., Li, C., Lizarraga, G., Marsiske, M., & Duara, R. (2018). Utilizing semantic intrusions to identify amyloid positivity in mild cognitive impairment. *Neurology*, *91*(10), e976–e984.

Matyas, N., Aschenberger, F. K., Wagner, G., Teufer, B., Auer, S., Gisinger, C., Kil, M., Klerings, I., & Gartlehner, G. (2019). Continuing education for the prevention of mild cognitive impairment and Alzheimer's-type dementia: a systematic review and overview of systematic reviews. BMJ Open, 9(7), e027719.

Mayeux, R., & Stern, Y. (2012). Epidemiology of Alzheimer Disease. Cold Spring Harbor Perspectives in Medicine, 2(8), a006239.

McDonnell, M., Dill, L., Panos, S., Amano, S., Brown, W., Giurgius, S., Small, G., & Miller, K. (2020). Verbal fluency as a screening tool for mild cognitive impairment. International psychogeriatrics, 32(9), 1055–1062.

Monsch, A. U., Bondi, M. W., Butters, N., Salmon, D. P., Katzman, R., & Thal, L. J. (1992). Comparisons of Verbal Fluency Tasks in the Detection of Dementia of the Alzheimer Type. *Archives of Neurology*, *49*(12), 1253–1258.

Mueller, K. D., Koscik, R. L., LaRue, A., Clark, L. R., Hermann, B., Johnson, S. C., & Sager, M. A. (2015). Verbal Fluency and Early Memory Decline: Results from the Wisconsin Registry for Alzheimer's Prevention. *Archives of Clinical Neuropsychology*, *30*(5), 448–457.

Paradise, M., Cooper, C., & Livingston, G. (2008). Systematic review of the effect of education on survival in Alzheimer's disease. International Psychogeriatrics, 21(1), 25–32.

Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni, L. (2014). Mild cognitive impairment: a concept in evolution. *Journal of Internal Medicine, 275*(3), 214-228.

Reitan, R., & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation* (2nd ed.). Tucson: Neuropsychology Press.

Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8*, 271–276.

Roberts, R., & Knopman, D. S. (2013). Classification and epidemiology of MCI. Clinics in geriatric medicine, 29(4), 753–772.

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solí, F. (2000). Verbal Fluency and Repetition Skills in Healthy Older Spanish-English Bilinguals. *Applied Neuropsychology*, *7*(1), 17–24.

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. Frontiers in Psychology, 5, 772.

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662. http://dx.doi.org/10.1037/h0054651

Thomas, K. R., Eppig, J., Edmonds, E. C., Jacobs, D. M., Libon, D. J., Au, R., Salmon, D. P., Bondi, M. W., & Alzheimer's Disease Neuroimaging Initiative* (2018). Word-list intrusion errors predict progression to mild cognitive impairment. Neuropsychology, 32(2), 235–245.

Trenerry, M., Crosson, B., DeBoe, J., & Leber, W. (1989). Stroop Neuropsychological Screening Test manual. Adessa, FL: Psychological Assessment Resources (PAR).

Torres, V. L., Rosselli, M., Loewenstein, D. A., Curiel, R. E., Uribe, I. V., Lang, M., Arruda, F., Penate, A., Vaillancourt, D. E., Greig, M. T., Barker, W. W., Bauer, R. M., & Duara, R. (2019). Types of errors on a semantic interference task in mild cognitive impairment and dementia. *Neuropsychology*, *33*(5), 670–684.

2021 Alzheimer's disease facts and figures. (2021). *Alzheimer's & dementia : the journal of the Alzheimer's Association*, *17*(3), 327–406.