**Supplementary Material – Pilot A**

**Methods**

**Participants**

This study involved 8 younger adults (ages 18-30) and 6 older adults (ages 60-85). Participants were recruited through the Baycrest Academy for Research and Education. Participants were required to be native English speakers or to have learned English before the age of five. Participants were excluded if they had medical conditions that affected their cognitive functioning (other than aMCI), such as brain injury, dementia, stroke, epilepsy, multiple sclerosis, chemotherapy, heart attacks, and heart disease. Participants with a diagnosis of anxiety, depression, or other psychiatric disorders were also excluded from the study. Individuals who were regularly taking any drugs (recreational or prescription) that could affect cognitive functioning were excluded as well. The study was completed in accordance with Helsinki Declaration and approved by the Baycrest Research Ethics Board (REB #17-16), and all participants provided written informed consent. Participants were given a $30 CAD e-gift card for completing the study.

**Materials**

The 225 images for the study were taken from a database developed by Konkle et. al (2010), available at http://cvcl.mit.edu/MM/. Twenty-five of the images were scrambled for the incidental encoding task using matPyrtools, source code developed by Portilla & Simoncelli (2013) on MATLAB (Mathworks, Natick, MA).

The 200 object images that were presented ranged from visual categories encompassing tools (16 items), vegetation (16 items), instruments (20 items), sports (16 items), animals, personal use (16 items), food (20 items), office appliances (20 items), kitchen utensils (16 items), clothing (16 items), furniture (16 items), and kitchen appliances (12 items). Each test version had images corresponding to 4 tools, 4 vegetation, 5 instruments, 4 sports, 4 animals, 4 personal use, 5 food, 5 office appliances, 4 kitchen utensils, 4 clothing, 4 furniture, and 3 kitchen appliances, and these were equally divided across studied (100 studied images) and new (100 new images) sets, as well as studied images in the long deadline (50 images) and studied images in the short deadline (50 images) sets. Images were counterbalanced to ensure equal distribution across both the studied and new sets, as well as the long and short deadline, into four total sets. Two sets were used in the study phase, and one of each set was used in the long and short deadline tests. Objects from the other two sets were used as lures in the long deadline, and objects from the remaining set were used as lures in the short deadline. Four combinations of these sets were used in counterbalancing for a total of four test versions, with each version being given to 6 participants per age group. Each individual item collected was also analyzed with a ‘memorability’ score (Isola et al., 2011; Bainbridge, 2019) to ensure the four counterbalanced stimulus sets did not statistically differ; *F* (3, 96) = .110, *p* = .954.

**Procedure**

The computer task for the study was developed and presented using Psychopy software (version 3.5, Open Science Tools, Ltd.). The task was uploaded online on Pavlovia, which created a URL for the experiment. The URL was emailed to participants during the Zoom call, and participants were assigned a number ranging from 1 to 4 as an ID corresponding to their test version.

Participants first underwent an incidental encoding task for a series of 125 object images. Twenty-five of the images were ‘scrambled’, meaning the image was impossible to identify. During this task, participants were asked to determine if an object was intact or scrambled. After a short break, participants were presented with new and old object images. In the first phase, 50 studied and 50 new images were presented for a longer duration (5s), while in the second phase, 50 studied and 50 new images (different from the first phase) were presented for a shorter duration (1s). The longer deadline was presented first to avoid participants adapting to quick responses. For both phases, participants selected whether they have seen the object before by pressing the ‘M’ key on the keyboard, or whether the object is new by pressing the ‘Z’ key on the keyboard.

**Results**

A 2 x 2 ANOVA as a function of Group (Younger Adults, Older Adults) and Deadline (Long, Short) was conducted for recognition accuracy (hit rate minus false alarms).

**Response Deadline Task Data**

Analyses of recognition accuracy, as evident on Figure 1, revealed no group differences between younger adults and older adults; *F*(1, 24) = 1.147, *p* = .295, *ηp2* = .03. Significant effects of deadline were found, *F*(1, 24) = 12.730, *p* < .001, *ηp2* = .32, in which participants had greater recognition accuracy with a longer, slower deadline, as opposed to a shorter, faster deadline. No significant interaction between deadline and group was found, *F*(1, 24) = 2.261, *p* = .146, *ηp2* = .15.

**Figure 1.** Recognition accuracy (hits – false alarms) in the Short and Long deadline, per Group.

****

**Discussion**

Our pilot data revealed that younger adults and older adults did not significantly differ in recognition accuracy performance. However, the data in Table 1 revealed that both age groups missed a significant amount of responses in the short deadline, with younger adults (*M* = 5.75; *SD* = 2.25) missing significantly fewer responses than older adults (*M* = 15.83; *SD* = 9.35), *F*(1, 12) = 8.86, *p* = .001, *ηp2* = .42. As such, we decided to extend the short deadline from 1000ms to 1200ms.

**Table 1**. Missed responses by each participant per Deadline.

|  |  |  |
| --- | --- | --- |
|  | Missed ‘Short’ Responses | Missed ‘Long’ Responses |
| Older Adult | 17 | 0 |
| Older Adult | 6 | 0 |
| Older Adult | 6 | 0 |
| Older Adult | 14 | 0 |
| Older Adult | 30 | 0 |
| Older Adult | 22 | 0 |
| Younger Adult | 5 | 1 |
| Younger Adult | 10 | 0 |
| Younger Adult | 4 | 0 |
| Younger Adult | 6 | 0 |
| Younger Adult | 8 | 0 |
| Younger Adult | 3 | 0 |
| Younger Adult | 5 | 0 |
| Younger Adult | 5 | 0 |

**Supplementary Material – Pilot B**

**Methods**

**Participants**

This study involved 35 younger adults (ages 18-30), 57 older adults (ages 60-85), and 16 older adults with aMCI (ages 60-85). Participant recruitment followed the same guidelines as Pilot A.

**Neuropsychological Testing**

Participants completed a neuropsychological assessment on Zoom. Participants began with questions to determine subjective cognitive complaints (Jessen et al., 2010), the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), Shipley Institute of Living Scale: Vocabulary (Shipley, 1946), and Wechsler Abbreviated Scale of Intelligence Matrix Reasoning (Wechsler, 1999). Older adults also completed the Rey Auditory Verbal Learning Test (Rey, 1964), Forward and Backward Digit Span (Wechsler, 2008), Logical Memory (Wechsler, 1997), the Boston Naming Test (Kaplan et al., 1983), FAS (Borkowski et al., 1967), and Animal Naming (Rosen, 1980). For participants that had impairments indicative of aMCI, an informant completed the Functional Activities Questionnaire (FAQ; Pfeffer et al., 1982), to rule out dementia. All test batteries were reviewed by a clinical neuropsychologist (NDA). The diagnostic criteria for aMCI was performance at least 1.5 standard deviations below the expected level (based on tests of intelligence) on two or more measures of memory, as well as concern regarding a change in cognition, independence of function in daily life, and no evidence of dementia (Albert et al., 2011).

**Materials**

 The study used the same materials from Pilot A.

**Procedure**

The study followed the same procedure as Pilot A, with one exception: the short deadline duration was 1200ms.

**Results**

A 3 x 2 ANOVA as a function of Group (Healthy Younger Adults, Healthy Older Adults, Adults with aMCI) and Deadline (Long, Short) was conducted for recognition accuracy (hit rate minus false alarms).

**Participant Data**

Table 2 reveals participants’ mean (and standard deviation) demographic and neuropsychological test data. There was a significant difference in education between groups; *F*(1, 106) = 18.61, *p* < .001, *ηp2* = .26, in which healthy older adults had significantly more education than younger adults and older adults with aMCI. Additionally, a significant difference was found in age between healthy older adults, and adults with aMCI; *F*(1, 72) = 4.16, *p* = .045, *ηp2* = .06.

 Healthy younger and older adults performed within normal ranges in neuropsychological functioning. Individuals with aMCI performed within normal ranges in the intellectual functioning tasks (SILS Vocabulary, WASI-MR) and the attention and working memory task (Digit Span) but performed below average or borderline across measures of memory. Individuals with aMCI also had an average MoCA score signifying cognitive impairment.

**Table 2.** Mean demographic and neuropsychological test data for Young, Old, and aMCI

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Young* | *Old* | *aMCI* |
| Age | 19.3 (1.8) | 73.0 (4.7) | 75.9 (5.9) |
| Sex (M/F) | 17/18 | 31/27 | 10/6 |
| Education | 13.8 (1.4) | 16.7 (2.3) | 14.6 (3.6) |
| **Cognitive Screening (raw)** |  |  |  |
| MoCA |  | 26.4 (2.0) | 22.9 (3.2) |
| **Intellectual Functioning** |  |  |  |
| Vocabulary | 105.7 (8.8) | 110.0 (7.1) | 104.4 (11.3) |
| Matrix Reasoning | 104.9 (9.6) | 119.6 (9.1) | 113.0 (15.9) |
| **Memory** |  |  |  |
| RAVLT Trials 1-5 Total |  | 108.9 (17.1) | 84.9 (14.4) |
| RAVLT List B Free Recall |  | 101.3 (15.8) | 84.4 (17.0) |
| RAVLT Short Delay Free Recall |  | 100.0 (15.3) | 77.8 (12.6) |
| RAVLT Long Delay Free Recall |  | 104.2 (20.0) | 70.8 (15.1) |
| Logical Memory Learning |  | 111.3 (13.9) | 87.3 (12.9) |
| Logical Memory Delayed Recall |  | 115.6 (12.7) | 89.3 (17.7) |
| **Language** |  |  |  |
| Boston Naming Test |  | 109.1 (17.2) | 94.1 (19.2) |
| F-A-S |  | 106.7 (12.9) | 97.7 (12.6) |
| Animals |  | 101.8 (15.7) | 83.7 (15.4) |
| **Attention and Working Memory** |  |  |  |
| WAIS Digit Span Total |  | 108.5 (14.6) | 104.3 (16.3) |

*Note*. Means are standard scores unless otherwise stated.

**Response Deadline Task Data**

Analyses of recognition accuracy, as evident on Figure 2, revealed significant group differences between younger adults, healthy older adults, and older adults with aMCI; *F*(2, 106) = 8.694, *p* < .001, *ηp2* = .14. Sidak’s post-hoc analyses revealed that older adults with aMCI had significantly lower recognition accuracy in the response deadline task than younger adults (*p* = .010) and older adults (*p* < .001). Additionally, significant effects of deadline were found, *F*(1, 106) = 164.094, *p* < .001, *ηp2* = .61, in which participants had greater recognition accuracy with a long deadline, as opposed to a short deadline. There was also a significant interaction between deadline and group, *F*(2, 106) = 5.718, *p* = .004, *ηp2* = .10. In the long deadline, individuals with aMCI had significantly lower recognition accuracy than younger adults (*p* < .001) and older adults (*p* < .001). Older adults and younger adults did not significantly differ. In the short deadline, individuals with aMCI had significantly lower recognition accuracy than older adults (*p* < .001) but not younger adults (*p* = .235), and younger adults had significantly lower recognition accuracy than older adults (*p* = .034). Numerically, younger adults did have greater recognition accuracy (*M* = .35; *SD* = .03) than older adults with aMCI (*M* = .26; *SD* = .05), but this difference was not significant.

**Figure 2.** Recognition accuracy (hits – false alarms) in the Short and Long deadline, per Group.

****

**Discussion**

 Extending the short deadline to 1200ms reduced the amount of missed responses per group, but younger adults [*M* = 0.34; *SD* = 0.64] still had significantly fewer missed responses than healthy older adults [*M* = 2.14; *SD* = 2.30] (*p* < .001) and older adults with aMCI [*M* = 3.06; *SD* = 2.91] (*p* < .001), with no difference between the latter two groups (*p* = .252), *F*(2, 105) = 12.62, *p* < .001, *ηp2* = .19.

Even though participants were instructed they would be provided a longer response time in the long deadline, across younger participants, 89% of responses remained within 1200ms, and across older participants, 75% of responses remained within 1200ms. Such a finding suggests that many of these responses in the long deadline can be classified as familiarity-based responses and could be considered similar to the short deadline responses. Therefore, we ran this study again but instead force participants to view the stimulus for 3 seconds before providing a response.