UC San Diego
UC San Diego Previously Published Works

Title
Refining CVLT-II recognition discriminability indices to enhance the characterization of recognition memory changes in healthy aging

Permalink
https://escholarship.org/uc/item/76j5r643

Journal
Aging Neuropsychology and Cognition, 25(5)

ISSN
1382-5585

Authors
Graves, Lisa V
Van Etten, Emily J
Holden, Heather M
et al.

Publication Date
2018-09-03

DOI
10.1080/13825585.2017.1372358

Peer reviewed

Refining CVLT-II recognition discriminability indices to enhance the characterization of recognition memory changes in healthy aging

Lisa V. Graves[a,b], **Emily J. Van Etten**[b], **Heather M. Holden**[a,b], **Lisa Delano-Wood**[a,c,d], **Mark W. Bondi**[a,c,d], **Jody Corey-Bloom**[a,e], **Dean C. Delis**[a,d], and **Paul E. Gilbert**[a,b]

[a]Joint Doctoral Program in Clinical Psychology, San Diego State University/University of California San Diego, San Diego/La Jolla, CA, USA

[b]Department of Psychology, San Diego State University, San Diego, CA, USA

[c]Center of Excellence for Stress and Mental Health, VA San Diego Healthcare System, La Jolla, CA, USA

[d]Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, CA, USA

[e]Department of Neurosciences, University of California San Diego, La Jolla, CA, USA

Abstract

The present study examined age-related differences on the four false-positive (FP) error subtypes found on the California Verbal Learning Test-Second Edition yes/no recognition memory trial and the influence of these subtypes on source and novel recognition discriminability (SoRD and NRD, respectively) index calculations. Healthy older (n = 55) adults generally made more FP errors than healthy young adults (n = 57). Accordingly, older adults performed worse than young adults on all SoRD and NRD indices. However, the manner in which FP error subtypes were incorporated into SoRD and NRD index calculations impacted the magnitudes of observed differences between and within the two age groups on SoRD and NRD indices. The present findings underline the importance of examining FP errors in assessments of recognition memory abilities, and using more refined indices of recognition discriminability to further elucidate the nature of age-related recognition memory impairment.

Keywords

Aging; source memory; item memory; recognition discriminability; California Verbal Learning Test

---

The human life expectancy in the United States has continuously risen over the last several decades. Age is currently the greatest known risk factor for neurodegenerative disease. As

the human life expectancy continues to rise, the burden of cognitive decline in older age and the prevalence of dementia due to neurodegenerative disease are expected to increase. The development and use of more refined assessments will be important for enhancing characterizations of the cognitive strengths and weaknesses associated with healthy aging.

Memory loss is one of the most common cognitive issues that arise in older age. Although memory loss, generally speaking, is associated with aging, evidence suggests that not all aspects of memory show an equal rate or magnitude of age-related decline. For example, several studies have shown that the effect of aging is greater on source memory than on item memory (Bayer et al., 2011; Dennis et al., 2008; Glisky & Kong, 2008; Hashtroudi, Johnson, & Chrosniak, 1989; McIntyre & Craik, 1987; Naveh-Benjamin & Craik, 1995; Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991; Spaniol, Madden, & Voss, 2006; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999). Source memory relates to the context from which information was learned or acquired, whereas item memory relates to content of such information regardless of its source. In other words, item memory refers to the ability to remember *what* happened, whereas source memory refers to the ability to remember *where, when, and how* it happened (Dennis et al., 2008). It has been suggested that impaired encoding of contextual information accounts for poorer performance among older adults on source memory tasks (Johnson, Hashtroudi, & Lindsay, 1993). Specifically, age-related dysfunction may result in the inability to engage mnemonic processes for integrating contextual information with item memory during encoding. Moreover, older adults may possess only enough cognitive resources to encode the stimulus itself, at the expense of also encoding contextual information (i.e., are stimulus bound), resulting in poorer recall of such contextual (or source) information (Glisky & Kong, 2008; Johnson et al., 1993).

Memory for the context (i.e., source) and content of an episodic event may rely on different brain regions. Neuroimaging studies and studies involving patients with focal brain lesions have shown that source memory may rely on the functional integrity of both the frontal and temporal lobes (Awipi & Davachi, 2008; Cansino, Maquet, Dolan, & Rugg, 2002; Ekstrom & Bookheimer, 2007; Janowsky, Shimamura, & Squire, 1989; Kirwan, Wixted, & Squire, 2008; Mitchell, Raye, Johnson, & Greene, 2006; Peters, Koch, Schwarz, & Daum, 2007; Peters et al., 2007). Accordingly, age-related pathology of the frontal and temporal regions may account for the source memory decline that is often observed in normal aging (Dennis et al., 2008; Fan, Snodgrass, & Bilder, 2003; Glisky & Kong, 2008; Glisky, Rubin, & Davidson, 2001; Henkel, Johnson, & De Leonardis, 1998; Mitchell et al., 2006). Other studies have indicated that the frontal lobes, in particular, are strongly implicated in source memory (Craik, Morris, Morris, & Loewen, 1990; Fan et al., 2003; Glisky, Polster, & Routhieaux, 1995; Glisky et al., 2001; Janowsky et al., 1989; Schacter, Harbluk, & McLachlin, 1984), whereas the medial temporal lobes may be more involved in item memory (Shimamura & Squire, 1987; Stark & Squire, 2000, 2003).

Recognition memory is a component of declarative memory that involves the ability to recognize previously encountered stimuli. Although not affected to the same degree as recall, recognition memory has been shown to decline with age (Craik & McDowd, 1987; Danckert & Craik, 2013). The original and second editions of the California Verbal Learning Test (CVLT-I and CVLT-II, respectively; Delis, Kramer, Kaplan, & Ober, 1987, 2000) are

widely used in research and clinical settings and have been utilized in efforts to characterize memory function and decline in healthy aging. In general, older adults have been shown to exhibit worse performances relative to young adults on indices of recall and, to a lesser extent, recognition (Delis et al., 1987; Ebert & Anderson, 2009; Kausler, 1994; Turner & Pinkston, 1993; Van der Linden, Philippot, & Heinen, 1997; Woodruff-Pak & Finkbiner, 1995).

Further exploration of more nuanced aspects of recognition memory function may provide additional valuable insight into the cognitive changes that accompany healthy aging. Since the mid to late twentieth century, signal detection theory has been applied in studies of recognition memory as a gold standard for assessing recognition memory function that takes sensitivity and response bias into account. Delis and colleagues included a recognition discriminability (RD) index on the CVLT-I (Delis et al., 1987) and introduced additional subtypes of RD on the CVLT-II (Delis et al., 2000). These CVLT-II indices are calculated using d' (Macmillan & Creelman, 1991). In addition to total RD, the CVLT-II provides an index of source recognition discriminability (SoRD), which captures the ability to distinguish List A target items from List B distractor items on the CVLT-II yes/no recognition memory trial. Thus, the SoRD index, although not a direct measure of source memory per se, taps into aspects of source memory by measuring one's ability to distinguish whether a word was included on List A or List B. Half of the List B distractor items are prototypical, or semantically related to target items, rendering them perhaps more challenging to identify as distractors than the other half of the List B distractor items, which are semantically unrelated (e.g., Baddeley, 1966). Thus, a SoRD index that excludes contributions from FP errors related to prototypical distractors and therefore more specifically captures the ability to distinguish List A target items from List B distractor items that are not semantically related to target items may yield a more refined assessment of SoRD.

The CVLT-II also includes an index of novel recognition discriminability (NRD), which captures the ability to distinguish List A target items from novel (i.e., non-List B) distractor items. Thus, the NRD index represents recognition memory in a more traditional sense, providing a measure of one's ability to distinguish "old" stimuli (i.e., target items) from "new" stimuli (i.e., novel distractor items). Half of the novel distractor items are prototypical, or semantically related to target items, rendering them more challenging to identify as distractors than the other half of the novel distractor items, which are semantically unrelated. Thus, a NRD index that excludes contributions from FP errors related to prototypical distractors and therefore more specifically captures the ability to distinguish List A target words from novel distractor items that are semantically unrelated to target items may provide a more refined assessment of NRD.

In addition to SoRD and NRD, the CVLT-II provides a third subtype of RD called semantic recognition discriminability that captures the ability to distinguish List A target items from distractor items that are semantically related to target items, including those that are from List B as well as those that are novel (Delis et al., 2000). In contrast, the SoRD and NRD indices reflect the ability to distinguish targets from List B and novel distractors,

respectively, without parsing the contributions of prototypical, or semantically related distractors from those of semantically unrelated distractors.

On that premise, the extent to which between- and within-group differences in SoRD and NRD performances may be influenced by the degree of semantic association between targets and distractors found on the CVLT-II yes/no recognition memory trial has not been explored. Thus, the present study has two main objectives. First, between- and within-group differences in FP errors in each of the four subtypes that are found on the CVLT-II yes/no recognition memory trial (prototypical List B, unrelated List B, prototypical novel, and unrelated novel) will be examined in healthy older ($n = 55$) and young ($n = 57$) adults. Second, between- and within-group differences in d' scores were examined on three variations of the SoRD and NRD indices: (1) original SoRD and NRD (which include both prototypical and semantically unrelated List B and novel distractors in d' calculations), (2) SoRD-prototypical and NRD-prototypical (which include prototypical List B and novel distractors only in d' calculations), and (3) SoRD-unrelated and NRD-unrelated (which include semantically unrelated List B and novel distractors only in d' calculations). Older adults are expected to make more FP errors than young adults, although it is hypothesized that the two age groups may exhibit different patterns of FP errors across the four subtypes. Additionally, older adults are expected to perform worse than young adults on all SoRD and NRD indices, albeit to a lesser extent on indices that exclude prototypical distractors. In particular, the SoRD-unrelated and NRD-unrelated indices are expected to be associated with smaller group differences than the SoRD and NRD indices. Furthermore, both age groups are expected to demonstrate better performances on SoRD-unrelated and NRD-unrelated indices than on SoRD and NRD indices, and older adults are generally expected to exhibit worse performances on SoRD than on NRD. Findings from the present study may help to elucidate the nature of recognition memory function and changes in healthy aging with the use of more refined measures of RD.

Method

Participants

Study participants included 57 healthy young adults (18–25 years of age) and 55 healthy older adults (65 years of age or older). Older adults were characterized as cognitively healthy based on Dementia Rating Scale-2 (DRS-2; Jurica, Leitten, & Mattis, 2001) scores (130 or above). Exclusionary criteria for all healthy adult participants included the following: a diagnosis of any neurological disorder, a diagnosis of any major medical condition (e.g., cancer), a diagnosis of any psychiatric disorder (with the exception of a mood disorder, for which any current symptoms must be well managed), a history of traumatic brain injury, and a history of substance abuse. All participants provided informed written consent and the study was approved by the Institutional Review Boards of San Diego State University (SDSU) and/or the University of California, San Diego (UCSD).

Healthy young adults were recruited from the San Diego community by the Center for Healthy Aging and Neurodegenerative Disease Research (CHANDR) at SDSU and the Huntington's Disease Clinical Research Program (HDCRP) at UCSD. Healthy older adults were recruited from the San Diego community by CHANDR at SDSU, the Normal Aging

Laboratory at UCSD, and the HDCRP at UCSD. Participants were administered a standardized battery of neuropsychological tests by trained research assistants or psychometrists. CVLT-II data from the subset of healthy older adults recruited by the Normal Aging Laboratory were extracted from an archival database that included data from a larger battery of neuropsychological tests administered at the Shiley-Marcos Alzheimer's Disease Research Center in La Jolla and the Veterans Affairs San Diego Healthcare System.

CVLT-II and RD indices

The CVLT-II was administered using standard procedures outlined by Delis and colleagues (2000). The CVLT-II is a list-learning test that provides a multitude of verbal learning and memory indices, including immediate recall, free and cued recall over short and long delays, and recognition memory. The RD indices that were of primary interest in the present study were generated using variables derived from the yes/no recognition memory trial on the CVLT-II. Short- and long-delay tests of recall were separated by an interval of approximately 20 min, during which other nonverbal neuropsychological measures were administered. CVLT-II data were scored using CVLT-II scoring software (Delis & Fridlund, 2000). Raw scores on hits, the four FP error subtypes [prototypical List B (used in calculating the SoRD-prototypical index), unrelated List B (used in calculating the SoRD-unrelated index), prototypical novel (used in calculating the NRD-prototypical index), and unrelated novel (used in calculating the NRD-unrelated index)], and the six RD indices (SoRD, SoRD-prototypical, SoRD-unrelated, NRD, NRD-prototypical, and NRD-unrelated) were examined.

SoRD and NRD indices are calculated using the following formulas (Delis et al., 2000):

1. SoRD (d') = z(hits) − z(FP errors associated with prototypical List B distractors + FP errors associated with semantically unrelated List B distractors).

2. NRD (d') = z(hits) − z(FP errors associated with prototypical novel distractors + FP errors associated with semantically unrelated novel distractors).

SoRD-prototypical, SoRD-unrelated, NRD-prototypical, and NRD-unrelated indices were generated using the following formulas:

3. SoRD-prototypical (d') = z(hits) − z(FP errors associated with prototypical List B distractor items only).

4. SoRD-unrelated (d') = z(hits) − z(FP errors associated with unrelated List B distractor items only).

5. NRD-prototypical (d') = z(hits) − z(FP errors associated with prototypical novel distractor items only).

6. NRD-unrelated (d') = z(hits) − z(FP errors associated with unrelated novel distractor items only).

Raw d' scores are computed by calculating inverse proportions of hits and respective FP errors and subtracting respective FP error rates from hit rates (see Macmillan & Creelman, 1991).

Statistical analyses

Analyses were conducted in the Statistical Package for the Social Sciences Version 24. Prior to examining age group differences in hits, the four FP error subtypes and the six RD indices of interest, chi-square analyses, and one-way analysis of variance (ANOVA) tests were conducted to determine whether gender and education, respectively, were significant predictors of the outcome variables (hits, FP errors, RD indices). Gender and education were not significant predictors of the particular outcome variables of interest in the present study and therefore were not controlled for in the primary analyses.

Shapiro–Wilk tests of normality revealed that all outcome variables were non-normally distributed ($ps < .05$). Thus, nonparametric analyses were conducted to address the aims of the present study.

Analysis of hits—A Mann–Whitney U test was conducted to examine the age group difference in the number of hits on the CVLT-II yes/no recognition memory trial.

Analyses of FP error subtypes—Due to the substantial number of zero FP errors across subtypes and individuals, separate chi-square analyses were conducted to examine differences between the two age groups in the number of individuals who made zero FP errors versus one or more FP errors in each of the four subtypes. Additionally, separate chi-square tests of independence were conducted to make pairwise comparisons of the four FP error subtypes within each age group. Effect size values (r) were calculated to quantify and compare the magnitudes of significant between- and within-group differences in FP error subtypes [$r = \sqrt{(\chi^2/N)}$].

Analyses of RD indices—Six separate Mann–Whitney U tests were conducted to examine age group differences in d' scores on the six RD indices of interest (SoRD, SoRD-prototypical, SoRD-unrelated, NRD, NRD-prototypical, and NRD-unrelated). Additionally, two separate Friedman tests were conducted to examine the effect of RD index type on d' scores within each group. If a significant omnibus effect of RD index type was observed, nine follow-up Wilcoxon signed-rank tests were conducted to make the following pairwise comparisons within a particular age group: (1) SoRD vs. NRD, (2) SoRD-prototypical vs. NRD-prototypical, (3) SoRD-unrelated vs. NRD-unrelated, (4) SoRD vs. SoRD-prototypical, (5) SoRD vs. SoRD-unrelated, (6) SoRD-prototypical vs. SoRD-unrelated, (7) NRD vs. NRD-prototypical, (8) NRD vs. NRD-unrelated, and (9) NRD-prototypical vs. NRD-unrelated. Effect size values (r) were calculated to quantify and compare the magnitudes of significant between- and within-group differences in d' scores on RD indices ($r = Z/\sqrt{N}$).

False discovery rate adjustment—Adjustments for a false discovery rate (FDR) of .05 (see Benjamini & Hochberg, 1995) were applied in the analyses of between- and within-group differences on FP error subtypes and RD indices. Original p-values are presented in the study tables, and asterisks indicate which p-values retained significance following FDR adjustments.

Results

Demographic information

A chi-square analysis revealed no difference between the older (49.09% women) and young (57.89% women) adult groups in their proportions of men and women, χ^2 (2, $N = 112$) = 0.87, $p = .45$. A one-way ANOVA revealed that older adults ($M = 16.36$, $SD = 2.08$) completed more years of education than young adults ($M = 14.28$, $SD = 2.21$), $F(1, 110)$ = 31.64, $p < .001$. All older adults had DRS-2 scores of 130 or higher ($M = 140.62$, $SD = 2.96$).

Analysis of hits

A Mann–Whitney U test revealed that older adults (*mean rank* = 50.15, *sum of ranks* = 2758.50) had significantly fewer hits on the CVLT-II yes/no recognition memory trial than young adults (*mean rank* = 62.62, *sum of ranks* = 3569.50), $U = 1218.50$, $p < .05$.

Analyses of FP error subtypes

Age group differences in FP error subtypes—Descriptive and inferential statistics for age group differences in the number of individuals who made zero versus one or more FP errors in each of the four subtypes are provided in Table 1. Proportions of older and young adults who made one or more FP errors in each of the four subtypes are illustrated in Figure 1. Chi-square analyses revealed that the extent to which the number of individuals who made zero FP errors was higher than the number of individuals who made one or more FP errors was smaller in the older adult group than in the young adult group in three of the four FP error subtypes: prototypical List B, unrelated List B, and prototypical novel (i.e., the proportion of individuals who made one or more FP errors was larger in the older adult group than in the young adult group in the three aforementioned subtypes). No age group difference in the extent to which the number of individuals who made zero FP errors was higher than the number of individuals who made one or more FP errors was observed in the unrelated novel subtype.

Within-group differences in FP error subtypes—Inferential statistics for comparisons within each age group in the number of individuals who made zero versus one or more FP errors across subtypes are provided in Table 2. Proportions of individuals who made one or more FP errors across subtypes within each age group are illustrated in Figure 2. Chi-square analyses revealed different patterns of FP errors within the older adult, χ^2 (3, $N = 55$) = 30.48, $p < .001$, and young adult, χ^2 (3, $N = 57$) = 14.15, $p < .01$, groups. In the older adult group, the proportion of individuals who made one or more FP errors was (1) greater for the prototypical List B and prototypical novel subtypes than the unrelated List B and unrelated novel subtypes, respectively, (2) greater for the unrelated List B subtype than the unrelated novel subtype, (3) greater for the prototypical novel subtype than the unrelated List B subtype, and (4) greater for the prototypical List B subtype than the unrelated novel subtype. In the young adult group, the proportion of individuals who made one or more FP errors was greater for the prototypical novel subtype than the unrelated novel subtype; however, no other comparisons within the young adult group were significant.

Analyses of RD indices

Age group differences on RD indices—Mean and standard deviation values as well as 25th, 50th (median), and 75th percentile values of the older and young adult groups on all six RD indices are provided in Table 3. Descriptive and inferential statistics for age group differences on RD indices are provided in Table 4. Mann–Whitney U tests revealed that older adults performed significantly worse than young adults on all six RD indices.

Comparisons of effect sizes for age group differences on RD indices—Although analyses revealed that older adults performed significantly worse than young adults on all RD indices, effect sizes associated with the observed age group differences on RD indices were compared to elucidate the extent to which incorporating FP errors associated with prototypical distractors only, unrelated distractors only, or both prototypical and unrelated distractors in calculations of SoRD and NRD scores impacts observed age group differences. The effect size associated with the age group difference on the SoRD-unrelated index (List A targets vs. unrelated List B distractors only) was 24.24% smaller than the effect size associated with the age group differences on the SoRD (List A targets vs. all List B distractors) and SoRD-prototypical (List A targets vs. prototypical List B distractors only) indices, which were comparable. Additionally, the effect size associated with the age group difference on the NRD-unrelated index (List A targets vs. unrelated novel distractors only) was 38.89% smaller than the effect size associated with the age group difference on the NRD index (List A targets vs. all novel distractors), but was comparable to the effect size associated with the age group difference on the NRD-prototypical index (List A targets vs. prototypical novel distractors only). In sum, the extent to which older adults performed worse than young adults was smaller on the SoRD-unrelated and NRD-unrelated indices than on the SoRD and NRD indices.

The effect size associated with the age group difference on the SoRD index was 8.33% smaller than the effect size associated with the age group difference on the NRD index. In contrast, the effect size associated with the age group difference on the SoRD-prototypical index was 50.00% larger than the effect size associated with the age group difference on the NRD-prototypical index. Finally, the effect size associated with the age group difference on the SoRD-unrelated index was 13.64% larger than the effect size associated with the age group difference on the NRD-unrelated index. Thus, a larger age group difference on SoRD relative to NRD was observed on SoRD and NRD indices that included either FP errors associated with prototypical distractors only (i.e., SoRD-prototypical and NRD-prototypical) or semantically unrelated distractors only (i.e., SoRD-unrelated and NRD-unrelated), although the difference was substantially smaller in the context of the latter indices.

Within-group differences on RD indices—Descriptive and inferential statistics for within-group differences on RD indices are provided in Table 5. Friedman tests revealed a significant effect of RD index type within both the older adult, χ^2 (5, $N = 55$) = 104.77, $p < .001$, and young adult, χ^2 (5, $N = 57$) = 161.39, $p < .001$, groups. Wilcoxon signed-rank tests revealed different patterns of performances on RD indices within the older and young adult groups.

In the older adult group, scores were higher on the SoRD-unrelated index than on the SoRD and SoRD-prototypical indices; however, scores were comparable on the latter two indices. Additionally, scores were higher on the NRD-unrelated index than on the NRD and NRD-prototypical indices, although scores were comparable on the latter two indices (after an FDR adjustment). Furthermore, in the older adult group, scores were comparable on the SoRD and NRD indices, on the SoRD-prototypical and NRD-prototypical indices, and on the SoRD-unrelated and NRD-unrelated indices (after an FDR adjustment).

Performances on RD indices in the young adult group largely mirrored the pattern of performances that was observed in the older adult group. For example, scores were higher on the SoRD-unrelated index than on the SoRD index; however, in contrast to the older adult group, scores also were higher on the SoRD-prototypical index than on the SoRD index, and scores on the SoRD-prototypical and SoRD-unrelated indices were comparable. Additionally, scores were higher on the NRD-unrelated index than on the NRD and NRD-prototypical indices, although scores were comparable on the latter two indices. Furthermore, in the young adult group, scores were comparable on the SoRD and NRD indices, on the SoRD-prototypical and NRD-prototypical indices, and on the SoRD-unrelated and NRD-unrelated indices.

Discussion

The present study demonstrated that FP errors associated with prototypical distractors substantially influence calculations of SoRD and NRD scores on the CVLT-II yes/no recognition memory trial. The examination of age group differences in FP errors revealed that, compared to young adults, a greater proportion of older adults made FP errors associated with prototypical List B, prototypical novel, and unrelated List B distractors. However, the two age groups did not differ in proportions of individuals who made FP errors associated with unrelated novel distractors. This finding is expected given that these items are generally conceptualized as the least challenging to identify as distractors or "non-targets" due to being both novel (i.e., were not presented at any point during task administration) and semantically unrelated – and therefore less similar – to target items.

Analyses also demonstrated that the two age groups yielded different patterns in proportions of individuals who made FP errors across the four subtypes. In particular, the pattern of FP errors within the older adult group suggests an age-related vulnerability to the effect of semantic interference from prototypical items on yes/no recognition testing, over and above an effect of source interference from List B items. In the older adult group, there was a greater proportion of individuals who made FP errors associated with distractors that are prototypical, or semantically related to targets (regardless of whether the items were from List B or novel) than there was of those who made FP errors associated with non-prototypical, or semantically unrelated distractors. Moreover, there was a greater proportion of individuals who made FP errors associated with prototypical novel distractors (which present only semantic interference) than there was of those who made FP errors associated with unrelated List B distractors (which present only source interference). Furthermore, the proportions of individuals who made FP errors associated with prototypical List B distractors (which present *both* semantic and source interference) and prototypical novel

distractors (which present *only semantic* interference) were comparable. These findings may imply that healthy older adults are even more vulnerable to semantic interference than source interference (i.e., experience even greater difficulty in identifying prototypical items as distractors, than in identifying List B items as distractors as a result of age-related source memory impairment) on the CVLT-II yes/no recognition memory trial. However, in the analysis of semantically unrelated distractors, there was a greater proportion of individuals who made FP errors associated with items that were from List B than those that were novel, which is not surprising given the research literature on age-related source memory impairment. Taken together, this set of findings regarding FP errors in the older adult group suggests that (1) older adults are particularly susceptible to inaccurately endorsing prototypical distractors over and above experiencing difficulty in identifying List B items as distractors and (2) in the context of semantically unrelated distractors only, continue to exhibit difficulty in identifying List B items as distractors. These findings provide more evidence of age-related source memory impairment as well as highlight that prototypical items, by introducing semantic interference, are an additional source of confusion or difficulty for healthy older adults on yes/no recognition memory testing.

In the young adult group, there was a greater proportion of individuals who made FP errors associated with prototypical distractors than there was of those who made FP errors associated with semantically unrelated distractors, only with regard to novel distractor items. Thus, young adults also may be prone to inaccurately endorsing prototypical distractors, albeit to a lesser extent than older adults based on an examination of effect sizes (see Table 2), although, in contrast to older adults, they are less likely to experience difficulty in aspects of yes/no recognition memory testing that rely on source memory (i.e., identifying List B items as distractors).

The examination of age group differences on SoRD, SoRD-prototypical, SoRD-unrelated, NRD, NRD-prototypical, and NRD-unrelated indices revealed that older adults performed significantly worse than young adults on all indices. Moreover, effect sizes associated with the observed age group differences on RD indices were compared to elucidate the extent to which incorporating FP errors associated with prototypical distractors only, unrelated distractors only, or both prototypical and unrelated distractors in calculations of SoRD and NRD scores impacted observed age group differences. A particular emphasis was made on comparing the degree to which older and young adults differed on new, more refined SoRD and NRD indices that exclude FP errors associated with prototypical distractors (i.e., SoRD-unrelated and NRD-unrelated) relative to existing CVLT-II SoRD and NRD indices that include FP errors associated with both prototypical and semantically unrelated distractors. As expected, the effect sizes associated with age group differences on SoRD-unrelated and NRD-unrelated indices were smaller than the effect sizes associated with age group differences on SoRD and NRD indices, respectively. The reduction in age group differences is likely driven by the notion that older adults showed greater improvements relative to young adults on indices that exclude FP errors associated with prototypical distractors, based on an examination of effect sizes. Analyses also revealed that the effect size associated with the age group difference on the SoRD index was smaller than the effect size associated with the age group difference on the NRD index. In contrast, the effect size associated with the age group difference on the SoRD-unrelated index was larger than the effect size associated

with the age group difference on the NRD-unrelated index (this pattern was even more evident in the context of SoRD-prototypical and NRD-prototypical indices). In sum, this set of findings indicates that, by excluding contributions from FP errors associated with prototypical distractor items that are semantically related to target items in the calculation of SoRD and NRD scores, (1) age group differences on SoRD and NRD are smaller in magnitude and (2) the extent to which older adults perform worse than young adults is greater on SoRD than on NRD, which further supports the notion that, relative to item memory, source memory is particularly vulnerable to age-related decline (Bayer et al., 2011; Dennis et al., 2008; Glisky & Kong, 2008; Hashtroudi et al., 1989; McIntyre & Craik, 1987; Naveh-Benjamin & Craik, 1995; Schacter et al., 1991; Spaniol et al., 2006; Trott et al., 1999). A possible limitation of these findings is that the older adult group in the study sample was relatively well educated and may not fully represent the general population of cognitively healthy older adults. However, it is reasonable to suspect that observed age group differences on RD indices would be larger in a sample of individuals with less cognitive reserve. Moreover, the present findings highlight the potential for these refined RD indices to demonstrate clinical utility in the assessment and characterization of recognition memory deficits in more cognitively impaired populations, such as individuals with neurodegenerative disease.

Analyses revealed different patterns of within-group differences on RD indices across the two age groups. In the older adult group, performances were higher on the SoRD-unrelated index than on the SoRD and SoRD-prototypical indices, whereas performances were comparable on the latter two indices. These findings suggest a significant influence of FP errors related to prototypical List B distractors on SoRD performances in older adults, and further highlight the cumulative effects of source and semantic interference on increasing the difficulty of identifying distractor items for older adults. Similarly, performances were higher on the NRD-unrelated index than on the NRD and NRD-prototypical indices, whereas performances were comparable on the latter two indices, suggesting a significant influence of FP errors related to prototypical novel distractors on NRD performances in older adults, and further highlighting the impact of semantic interference on yes/no recognition testing in older adults. Performances on the SoRD-prototypical and NRD-prototypical indices (and on the SoRD and NRD indices) were comparable, which is not surprising given that the proportions of older adults who made FP errors associated with prototypical List B and prototypical novel distractors were comparable. However, performances on the SoRD-unrelated and NRD-unrelated indices also were comparable despite the observation that the proportion of older adults who made FP errors associated with unrelated List B distractors was greater than the proportion of older adults who made FP errors associated with unrelated novel distractors. Nonetheless, this set of findings collectively suggests that older adults do benefit from the exclusion of FP errors associated with prototypical distractor items in the calculation of d' scores for SoRD and NRD. Moreover, the findings suggest that disproportionate source memory impairments in older adults may be more evident in the close examination of FP errors, rather than through comparisons of scores on SoRD indices relative to NRD indices.

In the young adult group, performances were higher on the SoRD-unrelated and NRD-unrelated indices than on the SoRD and NRD indices, respectively. Moreover, performances

were comparable on the SoRD and NRD indices, on the SoRD-prototypical and NRD-prototypical indices, and on the SoRD-unrelated and NRD-unrelated indices. Taken together, these results suggest that young adults also do benefit from the exclusion of FP errors associated with prototypical distractor items in the calculation of d' scores for SoRD and NRD. Nonetheless, disproportionate source or novel recognition memory impairments among young adults were not observed in analyses of FP errors or RD indices, which is not surprising given that relative weaknesses in source or item memory are not typically observed in young adulthood.

Overall, the present findings yield evidence for improved performances among both older and young adults on SoRD and NRD indices that exclude FP errors associated with prototypical distractor items in the calculation of d' scores. Moreover, the present findings highlight the important role of FP errors in the assessment of RD and efforts to characterize recognition memory function and changes in healthy aging.

Conclusion

The present study examined the impact of different FP error subtypes on assessments of SoRD and NRD using the CVLT-II in a cognitively healthy sample, and the degree of age-related differences on SoRD and NRD indices that exclude contributions from FP errors associated with prototypical distractors (i.e., SoRD-unrelated and NRD-unrelated) relative to age-related differences on original SoRD and NRD indices that include FP errors associated with both prototypical and semantically unrelated distractors. Both age groups demonstrated better performances on SoRD-unrelated and NRD-unrelated indices than on SoRD and NRD indices, respectively. Although older adults performed worse than young adults on all RD indices, age group differences were smaller in magnitude on the more refined SoRD-unrelated and NRD-unrelated indices relative to original SoRD and NRD indices, although older adults were shown to perform disproportionately worse than young adults on SoRD in the context of refined indices. Although CVLT-II indices of SoRD and NRD in their current form can reliably demonstrate age-related differences on these aspects of recognition memory function (i.e., those pertaining to source and item memory), the refined indices utilized in the present study may be used to further elucidate the extent towhich healthy older and young adults differ on these particular constructs. Furthermore, the present findings highlight the potential for these refined RD indices to exhibit clinical utility in improving assessments and characterizations of recognition memory deficits in more cognitively impaired populations.

Acknowledgments

References

Awipi T, Davachi L. Content-specific source encoding in the human medial temporal lobe. Journal of Experimental Psychology: Learning Memory and Cognition. 2008; 34(4):769–779. DOI: 10.1037/0278-7393.34.4.769

Baddeley AD. The influence of acoustic and semantic similarity on long-term memory for word sequences. Quarterly Journal of Experimental Psychology. 1966; 18(4):302–309. DOI: 10.1080/14640746608400047 [PubMed: 5956072]

Bayer ZC, Hernandez RJ, Morris AM, Salomonczyk D, Pirogovsky E, Gilbert PE. Age-related source memory deficits persist despite superior item memory. Experimental Aging Research. 2011; 37(4): 473–480. DOI: 10.1080/0361073X.2011.590760 [PubMed: 21800975]

Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995; 57(1): 289–300. DOI: 10.2307/2346101

Cansino S, Maquet P, Dolan RJ, Rugg MD. Brain activity underlying encoding and retrieval of source memory. Cerebral Cortex. 2002; 12(10):1048–1056. DOI: 10.1093/cercor/12.10.1048 [PubMed: 12217968]

Craik FI, McDowd JM. Age differences in recall and recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1987; 13(3):474–479. DOI: 10.1037/0278-7393.13.3.474

Craik FI, Morris LW, Morris RG, Loewen ER. Relations between source amnesia and frontal lobe functioning in older adults. Psychology and Aging. 1990; 5:148–151. DOI: 10.1037/0882-7974.5.1.148 [PubMed: 2317296]

Danckert SL, Craik FI. Does aging affect recall more than recognition memory? Psychology and Aging. 2013; 28(4):902–909. DOI: 10.1037/a0033263 [PubMed: 23978011]

Delis DC, Fridlund AJ. CVLT-II comprehensive scoring system and computerized report. San Antonio, TX: The Psychological Corporation; 2000.

Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test: Adult version. Manual. San Antonio, TX: The Psychological Corporation; 1987.

Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test–second edition: Adult version. Manual. San Antonio, TX: The Psychological Corporation; 2000.

Dennis NA, Hayes SM, Prince SE, Madden DJ, Huettel SA, Cabeza R. Effects of aging on the neural correlates of successful item and source memory encoding. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2008; 34(4):791–808. DOI: 10.1037/0278-7393.34.4.791

Ebert PL, Anderson ND. Proactive and retroactive interference in young adults, healthy older adults, and older adults with amnestic mild cognitive impairment. Journal of the International Neuropsychological Society. 2009; 15:83–93. DOI: 10.1017/S1355617708090115 [PubMed: 19128531]

Ekstrom AD, Bookheimer SY. Spatial and temporal episodic memory retrieval recruit dissociable functional networks in the human brain. Learning and Memory. 2007; 14(10):645–654. DOI: 10.1101/lm.575107 [PubMed: 17893237]

Fan J, Snodgrass J, Bilder RM. Functional magnetic resonance imaging of source versus item memory. Neuroreport. 2003; 14:2275–2281. DOI: 10.1097/00001756-200312020-00028 [PubMed: 14625462]

Glisky EL, Kong LL. Do young and older adults rely on different processes in source memory tasks? A neuropsychological study. Journal of Experimental Psychology: Learning, Memory, & Cognition. 2008; 34(4):809–822. DOI: 10.1037/0278-7393.34.4.809

Glisky EL, Polster MR, Routhieaux BC. Double dissociation between item and source memory. Neuropsychology. 1995; 9:229–235. DOI: 10.1037/0894-4105.9.2.229

Glisky EL, Rubin SR, Davidson PSR. Source memory in older adults: An encoding or retrieval problem? Journal of Experimental Psychology: Learning, Memory and Cognition. 2001; 27(5): 1131–1146. DOI: 10.1037/0278-7393.27.5.1131

Hashtroudi S, Johnson MK, Chrosniak LD. Aging and source monitoring. Psychology and Aging. 1989; 4:106–112. DOI: 10.1037/0882-7974.4.1.106 [PubMed: 2803603]

Henkel LA, Johnson MK, De Leonardis DM. Aging and source monitoring: Cognitive processes and neuropsychological correlates. Journal of Experimental Psychology: General. 1998; 127(3):251–268. DOI: 10.1037/0096-3445.127.3.251 [PubMed: 9742716]

Janowsky JS, Shimamura AP, Squire LR. Source memory impairment in patients with frontal lobe lesions. Neuropsychologia. 1989; 27(8):1043–1056. DOI: 10.1016/0028-3932(89)90184-X [PubMed: 2797412]

Johnson MK, Hashtroudi S, Lindsay DS. Source monitoring. Psychological Bulletin. 1993; 114:3–8. DOI: 10.1037/0033-2909.114.1.3 [PubMed: 8346328]

Jurica PJ, Leitten S, Mattis S. Dementia Rating Scale-2: Professional manual. Lutz, FL: Psychological Assessment Resources; 2001.

Kausler DH. Learning and memory in normal aging. San Diego, CA: Academic Press; 1994.

Kirwan CB, Wixted JT, Squire LR. Activity in the medial temporal lobe predicts memory strength, whereas activity in the prefrontal cortex predicts recollection. Journal of Neuroscience. 2008; 28(42):10541–10548. DOI: 10.1523/JNEUROSCI.3456-08.2008 [PubMed: 18923030]

Macmillan NA, Creelman DC. Detection theory: A user's guide. New York, NY: Cambridge University Press; 1991.

McIntyre JS, Craik FIM. Age differences in memory for item and source information. Canadian Journal of Psychology. 1987; 41(2):175–192. DOI: 10.1037/h0084154 [PubMed: 3502895]

Mitchell KJ, Raye CL, Johnson MK, Greene EJ. An fMRI investigation of short-term source memory in young and older adults. Neuroimage. 2006; 30(2):627–633. DOI: 10.1016/j.neuroimage. 2005.09.039 [PubMed: 16256377]

Naveh-Benjamin M, Craik FI. Memory for context and its use in item memory: Comparisons of younger and older persons. Psychology and Aging. 1995; 10(2):284–293. DOI: 10.1037/0882-7974.10.2.284 [PubMed: 7662187]

Peters J, Koch B, Schwarz M, Daum I. Domain-specific impairment of source memory following a right posterior medial temporal lobe lesion. Hippocampus. 2007; 17(7):505–509. DOI: 10.1002/hipo.20297 [PubMed: 17476681]

Schacter DL, Harbluk JL, McLachlin DR. Retrieval without recollection: An experimental analysis of source amnesia. Journal of Verbal Learning and Verbal Behavior. 1984; 23(5):593–611. DOI: 10.1016/S0022-5371(84)90373-6

Schacter DL, Kaszniak AW, Kihlstrom JF, Valdiserri M. The relation between source memory and aging. Psychology and Aging. 1991; 6(4):559–568. DOI: 10.1037/0882-7974.6.4.559 [PubMed: 1777144]

Shimamura AP, Squire LR. A neuropsychological study of fact memory and source amnesia. Journal Experimental Psychology: Learning, Memory, and Cognition. 1987; 13(3):464–473. DOI: 10.1037/0278-7373.13.3.464

Spaniol J, Madden DJ, Voss A. A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. Journal of Experimental Psychology: Learning, Memory, & Cognition. 2006; 32:101–117. DOI: 10.1037/0278-7393.32.1.101

Stark CE, Squire LR. Functional magnetic resonance imaging (fMRI) activity in the hippocampal region during recognition memory. Journal of Neuroscience. 2000; 20(20):7776–7781. [PubMed: 11027241]

Stark CE, Squire LR. Hippocampal damage equally impairs memory for single items and memory for conjunctions. Hippocampus. 2003; 13(2):281–292. DOI: 10.1002/hipo.10085 [PubMed: 12699335]

Trott CT, Friedman D, Ritter W, Fabiani M, Snodgrass JG. Episodic priming and memory for temporal source: Event related potentials reveal age-related differences in prefrontal functioning. Psychology and Aging. 1999; 14(3):390–413. DOI: 10.1037/0882-7974.14.3.390 [PubMed: 10509695]

Turner ML, Pinkston RS. Effects of a memory and aging workshop on negative beliefs of memory loss in the elderly. Educational Gerontology. 1993; 19(5):359–373. DOI: 10.1080/0360127930190501

Van der Linden M, Philippot P, Heinen P. Effect of age, education and verbal efficiency on memory performance and memory self-assessment. Archives De Psychologie. 1997; 65(254):171–185.

Woodruff-Pak DS, Finkbiner RG. Larger nondeclarative than declarative deficits in learning and memory in human aging. Psychology and Aging. 1995; 10(3):416–426. DOI: 10.1037/0882-7974.10.3.416 [PubMed: 8527062]
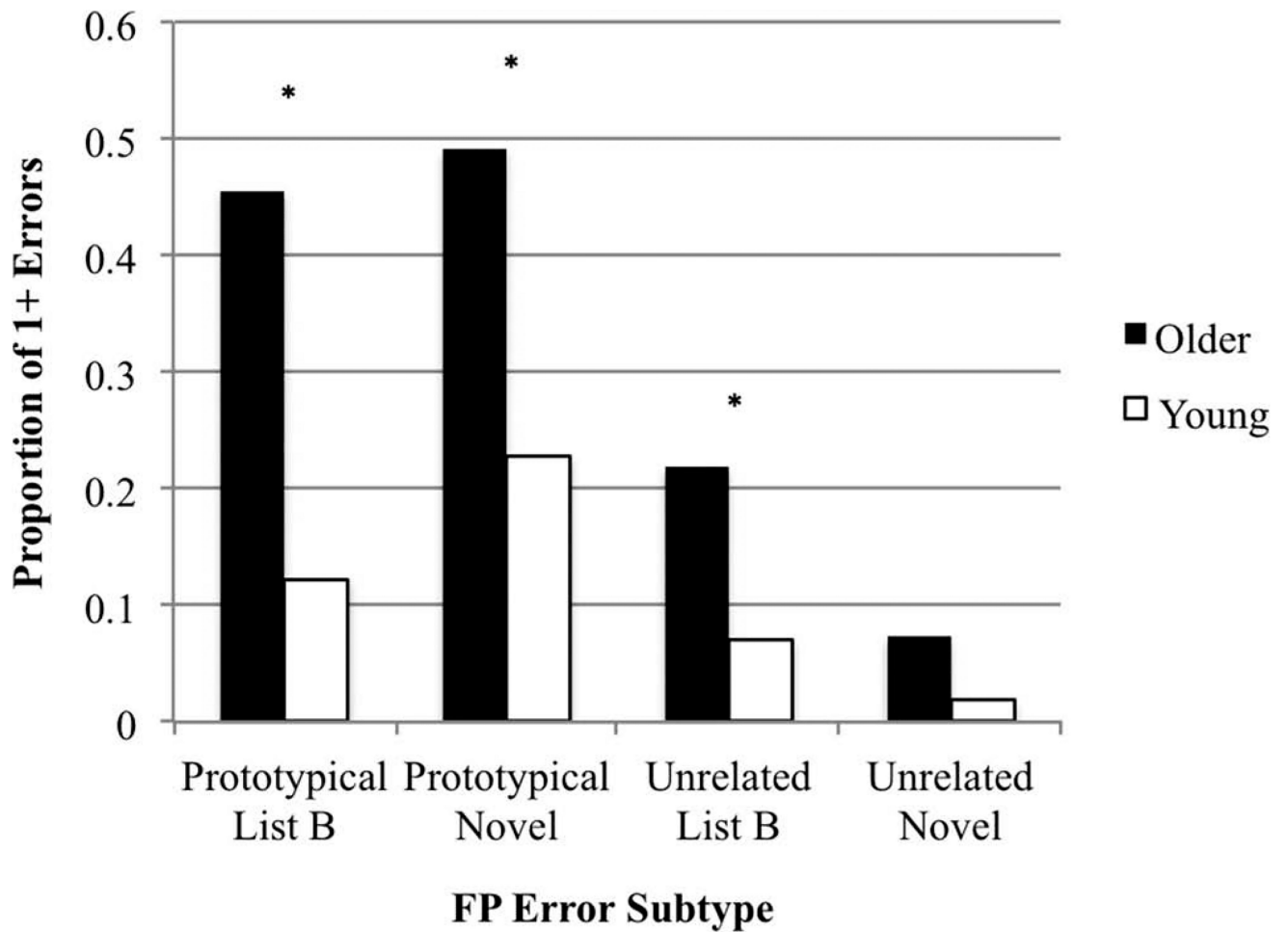
Figure 1.
Proportions of older and young adults who made one or more (1+) FP errors in each of the four subtypes: prototypical List B, prototypical novel, unrelated List B, and unrelated novel. Asterisks (*) indicate significant group differences.
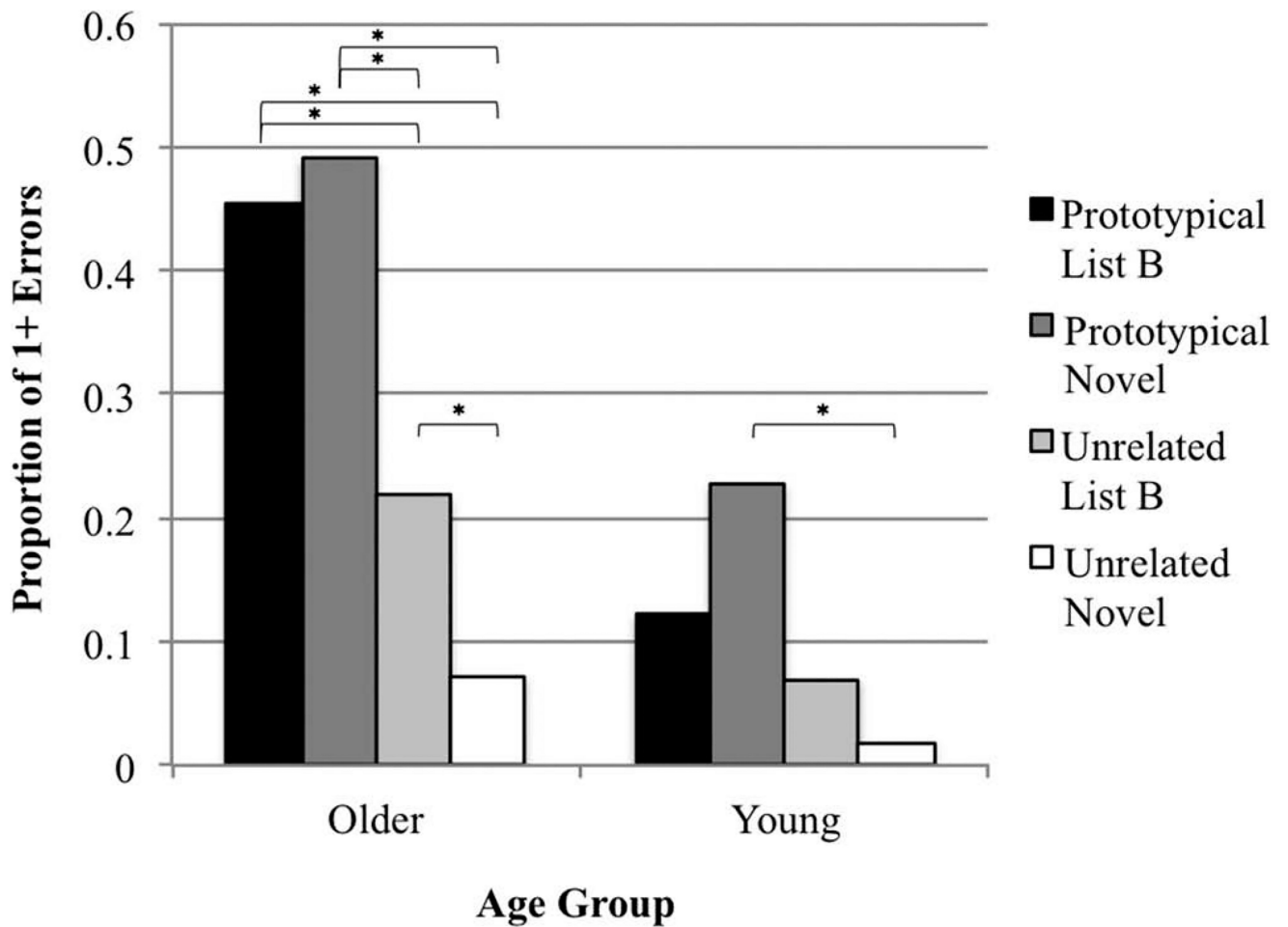
Figure 2.
Proportions of individuals who made one or more (1+) FP errors across subtypes
(prototypical List B, prototypical novel, unrelated List B, and unrelated novel) within each
age group. Asterisks (*) indicate significant pairwise comparisons within groups.

Table 1

Age group differences in the number of individuals who made zero (0 FP) versus one or more (1+ FP) FP errors in each of the four subtypes, based on chi-square analyses.

FP error subtype	Older			Young			χ^2	p	r
	0 FP	1+ FP	Proportion of 1+ FP	0 FP	1+ FP	Proportion of 1+ FP			
Prototypical List B	30	25	.455	50	7	.123	15.09	<.001[*]	.37
Prototypical Novel	28	27	.491	44	13	.228	8.42	<.01[*]	.27
Unrelated List B	43	12	.218	53	4	.070	5.01	<.05[*]	.21
Unrelated Novel	51	4	.073	56	1	.018	2.00	.202	.13

[*]
p-Value retains significance following FDR adjustment.

Table 2

Comparisons within each age group in the number of individuals who made zero versus one or more FP errors across subtypes, based on chi-square tests of independence.

Comparison	Older			Young		
	χ^2	p	r	χ^2	p	r
Prototypical List B vs. Unrelated List B	6.88	<.01 *	.25	0.91	.341	.09
Prototypical List B vs. Prototypical Novel	0.15	.702	.04	2.18	.140	.14
Prototypical List B vs. Unrelated Novel	20.65	<.001 *	.43	4.84	<.05	.21
Unrelated List B vs. Prototypical Novel	8.94	<.01 *	.28	5.60	<.05	.22
Unrelated List B vs. Unrelated Novel	4.68	<.05 *	.20	1.88	.170	.13
Prototypical Novel vs. Unrelated Novel	23.76	<.001 *	.46	11.73	<.001 *	.32

*
-Value retains significance following FDR adjustment.

Table 3

Descriptive information for older and young adults on all six recognition discriminability (RD) indices.

RD index	Older				Young			
	Mean (SD)	25th %ile	50th %ile (median)	75th %ile	Mean (SD)	25th %ile	50th %ile (median)	75th %ile
SoRD	2.96 (0.71)	2.40	3.10	3.70	3.40 (0.45)	3.40	3.70	3.70
SoRD-prototypical	3.00 (0.94)	2.04	3.05	4.00	3.61 (0.55)	3.25	4.00	4.00
SoRD-unrelated	3.34 (0.73)	2.99	3.63	4.00	3.67 (0.53)	3.63	4.00	4.00
NRD	2.94 (0.67)	2.40	3.10	3.70	3.40 (0.40)	3.25	3.40	3.70
NRD-prototypical	2.85 (0.98)	2.21	2.77	4.00	3.50 (0.65)	3.05	3.63	4.00
NRD-unrelated	3.52 (0.56)	2.99	3.63	4.00	3.76 (0.34)	3.63	4.00	4.00

RD = recognition discriminability; SoRD (source recognition discriminability) = List A targets vs. all List B distractors; SoRD-prototypical = List A targets vs. prototypical List B distractors only; SoRD-unrelated = List A targets vs. unrelated List B distractors only; NRD (novel recognition discriminability) = List A targets vs. all novel distractors; NRD-prototypical = List A targets vs. prototypical novel distractors only; NRD-unrelated = List A targets vs. unrelated novel distractors only.

Table 4

Age group differences on SoRD, NRD, SoRD-unrelated, and NRD-unrelated indices, with associated statistics based on Mann–Whitney U tests.

RD index	Older	Young			
	Mean rank (sum of ranks)	Mean rank (sum of ranks)	U	p	r
SoRD	45.90 (2524.50)	66.73 (3803.50)	984.50	<.001 *	.33
SoRD-prototypical	46.12 (2536.50)	66.52 (3791.50)	996.50	<.01 *	.33
SoRD-unrelated	48.81 (2684.50)	63.92 (3643.50)	1144.50	<.01 *	.25
NRD	45.06 (2478.50)	67.54 (3849.50)	938.50	<.001 *	.36
NRD-prototypical	45.35 (2494.50)	67.25 (3833.50)	954.50	<.001 *	.22
NRD-unrelated	49.70 (2733.50)	63.06 (3594.50)	1193.50	<.05 *	.22

*
p-Value retains significance following FDR adjustment.

RD = recognition discriminability; SoRD (source recognition discriminability) = List A targets vs. all List B distractors; SoRD-prototypical = List A targets vs. prototypical List B distractors only; SoRD-unrelated = List A targets vs. unrelated List B distractors only; NRD (novel recognition discriminability) = List A targets vs. all novel distractors; NRD-prototypical = List A targets vs. prototypical novel distractors only; NRD-unrelated = List A targets vs. unrelated novel distractors only.

Table 5

Within-group differences on SoRD, NRD, SoRD-unrelated, and NRD-unrelated indices, with associated statistics based on Wilcoxon signed-rank tests.

Comparison	Older			Young		
	Z	p	r	Z	p	r
SoRD vs. NRD	0.33	.744	.04	0.43	.670	.06
SoRD-prototypical vs. NRD-prototypical	1.36	.173	.18	1.24	.214	.16
SoRD-unrelated vs. NRD-unrelated	2.16	<.05	.29	1.63	.102	.22
SoRD vs. SoRD-prototypical	0.18	.860	.02	4.16	<.001*	.55
SoRD vs. SoRD-unrelated	5.83	<.001*	.79	5.03	<.001*	.67
SoRD-prototypical vs. SoRD-unrelated	3.65	<.001*	.49	1.19	.234	.16
NRD vs. NRD-prototypical	2.05	<.05	.28	1.06	.291	.14
NRD vs. NRD-unrelated	6.21	<.001*	.84	5.88	<.001*	.78
NRD-prototypical vs. NRD-unrelated	4.40	<.001*	.59	2.89	<.01*	.38

*
p-Value retains significance following FDR adjustment.

RD = recognition discriminability; SoRD (source recognition discriminability) = List A targets vs. all List B distractors; SoRD-prototypical = List A targets vs. prototypical List B distractors only; SoRD-unrelated = List A targets vs. unrelated List B distractors only; NRD (novel recognition discriminability) = List A targets vs. all novel distractors; NRD-prototypical = List A targets vs. prototypical novel distractors only; NRD-unrelated = List A targets vs. unrelated novel distractors only.