Research

The Retrograde Memory for News Events Test (RM-NET) and the relationship between news event memory and performance on standard neuropsychological tests

Andrew T.J. Cawley-Bennett,[1] Jennifer C. Frascino,[2,3] Isabel E. Asp,[3] Shahrokh Golshan,[2,3] Mark W. Bondi,[2,3] Zhishang Luo,[4] and Christine N. Smith[2,3,5]

[1]*Department of Psychology, Emory University, Atlanta, Georgia 30322, USA;* [2]*Department of Psychiatry, University of California at San Diego, San Diego, California 92093, USA;* [3]*Veterans Affairs San Diego Healthcare System, San Diego, California 92161, USA;* [4]*Halıcıoğlu Data Science Institute, University of California at San Diego, San Diego, California 92093, USA;* [5]*Center for the Neurobiology of Learning and Memory, University of California at Irvine, Irvine, California 92697, USA*

Novel tests of semantic memory (SM)—for example, memory for news events (NE; news facts) or famous personalities—are useful for estimating the severity of retrograde amnesia. Individuals with mild cognitive impairment exhibit relatively intact SM/language on traditional neuropsychological tests but exhibit consistent impairment on novel tests of SM, suggesting novel SM tests are dissimilar from traditional SM tests. To identify the relationship between NE memory and traditional cognitive measures, older adults ($N = 51$) completed a traditional neuropsychological battery and the Retrograde Memory News Events Test (RM-NET; a new test that robustly measures NE memory across the adult life span with high temporal resolution), and the relationship between performance on these tests was examined. Total RM-NET scores were more closely aligned with episodic memory scores than SM scores. The strength of the association between NE scores and episodic memory scores decreased as the age of NE memory increased. Tests of news events appear to reflect performance on traditional tests of episodic memory rather than SM, especially when recent news events are tested.

[Supplemental material is available for this article.]

Traditional neuropsychological tests are used to diagnose or characterize the nature of cognitive impairments in individuals, such as those with Alzheimer's disease, frontotemporal dementia, mild cognitive impairment (MCI), Huntington's disease, Korsakoff syndrome, or epilepsy. These tests tap into abilities that reflect different domains of cognition, such as episodic memory (EM), semantic memory (SM)/language, attention/processing speed, and executive functions. Many of these patient groups also exhibit impairment on novel tests of SM, such as news event memory or memory for famous personalities (Sanders and Warrington 1971; Kopelman 1989; Sadek et al. 2004; Milton et al. 2010; Irish et al. 2012; Smith 2014; Langlois et al. 2016).

Novel tests of SM, such as tests of notable news events or famous personalities, may detect impairment better than traditional neuropsychological tests of SM (Venneri et al. 2016). For example, individuals with MCI who do not typically exhibit impairment on traditional semantic tests (e.g., tests of language, category and letter fluency, and object naming) are consistently impaired on tests of news event memory or memory for famous personalities (Flicker et al. 1987; Murphy et al. 2008; Leyhe et al. 2009a, 2010; Seidenberg et al. 2009; Irish et al. 2010; Barbeau et al. 2012; Thomann et al. 2012; Smith 2014). Although these novel semantic tests reveal consistent impairment, it is unclear how performance on these tests relates to performance on the wider array of neuropsychological tests used to identify cognitively impaired individuals in the clinic. Therefore, these novel measures merit broader examination to identify which cognitive domains they reflect and whether they reflect new information about SM not measured by the traditional tests used clinically to identify cognitive impairment.

One way to understand the relationship between novel SM tests and traditional EM tests is to examine the brain structures that support performance on the two types of tests. By definition, news event memory is SM because it reflects memory for facts and knowledge about the world. Accordingly, like other fact memory, memory for news facts is impaired by damage to lateral temporal cortices (Bayley et al. 2005; Gilboa et al. 2005; Bright et al. 2006). Like other facts, news facts also typically lose information about the context in which they were learned (i.e., the time and place of the learning event). Therefore, according to the standard concepts described by Tulving (1983), news facts do not represent medial temporal lobe (MTL)-dependent EM because they do not retain information about the context of the learning event. Indeed, theories of memory consolidation agree that retrieval of recent, but not remote, SM depends on structures in the MTL (Marr 1971; Squire and Alvarez 1995; Nadel et al. 2000; Moscovitch et al.

Corresponding author: cnsmith@ucsd.edu
Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.053571.122.

2005). Thus, recent news event memory scores are additionally dependent on the MTL, which could drive an association between performance on news event tests and performance on other MTL-dependent tests (e.g., traditional EM tests of new learning).

Although SM and EM have distinct definitions and phenomenology (Tulving 1983), more contemporary theories suggest that EM and SM are more likely part of a continuum rather than distinct entities (Renoult et al. 2012, 2019; Irish and Vatansever 2020). Indeed, the neural substrates of EM and SM retrieval are highly overlapping (Dede and Smith 2016; Renoult et al. 2019; Irish and Vatansever 2020). Specifically, a common network is thought to support memory retrieval (parahippocampal cortex, middle temporal gyrus, ventral–lateral parietal cortex, and midline prefrontal and parietal regions), but depending on the retrieval requirements/memory content, the network differs in the degree that it involves regions more specific to EM (hippocampus) or more specific to SM (anterior temporal lobe and inferior frontal gyrus) (Renoult et al. 2019). This work is in line with the earlier idea that memory retrieval in everyday life involves dynamic interplay between these two types of memory (Tulving 1983). Consistent with these ideas, there is a continuum of SM that depends on the characteristics of the memory itself. Traditional SM tests assess memory for knowledge learned long ago (e.g., object knowledge and language), and this type of information is re-encountered frequently across the life span (common words, objects, and language use). In contrast, the novel SM tests most sensitive to mild brain injury reflect knowledge learned more recently and not re-encountered frequently (e.g., news events from the recent past that received limited news coverage). In this framework, novel SM tests likely tap into a mixture of EM and SM that better captures the memory retrieval network affected in mild cognitive impairment.

The relationship between novel measures of SM and performance on traditional neuropsychological tests has not been comprehensively examined. Across the five studies that have examined these relationships, there was variability in which domains were examined and the types of novel semantic tests used. As a whole, there were few consistent findings. Leyhe et al. (2010) used a 20-item test for news events from the previous 60 yr and found that mean performance across MCI and AD groups was significantly correlated with performance on a global measure of cognition (Mini-Mental State Examination [MMSE]) but was not correlated with individual tests of attention/processing speed or executive function (other domains were not examined). Johnson and Klingler (1976) used a 70-item test for news events from the previous 30 yr and found that performance for individuals with normal cognition (ranging from young to older adults) was significantly correlated with performance averaged across several standard tests of verbal and nonverbal EM (other domains were not examined). Seidenberg et al. (2009) used a test of famous personalities who came to prominence in the recent past (10 items) or remote past (10 items) and found that the number of semantic details reported by individuals with normal cognition or MCI was also significantly correlated with traditional measures of verbal EM (other domains were not tested). De Simone et al. (2020) also found significant correlations with measures of EM and executive function for their six-item recall and recognition test of news events spanning 45 yr. However, when Langlois et al. (2016) examined individuals with MCI or AD using a 40-item test for news events from the previous 45 yr, they did not detect a significant relationship with standard measures of verbal EM. Instead, they found significant correlations between news event memory and individual measures of SM (pyramids and palm trees test and the information subscale from the WAIS; other domains were not examined).

Only two of these studies examined these relationships as a function of the age of memory (Seidenberg et al. 2009; Leyhe et al. 2010). Measures of global cognition (Leyhe et al. 2010) and verbal EM (Seidenberg et al. 2009) were significantly correlated with novel measures of SM from both recent (1- to 10-yr-old) and remote (40- to 55-yr-old) time periods.

Thus, when identifying relationships with neuropsychological tests, there is variability in the types of novel SM tests used, in the number of time periods examined, and in the number of test items used to estimate retrograde memory. These relationships have never been examined systematically with a novel SM test that has high temporal resolution across the entire adult life span and where performance for each time period is estimated by more than a handful of questions. In addition, there has been no comprehensive analysis of these associations using robust estimates of five standard cognitive domains. Typically, a single neuropsychological test was used to estimate ability for a cognitive domain, even though using more than one test per domain increases the reliability of assessing a cognitive domain (Anastasi and Urbina 1997; Palmer et al. 1998; Loewenstein et al. 2007).

We describe the creation of a novel SM test of notable news facts (Retrograde Memory News Events Test [RM-NET]) that spans the entire adult life span (70 yr, separated into 3- to 5-yr time periods) (see Table 1). We examined behavioral findings from the RM-NET that reflect different elements of news event memory (e.g., accuracy, confidence, and response times). We also examined measures obtained from the RM-NET posttest that reflect incidental encoding during the RM-NET for events from the last 30 yr (see Table 2), as well as subjective reports of the amount of semantic knowledge available and the presence of concomitant autobiographical memories associated with the news event. Because of the relatively separate traditions of examining clinical populations using news events tests or standard neuropsychological tests, we asked how performance on the RM-NET relates to performance on traditional neuropsychological tests. Based on prior findings, we hypothesized that RM-NET memory accuracy would significantly predict performance on measures of EM. Furthermore, because news event memory is fact memory (i.e., a type of SM), we also predicted that RM-NET memory accuracy would significantly predict performance on measures of SM. Therefore, we tested how performance on the RM-NET was related to performance on standard neuropsychological tests in five cognitive domains: EM, SM/language, executive function, attention/processing speed,

Table 1. Time periods of the Retrograde Memory News Events Test (RM-NET)

Years before testing	From–to	Duration (years)	Number of items
1–3	2017–2015	3	20
4–6	2014–2012	3	20
7–9	2011–2009	3	20
10–12	2008–2006	3	20
13–15	2005–2003	3	20
16–20	2002–1998	5	20
21–25	1997–1993	5	20
26–30	1992–1988	5	20
31–35	1987–1983	5	10
36–40	1982–1978	5	8
41–45	1977–1973	5	8
46–50	1972–1968	5	10
51–55	1967–1963	5	8
56–60	1962–1958	5	9
61–65	1957–1953	5	10
66–70	1952–1948	5	8

Years before testing represents the approximate age of the memories queried. Testing occurred between March 2018 and April 2020. News events from the eight time periods spanning 1–30 yr before testing were queried in the RM-NET posttest.

Table 2. Example questions from the Retrograde Memory News Events Test (RM-NET)

Year	News event question	Recognition memory multiple-choice options	Surprise recognition memory test	Surprise recognition memory test multiple-choice options
2016	Which Asian country impeached its first female president?	(A) Indonesia **(B) South Korea** (C) India (D) Nepal	Which topic were you asked a question about?	**(1) The Asian country that impeached its first female president** (2) The company tied to the impeachment of the first female president of an Asian country (3) The reason an Asian country's first female president was impeached
2012	Why did skydiver Felix Baumgartner make news?	(A) First to jump off Freedom Tower (B) Used flying suit in Alps **(C) Skydived from space** (D) Parachuted from Mt. Everest	Which topic were you asked a question about?	(1) The age of skydiver Felix Baumgartner when he made worldwide news (2) The country that skydiver Felix Baumgartner is from **(3) The reason that skydiver Felix Baumgartner is known**
1997	What is Heaven's Gate?	**(A) Cult** (B) Computer game (C) Restaurant (D) Magazine	Which topic were you asked a question about?	**(1) What Heaven's Gate is** (2) The city that is associated with Heaven's Gate (3) Who is associated with Heaven's Gate
1988	Who was Ivan Boesky?	(A) Nazi leader convicted of war crimes (B) Ballet dancer who defected to the West **(C) Stock trader, convicted of insider trading** (D) Russian ambassador	Which topic were you asked a question about?	(1) What was Ivan Boesky's job before he became well known **(2) Why was Ivan Boesky famous** (3) The name of the U.S. commission Ivan Boesky was associated with

Correct answers appear in bold.

and visuospatial function. Because there appears to be a continuum of memory retrieval spanning EM to SM, another goal was to evaluate the relative effect sizes between RM-NET memory accuracy and these two domains. To obtain reliable estimates of ability for each domain of cognition, four to seven measures were used to estimate each domain. Finally, due to the dependence of recent news event memory and EM on the MTL, we hypothesized that the strength of the relationship between RM-NET memory accuracy and EM would decrease as memory age increased.

Results

Effects of covariates on variables of interest

The covariates are described in the Materials and Methods in "Identifying Relevant Covariates for Primary Analyses." Table 3 shows descriptive information for covariates and variables of interest, and Supplemental Table S1 shows significance values when using covariates to predict variables of interest. The RM-NET interval (the relative number of days that elapsed between when the first participant was tested and when each subsequent participant was tested) (see the Materials and Methods) significantly predicted RM-NET total accuracy, total confidence, and total response times. In addition, sex significantly predicted RM-NET total response times. These covariates were included in analyses of RM-NET components. Sex, education, total medical health burden, and total mental health burden significantly predicted one or more of the cognitive domain composite scores. None of the other covariates significantly predicted any of the variables of interest. In order to make the regression models comparable, all these significant covariates were included in the primary analyses.

Descriptive information about the RM-NET components and the RM-NET posttest components

Descriptive information about the sample, cognitive domain composite scores, and components of the RM-NET appear in Table 3. Estimated marginal means and SEM (corrected for covariates) for RM-NET accuracy, confidence, and response times appear in Figure 1 for each time period. There were no significant changes in these components across the 13 time periods (P-values > 0.13). Analysis of RM-NET posttest components revealed no significant changes in knowledge judgments ($P = 0.406$) or the number of autobiographical memories ($P = 0.072$) across the eight time periods. For autobiographical memories, this trend indicated an increasing number of autobiographical memories for older news events. Note that because reports of autobiographical memories were so rare (mean = 0.28 trials/time period) and they were not obtained for all time periods from the RM-NET, we did not investigate autobiographical memories further. There was a significant linear decrease in subsequent memory accuracy across the time periods ($F_{(1,50)} = 5.26$, $\eta_p^2 = 0.097$, $P = 0.026$).

Identifying relationships between total RM-NET memory accuracy and performance on traditional neuropsychological tests

We asked how performance on news events test relates to performance on traditional neuropsychological tests. We answered this using two approaches: (1) a theoretical approach and (2) a data-driven approach. For the theoretical approach, individual neuropsychological tests were grouped according to the primary domain of cognition they were thought to reflect (see the Materials and Methods), creating five cognitive domain composite scores. We carried out bivariate correlation analysis between RM-NET total accuracy scores and the cognitive composite scores to obtain estimates of effect size. The distributions of the RM-NET scores and domain composite scores appear in Figure 2, A and B, respectively. The largest effect sizes were observed for EM composite scores ($r = 0.425$, $P = 0.002$) and attention/processing speed composite scores ($r = 0.453$, $P = 0.001$). Smaller (and nonsignificant, P-values > 0.01) effect sizes were observed for SM/language ($r = 0.311$, $P = 0.026$), executive function ($r = 0.309$, $P = 0.027$), and visuospatial function ($r = 0.336$, $P = 0.016$).

Table 3. Descriptive statistics for participant characteristics and variables of interest

	Number or percentage	Mean	SD	Minimum–maximum
Participant characteristics				
Number of participants	51			
Age (years)		72.8	6.1	65–91
Education (years)		15.6	2.3	12–20
Sex (% male)	62.7%			
Race/ethnicity (% white and non-Hispanic)	78.4%			
Veteran (% veteran)	74.5%			
Medical health burden		3.3	1.8	0–7
Mental health burden		1.5	1.3	0–4
Frequency following news		3.6	0.7	1–4
Number of news sources		4.6	1.7	1–7
Cognitive domain composite scores				
Episodic memory (*Z*-score)		0.17	0.92	−1.51 to 1.99
Semantic memory/language (*Z*-score)		0.33	0.72	−1.33 to 1.84
Executive functions (*Z*-score)		0.08	0.64	−1.53 to 1.50
Attention/processing speed (*Z*-score)		0.10	0.53	−1.43 to 1.27
Visuospatial functions (*Z*-score)		0.19	0.36	−0.69 to 0.91
Retrograde Memory News Events Test (RM-NET)				
Accuracy (% correct)		65.3	12.0	40.8–89.1
Confidence rating		2.8	0.5	1.1–3.6
Response time (seconds)		8.3	1.1	6.2–11.0
Subsequent memory accuracy (% correct)		80.5	13.9	44.8–96.3
Knowledge rating		2.8	1.4	1.1–7.0
Autobiographical memory (number)		2.2	2.7	0–12
RM-NET interval (days)		324	254	0–727

Medical and mental health burden reflects the total number of comorbidities. Frequency following news ranges from 1 (never) to 4 (frequently). Number of news sources ranges from 0 to 7. Confidence rating scale was from 1 (pure guess) to 4 (definitely sure). The knowledge rating scale was from 1 (no information) to 10 (a lot of information about the news event topic). Autobiographical memory reflects the number of news events (out of 160) that were accompanied by autobiographical memories. The RM-NET interval reflects the relative number of days that elapsed between when the first participant was tested and when each subsequent participant was tested.

To account for the possible influence of covariates, multiple regression was also used to obtain measures of effect size (B [unstandardized beta coefficients ± SEM] and β [standardized beta coefficients]). RM-NET accuracy scores were used to predict each of the five cognitive composite scores (including significant covariates from Supplemental Table S1). Results from these analyses are concordant with the results from the correlational analysis (see Table 4; Fig. 3). These findings suggest that the RM-NET primarily reflects the EM and attention/processing speed domains of cognition measured by traditional neuropsychological tests, and these findings were robust to the inclusion of covariates.

Next, we used a data-driven approach to determine which individual neuropsychological tests were associated with RM-NET accuracy scores. Because there were so many neuropsychological tests, we opted for an exploratory factor analysis in lieu of bivariate correlations. We carried out the factor analysis on all the traditional neuropsychological test *Z*-scores and the RM-NET total accuracy scores. This method empirically identified clusters of these tests where performance varied together. The important finding was which factors the RM-NET loaded on mostly strongly and which individual tests were contained in these factors. We found that eight factors accounted for a significant amount of variance among the tests for a total of 73% cumulative variance explained. The RM-NET loaded highest on the first (0.355), fourth (0.432), and fifth (0.491) factors. The first factor explained 29.8% of variance and was composed of the EM tests (except for visual reproduction immediate recall) and the digit span sequencing test. This factor primarily reflected verbal and nonverbal EM, as well as verbal short-term memory and attention. The fourth factor explained 6.6% of variance and was composed of digit span forward and digit span backward tests. This factor reflected verbal short-term and working memory. The fifth factor explained 5.7% of variance and was composed of the RM-NET, the Multilingual Aphasia Exam token test, and the Clock Drawing command test. This factor appeared to reflect long-established SM. Thus, the RM-NET was most strongly related to individual neuropsychological tests that reflected verbal and nonverbal EM, verbal short-term and working memory, and long-established SM.

Identifying whether the relationship between RM-NET memory accuracy and episodic memory changes with memory age

The hypothesis that the relationship between RM-NET memory accuracy and EM changes as a function of memory age was tested by carrying out regression using memory age to predict the effect sizes between RM-NET memory accuracy for each time period and the single EM composite score (shown in Fig. 4A). We found that memory age significantly predicted effect sizes such that higher memory age was associated with lower effect size (Table 5; Fig. 4A). For the other domains of cognition, there were no significant effects of memory age on effect sizes (Table 5; Fig. 4B–E).

Discussion

The RM-NET was developed to robustly measure news event memory across the entire adult life span with high temporal resolution (Fig. 1). We examined relationships between RM-NET memory accuracy scores and robust measures of performance in five cognitive domains as estimated by traditional neuropsychological tests. Testing was carried out in older adults without dementia to provide ample variability in these measures (Fig. 2) and to obtain measures of the strength of the associations (effect sizes). RM-NET accuracy was most highly associated with composite scores derived from traditional measures of EM and attention/processing speed (Fig. 3).

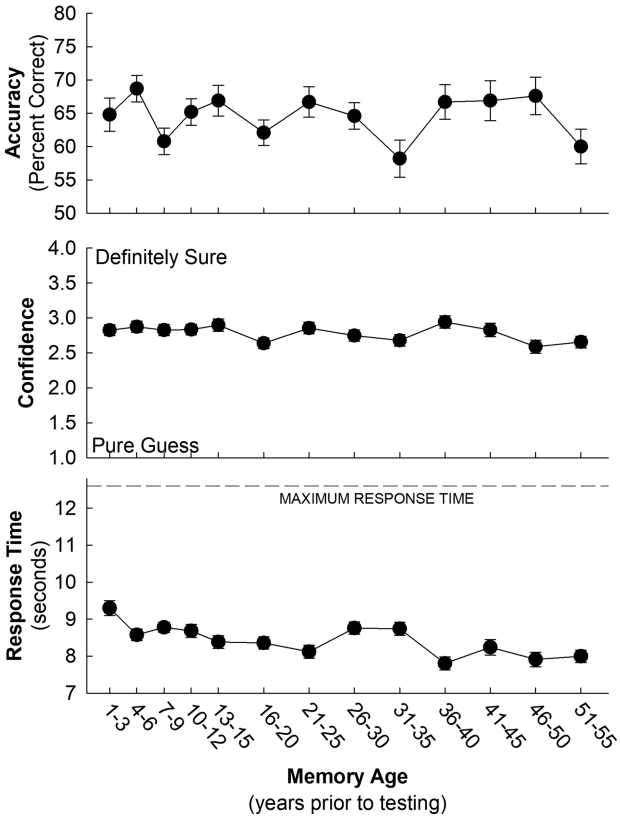

Figure 1. Mean accuracy scores (*top*), confidence judgments (*middle*), and response times (*bottom*) for the RM-NET for news events that occurred between 1 and 55 yr prior to testing (memory age) for 51 older adults. Performance did not change across time periods according to linear or power functions (P-values > 0.17). Means and SEMs were adjusted for the effects of covariates.

RM-NET accuracy was most highly associated with individual neuropsychological tests that reflected verbal and nonverbal EM, verbal short-term and working memory, attention, and long-established SM/language. The strength of the relationship between RM-NET accuracy and EM decreased as a function of the age of the news event memory (Fig. 4A). This pattern was not observed for other domains of cognition (Fig. 4B–E).

The RM-NET and episodic memory

RM-NET memory accuracy was strongly associated with EM scores on the traditional neuropsychological tests. The finding that EM scores were related to measures from novel SM tests (e.g., news events or famous names) is consistent with three earlier studies. These studies examined news event memory in young to older adults (Johnson and Klingler 1976) or older individuals with normal cognition, MCI, or AD (De Simone et al. 2020), or they examined memory for famous names in older adults with normal cognition or MCI (Seidenberg et al. 2009). These studies did not include covariates for demographic characteristics when examining these relationships; therefore, the associations could have been influenced by such factors (e.g., sex). Our findings demonstrate that significant relationships with EM can be detected even when the influence of demographic characteristics are taken into account.

We also observed that the effect size of the association between EM and RM-NET memory accuracy decreased as memory age increased (Fig. 4A), a pattern that was not observed for other domains of cognition (Fig. 4B–E). Because EM and recent news event memories depend on the integrity of the MTL (Kapur and Brooks 1999; Manns et al. 2003; Bayley et al. 2006), our findings likely reflect shared dependence on these brain structures, as well as more overlap between the more extended neural systems that support both EM and SM retrieval (Dede and Smith 2016; Renoult et al. 2019; Irish and Vatansever 2020). With regard to the idea that memory retrieval represents a continuum between EM (which contains highly detailed information related to the context of the learning event) and SM (which has lost this information), our findings support the idea that recent news event memories and remote news event memories lie on different parts of this continuum.

These ideas may also explain why Seidenberg et al. (2009) also found significant correlations between a measure of EM (delayed free recall of a word list) and a novel measure of people who became famous in the recent past (1–10 yr ago). Unlike the current findings, they also found a significant correlation between EM and famous personalities from the remote past (40–55 yr ago), and their effect sizes were similar for recent and remote memory (Pearson correlation coefficient values = 0.40). This difference between the findings of these studies may pertain to the different response requirements associated with the novel SM tests. The RM-NET data reported here reflect forced-choice recognition memory for news events where each test item was scored as either correct or incorrect. In contrast, Seidenberg et al.'s (2009) test allowed four opportunities to report semantic information about the famous personalities (i.e., reason they are well known, known works/accomplishments, names of other people associated with the individual, and history and background of the individual), resulting in a score from 0 to 12 for each name the participant previously identified as being famous. Recollection of this additional information for Seidenberg et al.'s (2009) test may be more reliant on EM, regardless of the age of the memory.

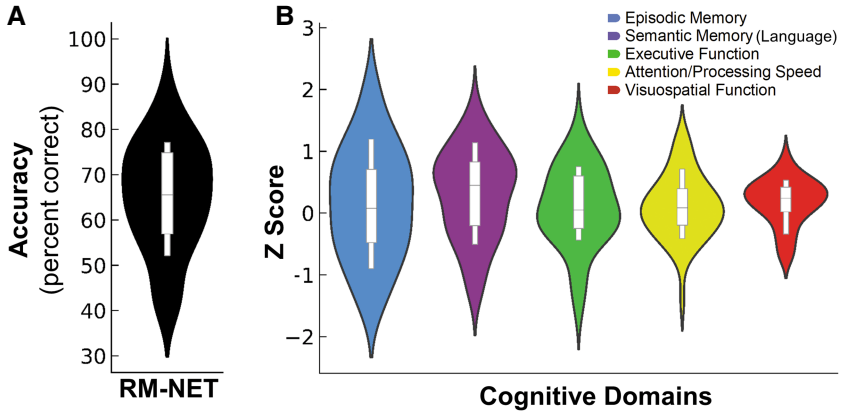

Figure 2. Descriptive measures (N = 51) for RM-NET accuracy scores (*A*) and cognitive domain composite scores (*B*). Violin plots illustrate medians (horizontal lines) and distributions (outer curved lines). *Below* the median is the first quartile (25th percentile, *bottom* white rectangle) and *above* the median is the third quartile (75th percentile, *top* white rectangle). Cognitive domain composite scores reflect mean Z-scores for four to seven tests relative to published norms.

Table 4. Multiple regression results using cognitive domain composite scores to predict mean RM-NET accuracy scores

Cognitive domain	B	SE B	β	*T*-value	*P*-value
Episodic memory	3.04	0.91	0.40	3.33	0.002
Semantic memory/language	1.55	0.80	0.26	1.94	0.058
Executive functions	1.41	0.71	0.27	2.00	0.052
Attention/processing speed	2.19	0.59	0.50	3.74	0.001
Visuospatial functions	0.98	0.44	0.33	2.23	0.031

P-values of <0.01 were considered statistically significant after correcting for multiple comparisons. Unstandardized beta coefficients (B) and SE B are shown in Figure 3.

The RM-NET and semantic memory/language

Although RM-NET accuracy scores were most highly related to EM and attention/processing speed, there was a trend for them to be related to SM ($P = 0.058$) (Table 4). In addition, factor analysis revealed that the RM-NET was grouped with a test of oral comprehension/language test (token test). Although the other test it was grouped with, the Clock Drawing command test, is considered to primarily reflect visuospatial functions, performance on this test also appears to be influenced by long-established semantic knowledge regarding the appearance of a clock (Rouleau et al. 1996; Leyhe et al. 2009b). Thus, these two neuropsychological tests likely reflect a correspondence between the RM-NET and tests sensitive to impairments in basic language comprehension and recall of long-established semantic knowledge.

Given that novel SM tests are more sensitive to mild impairment than traditional semantic tests (Seidenberg et al. 2013; Orlovsky et al. 2018), it is important to consider why this is the case and why it may not be surprising that RM-NET accuracy was only weakly (and nonsignificantly) associated with the SM composite scores. As described earlier, traditional SM tests typically assess memories that are very long established and re-experienced (relearned) frequently, whereas novel SM tests do not. These qualities of novel SM tests are the same ones that are associated with making information particularly vulnerable to disruption. Because novel SM tests differ from traditional SM tests in these two characteristics, it may explain why novel tests are sensitive to mild brain injury/disease (e.g., MCI) and why performance on the RM-NET is not highly associated with performance on these traditional tests.

It is also relevant that traditional and novel SM tests also lie on different parts of the SM continuum, with traditional SM tests solidly at the SM pole and reflecting highly conceptual entities, word knowledge, and language use, while novel semantic tests do not. In this way, the RM-NET shares characteristics with other types of SM that depart from the traditional concept of SM. For example, Petrican et al. (2010) found empirical evidence for the idea that as news event memories age, the memory traces lose the additional contextual information about how they were learned, which is thought to reflect EM. Moreover, personal SMs (e.g., knowledge about your primary school) that are accompanied by spatiotemporal and perceptual details (Renoult et al. 2012; Grilli and Verfaellie 2014; Grilli and Verfaellie 2016), as well as estimates of recollection-like knowledge about famous people (Waidergoren et al. 2012), can make these types of SM more similar to EM and make them MTL-dependent.

The RM-NET and attention/processing speed

RM-NET memory accuracy was also strongly associated with attention/processing speed scores (Fig. 3), and the strength of this association did not change with memory age (Fig. 4). This finding suggests that there were similar demands on attention/processing abilities for recent versus remote memories. Factor analysis also revealed that the RM-NET was highly associated with individual attention/processing speed tests that reflected verbal short-term and working memory and attention. Therefore, even though the RM-NET was timed (individuals had only ~13 sec to read and answer the question), which could have inflated relationships with processing speed measures, the strong relationship with attention/processing speed appears to be driven by measures of attention, not processing speed. This idea is in line with findings from another study that found no significant relationship between performance on an untimed news event memory test and a traditional measure of processing speed (Leyhe et al. 2010).

Limitations

This study had several limitations. First, because most of the data were acquired in the context of a functional neuroimaging study (which will be reported separately), the RM-NET was administered in a timed format and participants could not see their fingers when making responses. This procedure may have inflated the association between RM-NET accuracy and performance on cognitive domains (e.g., attention/processing speed, executive functions, or visuospatial function). Alternatively, these components could be retained if an assessment of these additional cognitive domains was desired. Second, it took ~2 yr to acquire these data and forgetting occurred across this interval (i.e., a significant effect of the date of testing on total RM-NET accuracy). To adapt to this challenge, we included the date of testing as a covariate in the analyses. Finally, the RM-NET takes 1 h to administer, precluding its widespread clinical utility. It would be fruitful to develop a shortened version of the test, fine-tuned to detect clinically useful outcomes (e.g., distinguishing between normal cognition and MCI).

Summary

We describe the RM-NET, which is a comprehensive measure of memory for news facts acquired across the entire adult life span. The test can be given in two formats: recognition memory and free recall. Recognition memory scores from the RM-NET were

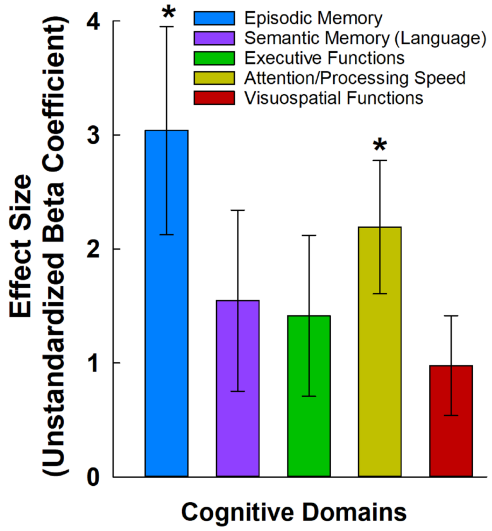

Figure 3. Relative effect sizes of the association between RM-NET accuracy scores and cognitive domain composite scores. RM-NET accuracy scores significantly predicted episodic memory and attention/processing speed composite scores. The *Y*-axis reflects the unstandardized beta coefficients (±SEM) from the regression equations (including covariates). (*) $P < 0.05$, corrected for multiple comparisons.

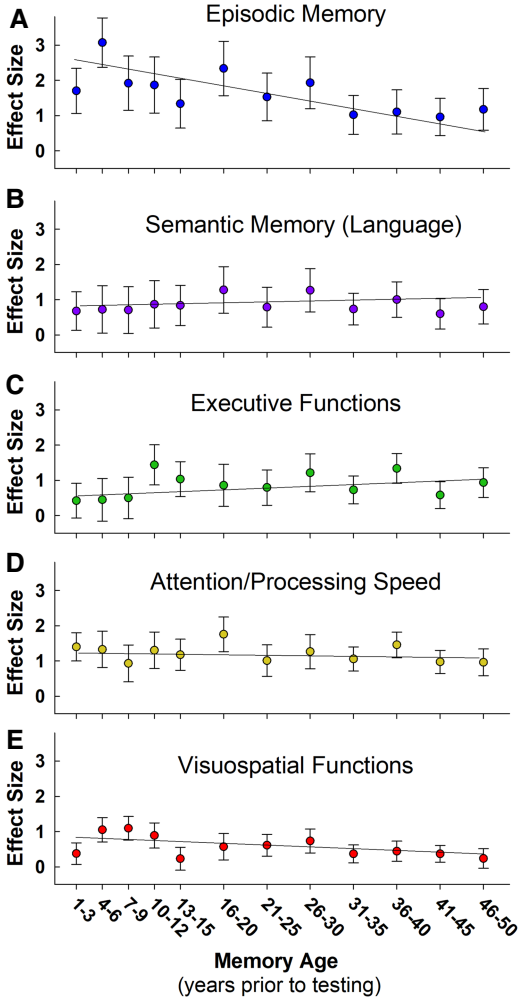

Figure 4. Effect sizes for the association between RM-NET accuracy scores for each time period and a single episodic memory domain composite score. *Y*-axis as in Figure 3. Memory age indicates the age of the news event memory relative to the testing date. Memory age significantly predicted episodic memory effect sizes (*A*; higher memory age was associated with lower effect size) but did not predict effect sizes for the other cognitive domains (*B–E*).

measured and primarily reflected traditional neuropsychological tests of EM, with news facts from the recent past more strongly associated with EM than facts from the remote past. A measure of incidental encoding of the RM-NET items also provides a measure of anterograde EM. In addition to EM, verbal working memory and short-term memory, as well as attention and long-established SM, also likely contribute to success on news events tests. Although the RM-NET was only weakly related to SM composite scores, and scores shared variance with one measure of oral comprehension/language, news event memory appears to reflect a type of SM not effectively measured by the traditional SM tests commonly used to detect cognitive impairment in the clinic (fluency and confrontational naming). This is likely due to the test's reliance on memory of news facts that have had less opportunity for relearning and that vary in the age of the memories queried. Novel SM tests, such as the RM-NET, continue to provide promise and opportunity for detecting mild impairments in cognition when traditional SM tests cannot. In addition, computerized administration and scoring of the RM-NET can occur without a trained administrator, increasing the ease and efficiency of identifying individuals with mild cognitive impairment.

Materials and Methods

Participants

Participants were recruited using flyers at the Veterans Affairs San Diego Healthcare System's (VASDHS) La Jolla and Mission Valley locations or other VA clinics in San Diego County, during a visit at the Neuropsychology Unit at VASDHS, and from the Alzheimer's Disease Research Center (ADRC) at University of California at San Diego. All participants provided written informed consent and were recruited and enrolled without regard to ethnicity or race.

Individuals were excluded if they met any of the following exclusion factors: diagnosis of dementia; <65 yr old; impaired activities of daily living; impaired score on MMSE; impaired reading ability; uncontrolled high blood pressure; not fluent in English; had not lived in the United States for most of adulthood; ineligible for MRI; left-handed (because functional neuroimaging was obtained for these participants); traumatic brain injury or head injury with loss of consciousness >30 min; strokes or transient ischemic attacks; chronic disorders of the lung or heart; seizures or other neurological conditions such as Parkinson's disease or dementia; type I diabetes or uncontrolled type II diabetes; general anesthesia in the previous 4 mo; chemotherapy or full-body radiation for cancer treatment; diagnosed with schizophrenia, bipolar disorder, or psychotic disorders; untreated major depression or exhibiting moderate to severe depression symptoms; or currently enrolled in an alcohol/drug treatment program. Exclusion criteria were identified via telephone interview prior to enrollment, except for depression symptoms, activities of daily living, MMSE, reading ability, and blood pressure, which were measured at the first visit. Participants were included in the study regardless of whether they exhibited impairment on the standard neuropsychological tests used to estimate ability in the five cognitive domains that were examined (see "Traditional Neuropsychological Assessment," below). Therefore, these participants likely included individuals with normal cognition or mild cognitive impairment.

Criteria used for exclusion

Participants were considered to have impaired activities of daily living if they obtained a scaled score of ≤40 (low functioning) on the "health and safety" or "managing money" subscales of the Independent Living Scales Test (Loeb 1996) or if they obtained a

Table 5. Multiple regression results where the age of the news event memory predicted the effect size between RM-NET accuracy scores and cognitive domain composite scores

Cognitive domain	B	SE B	*T*-value	*P*-value	Permutation *P*-value
Episodic memory	−0.0057	0.001	−4.23	0.002	0.003
Semantic memory/language	−0.0008	0.001	−0.70	0.502	0.494
Executive functions	0.0017	0.002	0.89	0.392	0.393
Attention/processing speed	−0.0002	0.002	−0.14	0.895	0.886
Visuospatial functions	−0.0034	0.002	−1.63	0.135	0.143

P-values of <0.01 were considered statistically significant after correcting for multiple comparisons.

score of ≤8 on the 15-item Functional Assessment Questionnaire (Pfeffer et al. 1982). They were considered impaired on the MMSE if they received a score <25 (Folstein et al. 1975). Reading ability was considered to be impaired if their sum of correct responses was less than two standard deviations below norms on the Wide Range Achievement Test (WRAT4) Blue Word Reading Test. They were considered to have moderate depressive symptoms if they obtained a score of ≥7 on the 15-item Geriatric Depression Scale (Sheikh and Yesavage 1986). Normal blood pressure was determined by using the age- and sex-adjusted values from the American Heart Association (Whelton et al. 2018). Blood pressure readings were obtained in visit 1 or visit 2 or reported by the participant (i.e., from a measurement obtained from another source, such as a recent doctor's visit). Finally, participants were excluded after the first visit if they could not successfully complete a news events practice test that simulated how the test was administered in visit 3. Unsuccessful performance on the practice test was evident if participants were unable to respond to 19 recognition memory news event questions within the 12.8-sec time window after two attempts at the practice test (see "Procedure," visit 3).

No individuals were excluded for exhibiting impaired activities of daily living, impaired MMSE score, or impaired reading ability after the first or second visit. Sixteen individuals were deemed ineligible after the first or second visit due to moderate or severe depression symptoms ($n = 3$), uncontrolled high blood pressure ($n = 2$), or inability to complete the news events practice test ($n = 1$), or because they developed a condition found in our exclusion criteria prior to completing RM-NET ($n = 1$) or because they withdrew from the study prior to completing RM-NET ($n = 5$). Because part of the testing occurred in the MRI scanner (see "Procedure," below), we also excluded participants due to brain abnormality (e.g., brain tumor) detected on structural MRI ($n = 2$) or MRI claustrophobia or incompatibility ($n = 2$). The remaining 51 participants (19 women) completed the study.

Materials

Traditional neuropsychological assessment

The Mini-Mental State Examination (MMSE) was administered as a general measure of cognition. We also completed a comprehensive neuropsychological assessment using 26 measures to assess cognitive ability in five domains (four to seven tests/domain): EM, SM/language, executive functions, attention/processing speed, and visuospatial functions. The measures for each domain were as follows:

Episodic memory. Verbal memory: CVLT-II: trials 1–5 total recall, recognition (d′), and long-delay free recall; and WMS-IV (or –R): logical memory immediate recall (sum of stories A and B) and logical memory delayed recall (sum of stories A and B). Nonverbal memory: WMS-IV visual reproduction: immediate recall (sum of items 1–5) and delayed recall (sum of items 1–5).

Semantic memory/language. DKEFS verbal fluency: letter (sum of correct responses) and category (sum of correct responses), Multilingual Naming Test (sum of correct responses with or without semantic cue), and Multilingual Aphasia Exam token test (sum of correct responses).

Executive functions. WAIS-IV digit span: backward (sum of correct items); DKEFS verbal fluency switching (total switching accuracy), trail making: condition 4 (Total Time); and Wisconsin card sorting-64 (number of categories completed and number of perseverative errors).

Attention/processing speed. DKEFS trail making: condition 2 (total time); WAIS-IV digit span sequencing (sum of correct items): digit span forward (sum of correct items) and digit vigilance test (total time and number of errors).

Visuospatial functions. Clock Drawing: command and copy; MMSE overlapping pentagons item; WASI-II block design (sum of correct items); and WMS-IV visual reproduction copy (sum of items 1–5).

Retrograde memory news events test (RM-NET)

A 231-item test was constructed to create a reliable measure of news event memory for the entire adult life span. The procedures used to carry out pilot testing during the creation of the test were approved by the Institutional Review Boards at the University of California at San Diego or the Veterans Affairs San Diego Healthcare System. Construction of the test occurred in four phases.

Phase 1: identification of time periods and refinement of test items available from prior tests. Data collection with the RM-NET was to begin in 2018, so that year was identified as the starting point of the test. Three-year time periods were created for the most recent 15 yr (2017–2003) to obtain high temporal resolution of memories as they transition from being hippocampus-dependent to hippocampus-independent (Kapur and Brooks 1999; Manns et al. 2003; Bayley et al. 2006). Five-year time periods were created for the more remote years (see Table 1) going back to 1948, so that memory would be queried from 2017 until the time participants were ∼15 yr old (anticipated age range of participants = 65–90 yr of age). In addition, we included 20 items per time period for the most recent 30 yr (to allow for FMRI analysis [to be reported in a separate publication]) and eight to 10 items for more remote time periods to obtain reliable estimates of older memories. Next, we identified suitable test items (recall and recognition questions) from the 314 items available from previous studies (Smith et al. 2010; Smith 2014). Specifically, there were 314 items covering events that occurred between 1931 and 2005. To identify the most suitable items to retain for detecting subtle impairment in SM, items associated with enduring events (news facts that are repeatedly encountered across the lifetime; e.g., what happened to the World Trade Center in New York City) were eliminated. Enduring events can be identified objectively when individuals who did and who did not live through the event obtain similar accuracy scores. Pilot testing of young adults ($N = 24$, aged 18–24 yr) via Amazon's Mechanical Turk was carried out in 2017 to obtain accuracy scores for questions from 1948 to 1997, when these individuals were <15 yr of age. Test items were eliminated if the young adults exhibited recall accuracy similar to that of the older adults ($N = 21$, aged 65–89 yr) (from Smith 2014). The remaining items were thought to represent transient events, associated with a single or limited encoding period, resulting in a reliable estimate of memory age.

Phase 2: creation of test items for events that occurred between 2006 and 2017. The investigators followed the same rules for selecting topics of interest as were used to select topics for previous versions of the test (Squire 1974; Manns et al. 2003; Smith and Squire 2009; Smith et al. 2010; Smith 2014). Test items were created for events that were likely to receive transient media coverage (transient events; e.g., South Korea impeaching their first female president) and attempting to exclude topics that were likely to become enduring events. Similar to how events were selected for these previous news events tests, the events were identified from year-end summaries of notable events and reflect different genres of events (e.g., entertainment, sports, politics, crime, and human interest). Unlike previous studies, we obtained this information by searching the Internet and reviewing reputable news sources (e.g., "2015 year in review") instead of reviewing published periodicals. We attempted to select events in the same way for each year and so that events from year to year would be expected to have experienced the same extent of initial exposure. Events of regional interest were avoided so that the test would have more widespread utility. Recall and recognition memory versions of the test were constructed by creating seven to eight questions associated with events from each year. For the recognition test, foils were first fashioned by creating plausible incorrect answers.

Pilot testing ($N = 10$ adults) was carried out to identify poorly worded questions and to generate more plausible foils for the recognition test (i.e., by using their incorrect answers from the recall portion of the test as alternative choices in the recognition test). See Table 2 for examples of new test items.

Phase 3: matching test difficulty with previous news event tests. The test difficulty was assessed and adjusted so that RM-NET accuracy scores matched accuracy scores from successful news event memory publications (Cohen and Squire 1981; Kopelman 1989; Manns et al. 2003). To achieve this, pilot testing of the available 165 questions for events that occurred between 2017 and 1988 ($N = 24$ older adults, aged 65–76 yr) was carried out via Mechanical Turk. The mean accuracy for each 5-yr time period was compared with the accuracy for the same 5-yr time period from the previous publications (i.e., accuracy from events that occurred 1–5 yr prior to testing). Questions were eliminated from each time period until the mean accuracy matched the mean accuracy from previous tests. Specifically, if the accuracy was higher than previous tests, questions with high accuracy were eliminated, and if the accuracy was lower than previous tests, questions with low accuracy were eliminated. Five questions were eliminated in this phase, leaving 160 questions for the events from 2017 to 1988.

Phase 4: creation of a surprise recognition memory posttest for the RM-NET to measure anterograde episodic memory. To assess incidental encoding ability during the news event test, we created a follow-up test that measured subsequent memory for the specific topics of the news event questions (Smith and Squire 2009). For the 160 questions for news events that occurred 1–30 yr before testing (2017–1988), we created a three-alternative, forced-choice recognition memory test that queried the specific topic asked about earlier. Foil options were created so that memory of the general news event topic would not be sufficient to guide memory judgments. Instead, participants had to remember the details of what the question had asked about for each event (see Table 2 for examples). Based on pilot testing ($N = 24$ older participants obtained via Mechanical Turk), accuracy on this test was 90% correct after a 20-min delay. By administering the 160-item posttest, one can obtain a robust measure of EM in addition to the robust measure of SM provided by the RM-NET itself.

News habits questionnaire

It may be important to know the extent to which participants follow the news when interpreting performance on the RM-NET (Johnson and Klingler 1976; Howes and Katz 1988; Kapur et al. 1999). Accordingly, participants were asked how often they follow the news to obtain a measure of general exposure frequency (never = 1 point, rarely = 2 points, somewhat often = 3 points, and frequently = 4 points). Next, participants were asked to rank seven news sources in sequential order by how often they used them to follow the news (1 = most used, 7 = least used, and 0 = never used). These news sources were television, word of mouth, radio, periodicals, news websites (e.g., BBC.com and CNN.com), news repository websites (e.g., Buzzfeed news and Yahoo news), and/or social media (e.g., Facebook and Twitter). The total number of news sources was computed by adding up the number of sources that were selected, irrespective of their ranked order.

Procedure

All procedures were approved by the Institutional Review Board at the Veterans Affairs San Diego Healthcare System. The experiment took place across four visits. Participants were first assessed with a comprehensive neuropsychological battery (visits 1 and 2). If neuropsychological testing had been completed elsewhere within 12 mo of administration of the RM-NET (visits 3 and 4), these data were used, reducing the number of tests administered in visit 2. A medical history questionnaire was also completed in visit 2. The recognition memory portion of the RM-NET was administered between March 2018 and April 2020. Test items that occurred between 2017 and 1988 were administered inside the MRI scanner (visit 3), while test items that occurred between 1987 and 1948 were administered outside the MRI scanner (visit 4) using similar methods.

Prior to completing the RM-NET, participants were given the opportunity to become familiar with the task by completing a 16-item practice version of the RM-NET and odd/even judgment test (see below). Participants completed five blocks of the task, where each block consisted of recognition memory judgments followed by confidence judgments for 32 news event questions (four items from each of the eight time periods covering 2017–1988) (see Table 1) using an MRI-compatible four-button response box (Current Designs). The order of the news event questions was counterbalanced across participants. Because data were collected during an FMRI experiment, limited time was allowed for responses. Each news event question was presented for 12.8 sec, during which time participants made a four-choice recognition memory judgment (the question remained on the screen for the full duration regardless of the participant's response). Because participants could not see their hands during the test, they used a finger-button map (presented next to the news event question) to indicate whether their response was A (index finger), B (middle finger), C (ring finger), or D (pinky finger). After the 12.8 sec had elapsed, participants then had 3.2 sec to provide a four-choice rating to indicate their confidence that they had selected the correct answer to the news event question (4 = definitely sure [index finger], 3 = probably sure [middle finger], 2 = somewhat sure [ring finger], or 1 = pure guess [pinky finger]).

In between news event questions, participants were presented odd/even judgment trials. For these trials, a single-digit number (between 1 and 8) was presented and participants had 3.2 sec to make their even (index finger) or odd (pinky finger) judgment. A variable number of odd/even judgment trials occurred after each news event trial (zero to seven odd/even trials). The odd/even judgments facilitated analysis of brain activity, which will be reported separately. For all questions, participants were allowed to change their responses so long as the reselection was made within the time allotted and their last response was taken as their final answer.

About 20 min after completing the first 160 test items, participants began the 160-item posttest to obtain additional information about each news event (i.e., news events from 2017 to 1988). Participants were asked questions about each news event. First, they were asked to identify the specific topic they had been asked about earlier (see Table 2). This surprise recognition memory test reflects incidental subsequent memory for the RM-NET content (EM). Next, they were shown the news event question and asked to indicate on a 10-point scale the depth of their knowledge about the news event (1 = none and 10 = a lot). They were instructed to make this judgment regardless of whether they believed they knew the answer to the news event question. This component provides a measure of the quality of the news event memory to facilitate neuroimaging analysis. Finally, participants indicated whether or not they had a specific autobiographical memory associated with the news event using the definition of an EM from the Autobiographical Memory Interview (i.e., a score of three on this measure) (Kopelman et al. 1989); for example, if they could report specific details about the time and place when they learned about the event and not whether they simply remembered hearing it on the radio or seeing it on television. This component provides information about whether episodic memories accompany the semantic, news event memories. Note that very few such events were reported by participants (see Table 3), limiting discussion of findings for this variable. Eprime software (Psychology Software Tools, Inc.) was used to administer the test. Responses were made using a computer keyboard and unlimited time was given. Participants typically took between 60 and 120 min to complete the posttest.

At the final visit (visit 4, 29.4 d ± 38.5 d after visit 3), participants completed the remaining items from the RM-NET, seated at a table using a laptop computer. The test was comprised of up to 71 questions about news events from 1987 to 1948 (see Table 1). News events that occurred when participants were <15 yr of age were not administered. The test was administered using the

same methods that were used for visit 3 (inside the MRI scanner), except that an external keyboard was used to make the responses. To make the testing experience more similar to the experience in the MRI scanner, participants were not allowed to view their hand to make the memory judgments. Instead, the external keyboard was placed out of view.

Data analysis

There were two primary objectives. First, we examined the relationship between performance on the RM-NET and performance on traditional neuropsychological tests. Second, we examined whether the strength of the relationship between the RM-NET and EM (i.e., the effect size) decreased as memory age increased. Variables of interest were created prior to carrying out group-level parametric statistics. Descriptive statistics and exploratory graphing were used to assess natural data distributions, normality, and homogeneity. Means and standard deviations are reported, unless explicitly noted otherwise. Analyses were carried out using SPSS version 27.

Variables of interest

RM-NET components

For each of the 16 time periods, mean news event memory accuracy, confidence, and response times were computed (2017–1948, 1–70 yr before testing) (see Table 1). News event memory accuracy reflected the percentage of correctly answered news event questions relative to the total number of questions in the time period. The same method was used for confidence and response time components. For trials where participants failed to provide a response for a news event question within the allotted time (12.8 sec), the trial was counted as incorrect, no confidence measure was assigned, and a response time of 12.8 sec was assigned. Next, we computed total accuracy, confidence, and response time scores across the entire adult life span by averaging the mean scores from each time period that was available for each participant (2017–2015, 2014–2012, and so on until reaching the last time period where the participant was 15 yr of age). In this way, each time period received equal weighting regardless of how many questions were available for the time period. For the three most remote time periods (see Table 1), data were only available from the oldest participants (<65% of the sample); therefore, these time periods were excluded from subsequent analyses as a function of time period.

For news events that occurred between 2017 and 1988 (1–30 yr before testing), additional information was available from the posttest: subsequent memory accuracy, amount of knowledge reported, and presence of autobiographical memories. For these components, mean scores were computed for each time period in the same way as described above, except that there were only eight time periods represented instead of 13. Total mean scores were also computed for these variables by averaging the mean scores across the time periods.

Cognition domain composite scores

Performance on individual tests from the neuropsychological test battery were converted into *Z*-scores based on published norms. Composite scores (*Z*-scores) were computed for each participant and domain by averaging the individual *Z*-scores for the tests in that domain, resulting in five composite scores representing the five cognitive domains assessed. Normative data for all but three tests were developed by the University of California at San Diego Alzheimer's Disease Research Center normal cohort or elsewhere (e.g., National Alzheimer's Coordinating Center Uniform Data Set; Mayo older American normative studies [Heaton et al. 2004; Steinberg et al. 2005]). For three measures, norms were not available and performance was rated as impaired or unimpaired according to published methods. For both Clock Drawing command and Clock Drawing copy, *Z*-scores were based on the mean and SD reported in Rouleau et al. (1996). The drawing for overlapping pentagons was taken from the MMSE and scored on eight criteria (Jefferson et al. 2002), where a score of three or more errors was considered impaired. For WMS-IV visual reproduction copy and Wisconsin card sorting number of categories completed, scoring was based on the normed percentile ranking, and a score less than the 16th percentile was considered impaired. In order to include these five tests in the cognitive domain composite scores, impaired scores on these measures were converted to a *Z*-score of −1.1. This value was selected because performance of more than one standard deviation below norms has been used as a cutoff to identify mild cognitive impairment (Jak et al. 2009; Bondi et al. 2014). Unimpaired scores on these three measures were converted to a *Z*-score of 0. For the Wisconsin card sorting task, the lower *Z*-score from its two measures was used for the executive function composite score.

Identifying relevant covariates for primary analyses

Participant characteristics, such as age, education, and news habits, can sometimes affect performance on news events tests (Warrington and Sanders 1971; Johnson and Klingler 1976; Kapur et al. 1999). Therefore, prior to carrying out the primary analyses, we used stepwise, multiple regression to identify covariates that significantly predicted dependent variables of interest (i.e., RM-NET components and cognitive domain composite scores) (see Table 3). Covariates that were significant predictors ($P < 0.05$) for any variables of interest were included as covariates in the primary analyses. Estimated marginal means and SDs are reported and reflect scores adjusted for covariates.

The participant characteristics evaluated included traditional demographic characteristics (age, gender, education, and race/ethnicity), medical health and mental health comorbidity burden (see "Comorbidities," below), and news habits (sum of frequency of news exposure and number of news sources). Finally, because all participants were administered the same RM-NET items during the 2-yr data collection interval, each participant had a unique duration between the date of testing and the years in which the news events occurred. To take this test interval into consideration and to control for possible effects of forgetting over that interval, we created a covariate (RM-NET interval) for each participant. The RM-NET interval reflected the number of days that elapsed between when each participant was tested relative to when the first participant was tested (i.e., the relative dates of visit 3).

Comorbidities

Due to the small sample size and wide variety of comorbidities and medications reported by participants, composite measures of comorbidity burden were created following the method developed by Charlson et al. (1987). The measures were modified so they would reflect the disorders queried in the self-reported medical history questionnaire but without assigning weightings based on disease severity. Composite scores for total medical health burden and total mental health burden were computed separately for each participant. Each composite score reflected the sum of the reported comorbidities. For comorbidities that comprised multiple subconditions (listed in parentheses below), the comorbidity was counted if at least one of the subconditions was reported.

The total mental health burden composite reflected five comorbidities (scale 0–5): posttraumatic stress disorder (PTSD), substance abuse (substance abuse, history of alcohol/drug treatment, or history of nicotine abuse), mood disorders (depression, taking antidepressant medication, anxiety, or taking antianxiety medication), sleep disorders (sleep apnea or sleep disorder), and mental health treatment (history of mental health treatment).

The total medical health burden composite score reflected 10 comorbidities (scale 0–10): cancer, hypertension (hypertension or taking antihypertensive medication), vascular (blood clots, high cholesterol, or other vascular disorder), cardiac (heart disease), immune system (autoimmune disease or taking corticosteroids), renal (kidney disorder or adrenal gland disorder), hepatic (liver disorder), respiratory (emphysema or asthma), endocrine–metabolic (type II

diabetes, taking diabetes medication, or thyroid gland disorder), pain/fatigue (chronic fatigue, fibromyalgia, or pain), and neurological disorders (concussion, headaches, or other neurological disorders).

Primary analyses

Identifying relationships between RM-NET memory accuracy and performance on traditional neuropsychological tests

We sought to answer the question of how performance on news events tests relates to performance on traditional neuropsychological tests. We hypothesized that performance on the RM-NET would significantly predict performance on tests of EM and SM/language. To test these hypotheses, we used two approaches: (1) a theoretical approach and (2) a data-driven approach. For the theoretical approach, bivariate Pearson correlations were computed between RM-NET memory accuracy scores and each cognitive domain composite score. To correct for the possible influence of covariates, we also carried out hierarchical, multiple regression analyses using RM-NET memory accuracy to predict mean performance in each domain of cognition (composite *Z*-scores for EM, SM/language, executive functions, attention/processing speed, and visuospatial functions). Relevant covariates were entered into the model followed by RM-NET memory accuracy. Each cognitive domain composite score served as the dependent variable. Although our goal was to report relative effect sizes (correlation coefficients or beta coefficients), probability values are also reported, correcting for multiple comparisons across each of the five domains ($P < 0.05/5 = P < 0.01$).

We also carried out a data-driven approach because it is possible that strong relationships between individual neuropsychological tests and the RM-NET could be concealed when averaging across individual tests to create the cognitive domain composite scores. Due to the large number of tests, bivariate correlational analysis was not used. Instead, we carried out an exploratory factor analysis that included all standard neuropsychological test *Z*-scores as well as RM-NET accuracy scores. The principal component method was used for extraction of factors, and factors with eigenvalues >1 were retained. The goal was to identify which factors the RM-NET loaded on most strongly according to rotated factor loadings.

Identifying whether relationships between RM-NET memory accuracy and episodic memory change with memory age

We tested the hypothesis that the effect size between news event memory and EM changed with the age of the news event memory (memory age). First, for each of the 12 time periods, a regression model (same as described above) was used to obtain the effect sizes (unstandardized beta coefficients) when using the RM-NET time period accuracy score to predict the EM composite score. Next, multiple regression was used to test whether the age of the news event memory predicted these 12 effect sizes. The age of the news event memory variable was created by taking the first number from each time period (1- to 3-yr time period = 1, 4- to 6-yr time period = 4, 7- to 9-yr time period = 7, and so on), resulting in a variable with 12 values. We then tested for a linear relationship between this variable and the 12 unstandardized beta coefficients. The robustness of this analysis was tested by permuting the RM-NET accuracy scores across participants 1000 times and then repeating the regression analysis described above. The number of times that total variance accounted for (R^2) of the permuted data exceeded the R^2 of the nonpermuted data was counted and taken to reflect the probability of observing changes in the effect sizes that could have occurred by chance. The analysis of time periods was limited to news events from the last 50 yr (12 time periods). For comparison, we also carried out this analysis for the other domains of cognition.

Acknowledgments

This work was supported by Merit Award I01CX001375 (to C.N.S) from the US Department of Veterans Affairs Clinical Sciences Research and Development Service. We thank Julia Gulliver for assistance developing the RM-NET, and Irene Chu, Aaron Creswell, Matthew Koester, and Eunice Lee for assistance with data collection and data preprocessing. We thank Du Xiang for assistance with violin plots. The contents of this publication do not represent the views of VA or the United States Government.

References

Anastasi A, Urbina S. 1997. *Psychological testing*. Prentice Hall, Upper Saddle River, N.J.

Barbeau EJ, Didic M, Joubert S, Guedj E, Koric L, Felician O, Ranjeva JP, Cozzone P, Ceccaldi M. 2012. Extent and neural basis of semantic memory impairment in mild cognitive impairment. *Journal of Alzheimer's Dis* **28:** 823–837. doi:10.3233/JAD-2011-110989

Bayley PJ, Gold JJ, Hopkins RO, Squire LR. 2005. The neuroanatomy of remote memory. *Neuron* **46:** 799–810. doi:10.1016/j.neuron.2005.04.034

Bayley PJ, Hopkins RO, Squire LR. 2006. The fate of old memories after medial temporal lobe damage. *J Neurosci* **26:** 13311–13317. doi:10.1523/JNEUROSCI.4262-06.2006

Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, Nation DA, Libon DJ, Au R, Galasko D, et al. 2014. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. *J Alzheimer's Dis* **42:** 275–289. doi:10.3233/JAD-140276

Bright P, Buckman JR, Fradera A, Yoshimasu H, Colchester ACF, Kopelman MD. 2006. Retrograde amnesia in patients with hippocampal, medial temporal, temporal lobe, or frontal pathology. *Learn Mem* **13:** 545–557. doi:10.1101/lm.265906

Charlson ME, Pompei P, Ales KL, MacKenzie CR. 1987. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* **40:** 373–383. doi:10.1016/0021-9681(87)90171-8

Cohen NJ, Squire LR. 1981. Retrograde amnesia and remote memory impairment. *Neuropsychologia* **19:** 337–356. doi:10.1016/0028-3932(81)90064-6

Dede AJO, Smith CN. 2016. The functional and structural neuroanatomy of systems consolidation for autobiographical and semantic memory. *Curr Top Behav Neuroscis* **37:** 119–150. doi:10.1007/7854_2016_452

De Simone MS, De Tollis M, Fadda L, Perri R, Caltagirone C, Carlesimo GA. 2020. Lost or unavailable? Exploring mechanisms that affect retrograde memory in mild cognitive impairment and Alzheimer's disease patients. *J Neurol* **267:** 113–124. doi:10.1007/s00415-019-09559-8

Flicker C, Ferris SH, Crook T, Bartus RT. 1987. Implications of memory and language dysfunction in the naming deficit of senile dementia. *Brain Lang* **31:** 187–200. doi:10.1016/0093-934X(87)90069-1

Folstein MF, Folstein SE, McHugh PR. 1975. Mini-mental state—practical method for grading cognitive state of patients for clinician. *J Psychiatr Res* **12:** 189–198. doi:10.1016/0022-3956(75)90026-6

Gilboa A, Ramirez J, Kohler S, Westmacott R, Black SE, Moscovitch M. 2005. Retrieval of autobiographical memory in Alzheimer's disease: relation to volumes of medial temporal lobe and other structures. *Hippocampus* **15:** 535–550. doi:10.1002/hipo.20090

Grilli MD, Verfaellie M. 2014. Personal semantic memory: insights from neuropsychological research on amnesia. *Neuropsychologia* **61:** 56–64. doi:10.1016/j.neuropsychologia.2014.06.012

Grilli MD, Verfaellie M. 2016. Experience-near but not experience-far autobiographical facts depend on the medial temporal lobe for retrieval: evidence from amnesia. *Neuropsychologia* **81:** 180–185. doi:10.1016/j.neuropsychologia.2015.12.023

Heaton RK, Miller SW, Taylor MJ, Grant I. 2004. *Revised comprehensive norms for an expanded Halstead-Reitan Battery: demographically adjusted neuropsychological norms for African-American and Caucasian adults*. Psychological Assessment Resources, Inc., Lutz, FL.

Howes JL, Katz AN. 1988. Assessing remote memory with an improved public events questionnaire. *Psychol Aging* **3:** 142–150. doi:10.1037/0882-7974.3.2.142

Irish M, Vatansever D. 2020. Rethinking the episodic-semantic distinction from a gradient perspective. *Curr Opin Behav Sci* **32:** 43–49. doi:10.1016/j.cobeha.2020.01.016

Irish M, Lawlor BA, O'Mara SM, Coen RF. 2010. Exploring the recollective experience during autobiographical memory retrieval in amnestic mild cognitive impairment. *J Int Neuropsychol Soc* **16:** 546–555. doi:10.1017/S1355617710000172

Irish M, Addis DR, Hodges JR, Piguet O. 2012. Considering the role of semantic memory in episodic future thinking: evidence from semantic dementia. *Brain* **135:** 2178–2191. doi:10.1093/brain/aws119

Jak AJ, Bondi MW, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon DP, Delis DC. 2009. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry* **17:** 368–375. doi:10.1097/JGP.0b013e31819431d5

Jefferson AL, Cosentino SA, Ball SK, Bogdanoff B, Leopold N, Kaplan E, Libon DJ. 2002. Errors produced on the mini-mental state examination and neuropsychological test performance in Alzheimer's disease, ischemic vascular dementia, and Parkinson's disease. *J Neuropsych Clin N* **14:** 311–320. doi:10.1176/jnp.14.3.311

Johnson JH, Klingler DE. 1976. Questionnaire technique for measurement of episodic long-term memory. *Psychol Rep* **39:** 291–298. doi:10.2466/pr0.1976.39.1.291

Kapur N, Brooks DJ. 1999. Temporally-specific retrograde amnesia in two cases of discrete bilateral hippocampal pathology. *Hippocampus* **9:** 247–254. doi:10.1002/(SICI)1098-1063(1999)9:3<247::AID-HIPO5>3.0.CO;2-W

Kapur N, Thompson P, Kartsounis LD, Abbott P. 1999. Retrograde amnesia: clinical and methodological caveats. *Neuropsychologia* **37:** 27–30. doi:10.1016/S0028-3932(98)00065-7

Kopelman MD. 1989. Remote and autobiographical memory, temporal context memory and frontal atrophy in Korsakoff and Alzheimer patients. *Neuropsychologia* **27:** 437–460. doi:10.1016/0028-3932(89)90050-X

Kopelman MD, Wilson BA, Baddeley AD. 1989. The autobiographical memory interview: a new assessment of autobiographical and personal semantic memory in amnesic patients. *J Clin Exp Neuropsychol* **5:** 724–744. doi:10.1080/01688638908400928

Langlois R, Joubert S, Benoit S, Dostie V, Rouleau I. 2016. Memory for public events in mild cognitive impairment and Alzheimer's disease: the importance of rehearsal. *J Alzheimers Dis* **50:** 1023–1033. doi:10.3233/JAD-150722

Leyhe T, Muller S, Milian M, Eschweiler GW, Saur R. 2009a. Impairment of episodic and semantic autobiographical memory in patients with mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia* **47:** 2464–2469. doi:10.1016/j.neuropsychologia.2009.04.018

Leyhe T, Saur R, Eschweiler GW, Milian M. 2009b. Clock test deficits are associated with semantic memory impairment in Alzheimer disease. *J Geriatr Psychiatry Neurol* **22:** 235–245. doi:10.1177/0891988709335798

Leyhe T, Muller S, Eschweiler GW, Saur R. 2010. Deterioration of the memory for historic events in patients with mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia* **49:** 4093–4101. doi:10.1016/j.neuropsychologia.2010.10.011

Loeb P. 1996. *Independent living scales*. The Psychological Corporation, San Antonio, TX.

Loewenstein DA, Acevedo A, Agron J, Duara R. 2007. Stability of neurocognitive impairment in different subtypes of mild cognitive impairment. *Dement Geriatr Cogn Disord* **23:** 82–86. doi:10.1159/000097304

Manns JR, Hopkins RO, Squire LR. 2003. Semantic memory and the human hippocampus. *Neuron* **37:** 127–133. doi:10.1016/S0896-6273(03)00146-6

Marr D. 1971. Simple memory: a theory for archicortex. *Philos Trans R Soc Lond Series B* **262:** 23–81. doi:10.1098/rstb.1971.0078

Milton F, Muhlert N, Pindus DM, Butler CR, Kapur N, Graham KS, Zeman AZ. 2010. Remote memory deficits in transient epileptic amnesia. *Brain* **133:** 1368–1379. doi:10.1093/brain/awq055

Moscovitch M, Rosenbaum RS, Gilboa A, Addis DR, Westmacott R, Grady C, McAndrews MP, Levine B, Black S, Winocur G, et al. 2005. Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory. *J Anat* **207:** 35–66. doi:10.1111/j.1469-7580.2005.00421.x

Murphy KJ, Troyer AK, Levine B, Moscovitch M. 2008. Episodic, but not semantic, autobiographical memory is reduced in amnestic mild cognitive impairment. *Neuropsychologia* **46:** 3116–3123. doi:10.1016/j.neuropsychologia.2008.07.004

Nadel L, Samsonovich A, Ryan L, Moscovitch M. 2000. Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus* **10:** 352–368. doi:10.1002/1098-1063(2000)10:4<352::AID-HIPO2>3.0.CO;2-D

Orlovsky I, Huijbers W, Hanseeuw BJ, Mormino EC, Hedden T, Buckley RF, LaPoint M, Rabin JS, Rentz DM, Johnson KA, et al. 2018. The relationship between recall of recently versus remotely encoded famous faces and amyloidosis in clinically normal older adults. *Alzheimers Dement* **10:** 121–129. doi:10.1016/j.dadm.2017.11.003

Palmer BW, Boone KB, Lesser IM, Wohl MA. 1998. Base rates of 'impaired' neuropsychological test performance among healthy older adults. *Arch Clin Neuropsychol* **13:** 503–511.

Petrican R, Gopie N, Leach L, Chow TW, Richards B, Moscovitch M. 2010. Recollection and familiarity for public events in neurologically intact older adults and two brain-damaged patients. *Neuropsychologia* **48:** 945–960. doi:10.1016/j.neuropsychologia.2009.11.015

Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. 1982. Measurement of functional activities in older adults in the community. *J Gerontol* **37:** 323–329. doi:10.1093/geronj/37.3.323

Renoult L, Davidson PS, Palombo DJ, Moscovitch M, Levine B. 2012. Personal semantics: at the crossroads of semantic and episodic memory. *Trends Cogn Sci* **16:** 550–558. doi:10.1016/j.tics.2012.09.003

Renoult L, Irish M, Moscovitch M, Rugg MD. 2019. From knowing to remembering: the semantic-episodic distinction. *Trends Cogn Sci* **23:** 1041–1057. doi:10.1016/j.tics.2019.09.008

Rouleau I, Salmon DP, Butters N. 1996. Longitudinal analysis of clock drawing in Alzheimer's disease patients. *Brain Cognition* **31:** 17–34. doi:10.1006/brcg.1996.0022

Sadek JR, Johnson SA, White DA, Salmon DP, Taylor KI, Delapena JH, Paulsen JS, Heaton RK, Grant I. 2004. Retrograde amnesia in dementia: comparison of HIV-associated dementia, Alzheimer's disease, and Huntington's disease. *Neuropsychology* **18:** 692–699. doi:10.1037/0894-4105.18.4.692

Sanders HI, Warrington DK. 1971. Memory for remote events in amnesic patients. *Brain* **94:** 661–668. doi:10.1093/brain/94.4.661

Seidenberg M, Guidotti L, Nielson KA, Woodard JL, Durgerian S, Zhang Q, Gander A, Antuono P, Rao SM. 2009. Semantic knowledge for famous names in mild cognitive impairment. *J Int Neuropsychol Soc* **15:** 9–18. doi:10.1017/S1355617708090103

Seidenberg M, Kay CD, Woodard JL, Nielson KA, Smith JC, Kandah C, Breting LMG, Novitski J, Lancaster M, Matthews M, et al. 2013. Recognition of famous names predicts cognitive decline in healthy elders. *Neuropsychology* **27:** 333–342. doi:10.1037/a0032226

Sheikh J, Yesavage J. 1986. Geriatric Depression Scale (GDS): recent findings and development of a shorter version. In *Clinical gerontology: a guide to assessment and intervention* (ed. Brink T), pp. 165–173. Howarth Press, New York.

Smith CN. 2014. Retrograde memory for public events in mild cognitive impairment and its relationship to anterograde memory and neuroanatomy. *Neuropsychology* **28:** 959–972. doi:10.1037/neu0000117

Smith CN, Squire LR. 2009. Medial temporal lobe activity during retrieval of semantic memory is related to the age of the memory. *J Neurosci* **29:** 930–938. doi:10.1523/JNEUROSCI.4545-08.2009

Smith CN, Frascino JC, Kripke DL, McHugh PR, Treisman GJ, Squire LR. 2010. Losing memories overnight: a unique form of human amnesia. *Neuropsychologia* **48:** 2833–2840. doi:10.1016/j.neuropsychologia.2010.05.025

Squire LR. 1974. Remote memory as affected by aging. *Neuropsychologia* **12:** 429–435. doi:10.1016/0028-3932(74)90073-6

Squire LR, Alvarez P. 1995. Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr Opin Neurobiol* **5:** 169–177. doi:10.1016/0959-4388(95)80023-9

Steinberg BA, Bieliauskas LA, Smith GE, Ivnik RJ. 2005. Mayo's older Americans normative studies: age- and IQ-adjusted norms for the Wechsler memory scale—revised. *Clin Neuropsychol* **19:** 378–463. doi:10.1080/13854040590945201

Thomann PA, Seidl U, Brinkmann J, Hirjak D, Traeger T, Wolf RC, Essig M, Schroder J. 2012. Hippocampal morphology and autobiographic memory in mild cognitive impairment and Alzheimer's disease. *Curr Alzheimer Res* **9:** 507–515. doi:10.2174/156720512800492558

Tulving E. 1983. *Elements of episodic memory*. Oxford University Press, Cambridge.

Venneri A, Mitolo M, De Marco M. 2016. Paradigm shift: semantic memory decline as a biomarker of preclinical Alzheimer's disease. *Biomark Med* **10:** 5–8. doi:10.2217/bmm.15.53

Waidergoren S, Segalowicz J, Gilboa A. 2012. Semantic memory recognition is supported by intrinsic recollection-like processes: 'the butcher on the bus' revisited. *Neuropsychologia* **50:** 3573–3587. doi:10.1016/j.neuropsychologia.2012.09.040

Warrington EK, Sanders HI. 1971. The fate of old memories. *Q J Exp Psychol* **23:** 432–442. doi:10.1080/14640747108400255

Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, DePalma SM, Gidding S, Jamerson KA, Jones DW, et al. 2018. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* **138:** e484–e594. doi:10.1161/HYP.0000000000000066

Received February 10, 2022; accepted in revised form July 28, 2022.

The Retrograde Memory for News Events Test (RM-NET) and the relationship between news event memory and performance on standard neuropsychological tests

Andrew T.J. Cawley-Bennett, Jennifer C. Frascino, Isabel E. Asp, et al.

Learn. Mem. 2022, **29:**
Access the most recent version at doi:10.1101/lm.053571.122