Neuropsychology
2014, Vol. 28, No. 5, 685–694

0894-4105/14/$12.00 DOI: 10.1037/neu0000070

Long-Term Spacing Effect Benefits in Developmental Amnesia: Case Experiments in Rehabilitation

Janet L. Green, Tina Weston, and Melody Wiseheart
York University

R. Shayna Rosenbaum
York University and Rotman Research Institute, Baycrest, Toronto, Ontario Canada

Objective: The spacing effect describes the typical finding that repeated items are remembered best when additional items are introduced between each repetition than when the repetitions occur in immediate succession. In this study, we investigated the nature and limits of the spacing effect in the developmental amnesic case H.C. ***Method:*** In Experiment 1, we compared the performance of H.C. to that of controls on a short-term, free recall, verbal learning spacing paradigm while controlling for retention interval (timing of item review and recall). In Experiment 2, we compared the performance of H.C. to that of controls on a multiday, cued recall, verbal learning spacing paradigm, in which memory was assessed after 1 week. ***Results:*** In both experiments, H.C. demonstrated a spacing effect comparable to the effect exhibited by controls. In Experiment 1, her final recall memory for long-lag (spaced) items was better than recall for no-lag (massed) items $t(23) = 10.99$, $p < .001$, $d = 2.5$. In Experiment 2, her final cued recall memory for next-day-reviewed (spaced) items was better than cued recall for same-day-reviewed (massed) items, $t(20) = 17.6$, $p < .001$, $d = 4.1$. ***Conclusions:*** This study demonstrates the spacing effect in a person with impaired episodic memory development and is the first to show long-term benefits of spacing in amnesia. Substantially slower learning-to-criterion suggests an alternate mechanism supporting the spacing effect, perhaps independent of the hippocampus. Spacing should be considered as a candidate memory intervention technique given its effectiveness in both short- and long-term learning settings.

Keywords: spacing effect, developmental amnesia, episodic memory, hippocampus

It is widely accepted that a to-be-learned item benefits most from repetition when additional items are introduced before it is repeated than when learning presentations occur in immediate succession. This "spacing effect" is already shaping clinical and educational practices, but its underlying mechanisms and the extent of its effectiveness across a range of populations are not yet fully understood (for a review, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). For this technique to reach its potential in applied settings, it is important to delineate how and under what conditions it works best. Here, we investigate the potential of the spacing effect to provide rehabilitation benefits in a unique person with developmental amnesia. After validating a short-term list-learning paradigm in which retention interval was held constant across massed (immediate succession) and spaced conditions, we examined whether spacing benefits would persist in a long-term paired associate paradigm.

The spacing effect was first described more than a century ago (Ebbinghaus, 1964). Since that time, researchers have found the effect to be widespread and robust to a number of manipulations. It has been demonstrated in different modalities (Hintzman, Block, & Summers, 1973); in tests requiring free recall, cued recall, and recognition (Glenberg, 1976; Madigan, 1969; Shaughnessy, Zimmerman, & Underwood, 1972); in short-term lab-based episodic memory tasks (e.g., Madigan, 1969); and in real-world learning paradigms that occur over weeks or months (Carpenter, Pashler, & Cepeda, 2009; Sobel, Cepeda, & Kapler, 2011). Importantly, there is evidence that the spacing effect helps to improve memory in patient populations (Goverover, Arango-Lasprilla, Hillary, Chiaravalloti, & Deluca, 2009; Hawley, Cherry, Boudreaux, & Jackson, 2008), including individuals with anterograde amnesia due to damage of the medial temporal lobe (MTL) and extended MTL system, who otherwise perform poorly on tests of recognition and recall (Cermak et al., 1996).

That the spacing effect can be demonstrated in a heterogeneous group of amnesic individuals is encouraging. This finding reinforces the robustness of the spacing effect, and suggests that spacing could be an effective rehabilitation strategy for a broad range of memory disorders. However, further systematic examination of the spacing effect is needed before the benefits of this strategy can be maximized in clinical settings. In order for the

This article was published Online First April 21, 2014.

Janet L. Green, Department of Psychology, York University, Toronto, Ontario, Canada; Tina Weston and Melody Wiseheart, Department of Psychology and LaMarsh Centre for Child and Youth Research, York University; R. Shayna Rosenbaum, Department of Psychology, York University, and Rotman Research Institute, Baycrest, Toronto, Ontario, Canada.

We thank H.C. and her family for their continued participation and contributions to the scientific study of memory. Thank you to R. Olsen for assistance with the MR figure. This research was supported by an Alfred P. Sloan Research Fellowship and a Canadian Institutes of Health Research New Investigator Award to R. Shayna Rosenbaum.

Correspondence concerning this article should be addressed to R. Shayna Rosenbaum, Department of Psychology, York University, 4700 Keele St., Toronto, ON M3J 1P3, Canada. E-mail: shaynar@yorku.ca

spacing effect to be clinically useful, the benefits from spacing must extend to longer, more meaningful delays. It is also important to replicate the effect while controlling for the retention interval (RI; i.e., the time that has elapsed between the second repetition of an item and its recall), especially for longer lags, as this is often overlooked but is essential in determining if spaced repetition is, in fact, responsible for any benefits to memory (Balota, Duchek, & Paullin, 1989; Cepeda et al., 2006). If RI is not controlled, the most recently presented items, which typically have the longest lags between presentations, might be remembered best on a final test.

To this end, we determined the effectiveness of two spacing-effect paradigms in H.C., a young woman who experienced significant volume reduction of her hippocampi in relation to a diagnosis of developmental amnesia (Hurley, Maguire, & Vargha-Khadem, 2011; Olsen et al., 2013; Vargha-Khadem et al., 2003). H.C. never developed the ability to fully re-experience personal episodes in memory and has tremendous ongoing difficulty committing details of new encounters to memory, though over time, she has managed to accumulate general and personal facts that are not tied to any particular event or context. Standard neuropsychological tests confirm that H.C. has relatively selective episodic memory impairment on tests of recall and, to a lesser extent, on tests of recognition (Rosenbaum et al., 2011). By testing H.C., the current study sought to determine whether the spacing effect is sufficiently robust to withstand, and perhaps alleviate, the burden of episodic memory impairment in a developmental amnesic person when controlling for retention interval and after a 1-week delay. In the event that H.C. does not show a spacing effect, it is possible to conclude with greater certainty that this is due to her hippocampal damage, which is more selective in nature than that of the adult-onset cases examined to date.

Experiment 1

Experiment 1 was conducted to confirm that H.C., like individuals with adult-onset amnesia (Cermak et al., 1996), would benefit from spaced presentations of items in a simple word-list learning task. In the list-learning task, RI was carefully controlled, such that it was similar across four lag conditions. These parameters were chosen to stay true to the paradigm used by Cermak and colleagues while eliminating the potential confound of unequal RI between lags.

Method

Participants. At the time of testing, H.C. was 22 years old. She is believed to have suffered an anoxic event in her first week of life. In a recent examination of H.C.'s medial temporal lobe subregion volumes with high-resolution MRI, it was confirmed that she experienced approximately 30% volume loss that was relatively limited to her hippocampus and equally distributed across subfields (Figure 1; Olsen et al., 2013). This selective reduction is likely responsible for H.C.'s impaired performance on real-world tests of personal episodic memory and public-event memory (Kwan, Carson, Addis, & Rosenbaum, 2010; Rosenbaum et al., 2011), as well as on lab-based list-learning tasks on which, unlike control participants, H.C.'s recognition memory failed to benefit from elaborative encoding of the words (Rosenbaum et al., 2011). Despite her impaired episodic memory, H.C. has completed 14 years of education, graduating from high school and completing 1 year of community college. Results of neuropsychological evaluations confirm a significant deficit in episodic memory in the context of intellectual function, semantic knowledge, and verbal fluency that are within normal limits, based on neuropsychological testing (see Table 1).

H.C.'s performance was compared with that of 24 undergraduate psychology students (20 females) from York University with a mean age of 20.0 years (SD = 2.66 years) and a mean of 14.64 years of education (SD = 0.86 years). The control participants had no known history of neurological or psychiatric illness and were fluent in English. All participants gave written informed consent and received monetary compensation or course credit for participation, as approved by the York University and Baycrest ethics committees before participating.

Materials and design. Word lists were created from a pool of 280 nouns from the MRC Psycholinguistics data base

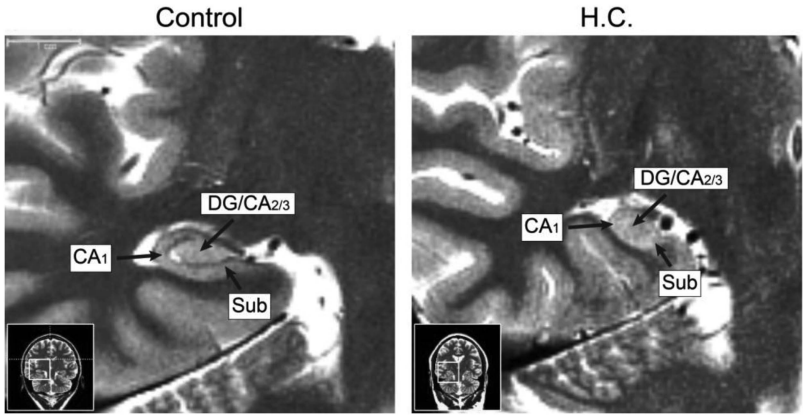

Figure 1. High-resolution T2-weighted MR images of the right hippocampus of a healthy control participant (left) and H.C. (right) in coronal views with subfields labeled (voxel size: 0.43 × 0.43 × 3 mm, no skip), acquired on a 3T Siemens Trio scanner in an oblique-coronal plane for the purpose of segmentation. See Olsen et al. (2013) for further details of volumetric analyses. Sub = subiculum; DG = dentate gyrus.

Table 1
Neuropsychological Profile of H.C.

Test	Raw score	Normed score
Intellectual function/Academic attainment		
WASI		percentile
Verbal IQ	104	61
Performance IQ	106	66
Full Scale IQ	106	66
AM-NART		standard score
Total correct	27	101.28 (estimated FSIQ)
WAIS-III		scaled score
Arithmetic	10	8
Information	19	12
Language		percentile
Boston Naming Test[1]	58	77–79
Semantic Fluency (animals)	32	> 90
Phonemic Fluency (FAS)[2]	53	70–80
WASI		T-score
Vocabulary	58	55
Anterograde memory		
WMS-III		scaled score
Logical Memory I	27	4
Logical Memory II	3	1
California Verbal Learning Test-II		z-score
Total Trials 1–5	44	38 (T-score)
Short-delay free recall	0	−4
Short-delay cued recall	5	−3.5
Long-delay free recall	3	−3
Long-delay cued recall	4	−3.5
Recognition	13	−2
Rey Osterreith complex figure[3]		T-score
Immediate recall	4	< 20
Delayed recall	3	< 20
Delayed recognition	17	22
Processing speed		
WAIS-III		scaled score
Digit Symbol	96	13
Symbol Search	45	14
Visuospatial function		Percentile
Judgement of Line Orientation	24	56
Benton Facial Recognition	45	33–59
Rey-Osterrieth Complex Figure – Copy[3]	33	> 16
WASI		T-score
Block Design	52	54
Attention and executive function		
Stroop[4]		z-score
Word full (sec)	45	3.65
Color full (sec)	48	−0.03
Interference full (sec)	80	−0.57
Trail Making Test[1]		z-score
Part A (sec)	34	0.69
Part B (sec)	55	−0.23
WASI		T-score
Similarities	35	50
Matrix Reasoning	29	55
WAIS-III		scaled score
Digit span forward	10	
Digit span backward	5	
Digit span total	15	8
Wisconsin Card Sorting Task		T-score
Categories[5]	10	57
Perseverative errors	10	

Note. The 128-card version of the Wisconsin Card Sorting Task was administered according to the Grant and Berg (1948) method. Additional results of neuropsychological testing are reported in Rosenbaum et al. (2011), Hurley et al. (2012), and Rabin et al. (2012). WASI = Wechsler Abbreviated Scale of Intelligence; AM-NART = American National Adult Reading Test; WAIS-III = Wechsler Adult Intelligence Scale–III; WMS-III = Wechsler Memory Scale-III.
[1] Spreen & Strauss (1998). [2] Tombaugh, Kozak, & Rees (1996). [3] Meyers & Meyers (1996). [4] In-house unpublished normative data. [5] Heaton et al. (1993).

(Coltheart, 1981) that were equated with respect to familiarity, concreteness, and imageability (all ratings between 500 and 700), and of low to moderate frequency, non-taboo, one to three syllables in length, and three to six letters long. Using MATLAB 7.4.0 (R2007a), words were randomly assigned to one of five lists, with each list having a total of 88 presentation positions containing 56 unique words. The first and last eight words served as primacy and recency buffers, respectively. Within the remaining 72 presentation positions, eight words were presented once and 32 words were repeated with lags of zero, one, six, and 24 intervening items (eight words per lag condition). At first, we attempted to control for RI by controlling the list position of each lag in each list. To do this, we constrained the MATLAB program, selecting the presentation order such that the average RI for all lag conditions (measured as the number of positions between the second presentation of each word and the start of the final recall test) would be within five positions of one another. However, we were unable to find more than one list that matched these criteria. Instead, we chose five lists that, when averaged together, had a difference of RI less than five serial positions. More specifically, for the 40 items at Lag 0 (8 items $\times$ 5 lists), their average RI was 39.6 positions (SD = 3.8). For the 40 items at Lag 1, their average RI was 37.8 positions (SD = 5.3). For the 40 items at Lag 6, their average RI was 42.0 positions (SD = 4.7). For the 40 items at Lag 24, their average RI was 37.6 positions (SD = 1.7). Of these four averages, each was within five positions of each other (i.e., 39.6, 37.8, 42.0, 37.6).

Procedure. Participants were given both written and verbal instructions, and completed a brief practice session before participating. All stimuli were presented using Presentation software (Version 14.1 09.21.09; www.neurobs.com). Words were presented sequentially in white uppercase letters on a black background for 1.5 s. After presentation of each list, participants were given 3 min for free recall, during which they were asked to recall as many words as possible by typing out their responses. The experiment took approximately 40 min to complete.

Results and Discussion

A single case, which does not have a mean or variance with which to compare with a group of control participants, presents a unique problem for traditional statistical analyses. Several different approaches to analyzing H.C.'s data were taken, with conclusions based on results that were consistent across the approaches. Our first approach was to visually inspect the raw data (Figures 2a and b), and then to conduct an ANOVA, which assumes H.C. to have the same variance as the control group (Corballis, 2009). Post hoc modified t tests (Crawford & Howell, 1998) were then used to compare H.C.'s performance with that of the control participants. Our second approach employed a resampling technique. Because of multiple t tests, significance levels for post hoc comparisons were calculated using Bonferroni correction.

A 2 (group: H.C., controls) $\times$ 4 (lag: 0, 1, 6, and 24 intervening items) between-within repeated measures ANOVA revealed a significant main effect of lag, $F(3, 69) = 3.1$, $p = .03$, $\eta_p^2 = .12$. For the control data only, post hoc pairwise comparisons confirmed that there was a significant difference between Lag 0 and Lag 6 ($p < .001$), Lag 0 and Lag 24 ($p < .001$), Lag 1 and Lag 6 ($p = .004$), and Lag 1 and Lag 24 ($p < .001$). When H.C.'s data were added to the pairwise comparisons (i.e., all data), the same lags were still significant (ps $< .008$). The main effect of group and the interaction were not significant (see Figure 2).

Our first statistical approach demonstrated a main effect of lag, but there are obvious difficulties in assuming that H.C. would have the same variance as a group of controls. Our second approach employed a resampling technique to better estimate H.C.'s mean and variance. More specifically, we randomly sampled (with replacement) 100% of the total 40 trials of H.C.'s performance for each lag (8 items per lag $\times$ 5 lists) and calculated the accuracy on these 40 trials. This process was repeated 24 times (per lag) to match the control group sample size in an effort to create an artificial "H.C. group" with an artificial variance that was more likely to be representative of an H.C. "population." These data were compared with the control group's data using ANOVA. This approach revealed main effects of group, $F(1, 46) = 43.2$, $p < .001$, $\eta_p^2 = .48$, and lag, $F(3, 138) = 47.3$, $p < .001$, $\eta_p^2 = .51$, and a significant interaction, $F(3, 138) = 4.7$, $p < .01$, $\eta_p^2 = 0.10$. Post hoc t tests revealed that controls significantly outperformed H.C. at lags 0, 1 and 6 (ps $< .0125$). For the control group, post hoc pairwise comparisons confirmed there were significant differences (ps < 0.008) between all lags except for Lags 0 and 1 ($p = .086$), and Lags 6 and 24 ($p = .334$). For the H.C. group, post hoc pairwise comparisons confirmed that there were significant differences (ps < 0.008) between all lags except for Lags 1 and 6 ($p = .162$). The largest spacing effect for both groups, as expected, was between Lags 0 ($M_{control} = 0.25$, $SD_{control} = 0.12$; $M_{H.C.} = 0.05$, $SD_{H.C.} = 0.05$) and 24 ($M_{control} = 0.39$, $SD_{control} = 0.16$; $M_{H.C.} = 0.25$, $SD_{H.C.} = 0.08$). Although this effect was strong for both groups, the H.C. group exhibited a larger spacing effect ($d = 2.5$) than did controls ($d = 1.4$).

Although there are slightly different results generated by the two statistical approaches, the common result that emerges is that H.C. reliably benefits from a longer lag between repeated items. These results are consistent with adult-onset amnesia data (Cermak et al., 1996) and show that an individual with lifelong episodic memory impairment can also benefit from spaced repetition of items in a word list.

Experiment 2

To see if spacing would benefit H.C. over a longer, more meaningful delay, we conducted a second experiment. Participants learned a list of paired associates, reviewed the pairs either immediately after learning or 1 day later, and were tested for their final memory of the stimuli 8 days after each review (to properly control for RI; see Figure 3). Paired associates allowed for cued recall, which is less demanding than free recall after a long RI.

Method

Participants. H.C. was 23 years old at the time of testing and continued to report 14 years of education. Her performance was compared with that of 21 new control participants (15 females) with a mean age of 21.65 years (SD = 2.32 years) and a mean of 14.94 years of education (SD = 1.54 years). Participants were

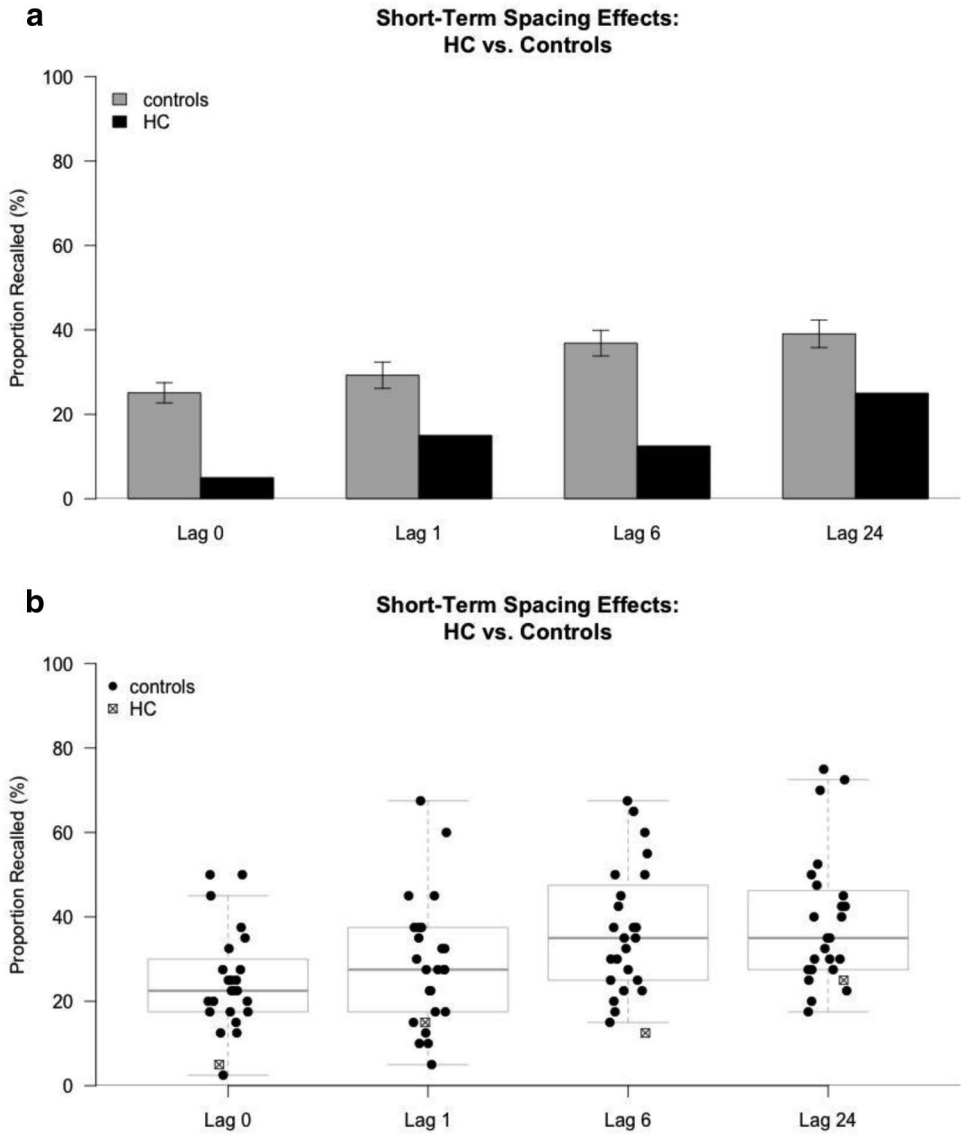

Figure 2. Both controls and H.C. demonstrated a spacing effect. (a) Bar graph of participants' averaged data. Error bars represent *SE*. (b) Dot plot of participants' raw data.

fluent in English and had no known history of neurological or psychiatric illness. Ethical considerations and participant incentives were the same as Experiment 1.

Materials. The stimuli consisted of 40 semantically unrelated word pairs (e.g., PIN–YAWN) generated by randomly combining 80 of the words used in Experiment 1 and ensuring that each word of a pair was semantically independent from the other word of the pair.[1] To ensure that both H.C. and the control group learned the stimuli well during the initial episode, we used a computer program that implemented a learning-to-criterion requirement for the initial learning episode (Cepeda et al., 2009).

For the review phase, the 40 paired associates were randomly split into two lists of 20 items. All participants reviewed one list immediately after learning ("massed review") and the other list 1 day after learning ("spaced review"). Lists were not counterbalanced (i.e., control participants completed the exact same experiment as H.C.). Piloting and post hoc analyses confirmed that learning-to-criterion rates were matched between lists. The final test was conducted online via SurveyMonkey.com.

Procedure.

Session 1 (initial learning and massed review). Participants were given written and verbal instructions prior to beginning the session. During Session 1, the 40 paired associates were presented

[1] Although H.C. viewed these words in isolation in Experiment 1, she was now being asked to use them in a new task (i.e., associating two words together). It had also been over one year since Experiment 1. We therefore felt that it was fair to assume that the words were no longer salient in her memory. The control participants had not participated in Experiment 1 and therefore had never seen the stimuli before.

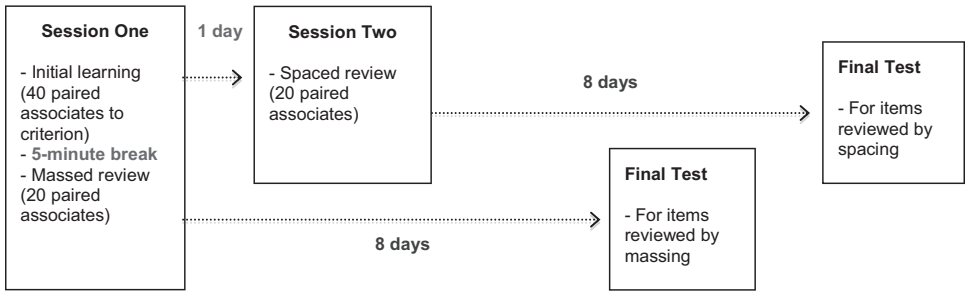

Figure 3. Visual depiction of the design of Experiment 2. Participants learned 40 paired associates in an initial learning session and, after a 5-min break, they reviewed half of the paired associates ("massed review"). A day later, they reviewed the other half of the paired associates ("spaced review"). Eight days after each respective review (to properly control for retention interval), participants were tested for their final recall memory of the items.

on a computer screen sequentially for 7 s each. Each pair was presented in black uppercase letters against a gray background, with one word presented above the other. After presentation of all 40 pairs, a cued-recall test prompted participants' memory for the stimuli. One word from the pair was presented (e.g., PIN) with a textbox underneath for participants to type the associated word from memory (e.g., YAWN). There was no time limit and a response was not required. After each trial, participants were given corrective feedback for 5 s, regardless of accuracy. The computer program tested all paired associates once in random order, then a second time in random order, and then again until the participant had correctly answered each paired associate a total of two times. Once a paired associate was correctly answered two times (i.e., learned to the criterion of twice correct), the computer program discarded it from the testing rotation.

After learning to criterion, participants had a brief 5-min delay before they reviewed half of the paired associates (i.e., the massed list). The review consisted of the same cued-recall test, except this time the 20 paired associates were tested twice in total, regardless of participants' responses. Session 1 lasted approximately one hour.

Session 2 (spaced review). One day later, participants returned to the lab to complete a review of the other half of the paired associates (i.e., the spaced list). The procedure was the same as the massed review explained in the *Session 1* section. Session 2 lasted approximately 15 min.

Session 3 (massed final test). Eight days after the massed review, participants were e-mailed an Internet link to complete a final test of the 20 massed paired associates. Again, participants viewed one word of the pair and were asked to type the associated word that they could remember. Words were presented one at a time, appeared only once, and no feedback was provided.

Session 4 (spaced final test). Eight days after the spaced review, participants were e-mailed a second Internet link to complete a final test of the 20 spaced paired associates. The procedure was the same as the massed final test explained in the *Session 3* section.

Results and Discussion

Review data (Sessions 1 and 2). For the massed review, controls performed with 98% recall accuracy and H.C. performed with 25% recall accuracy. For the spaced review, controls performed with 89% accuracy and H.C. performed with 30% accuracy.

Final test data (Sessions 3 and 4). The final test data were used as the main outcome variable in Experiment 2. As in Experiment 1, we first visually inspected the raw data of the control group compared with H.C. (Figures 4a and b). We then compared H.C.'s final test data with controls' final test data using an ANOVA. In the first analysis, a 2 (group: H.C., controls) × 2 (lag: massed, spaced) between-within repeated measures ANOVA revealed significant main effects of lag, $F(1, 20) = 17.7$, $p < .001$, $\eta_p^2 = .47$, and group, $F(1, 20) = 6.6$, $p = .019$, $\eta_p^2 = .25$. The interaction was not significant (Figures 4a and 4b).

To perform the resampling analysis, we used the same technique as in Experiment 1 to randomly generate 21 resampled data points (per lag) based on H.C.'s data. These data were compared with the control group's data using ANOVA. This approach revealed significant main effects of group, $F(1, 40) = 117.2$, $p < .001$, $\eta_p^2 = .75$, and lag, $F(1, 40) = 282.1$, $p < .001$, $\eta_p^2 = .88$. The interaction was significant, $F(1, 40) = 13.6$ $p < .01$, $\eta_p^2 = .25$. Post hoc t tests revealed that items reviewed by spacing ($M_{control} = 0.74$, $SD_{control} = 0.14$; $M_{H.C.} = 0.47$, $SD_{H.C.} = 0.10$) were better remembered than items reviewed by massing ($M_{control} = 0.46$, $SD_{control} = 0.17$; $M_{H.C.} = 0.04$, $SD_{H.C.} = 0.05$), and that the H.C. group exhibited a larger spacing effect ($d = 4.1$) than did controls ($d = 1.8$). As in Experiment 1, regardless of the statistical method employed (ANOVA vs. resampling), H.C.'s final test performance was positively influenced by spacing.

We also examined how many trials it took participants to reach criterion during Session 1. A 2 (group: H.C., controls) × 2 (review list: massed, spaced) between-within repeated measures ANOVA revealed a significant main effect of group, $F(1, 38) = 59.1$, $p < .001$, $\eta_p^2 = .61$. H.C. took more trials to reach criterion ($M = 15.10$, $SE = 1.03$) than controls ($M = 3.93$, $SE = 1.03$). The main effect of review list and the interaction were not significant.

Given the additional number of trials H.C. needed to reach criterion in Session 1, we were interested in whether rate of learning might be associated with the magnitude (or success) of the spacing effect. First, we examined whether the slowest control participant to reach criterion ($M = 6.95$, $SD = 3.46$) differed in his or her susceptibility to a spacing effect compared with the rest of

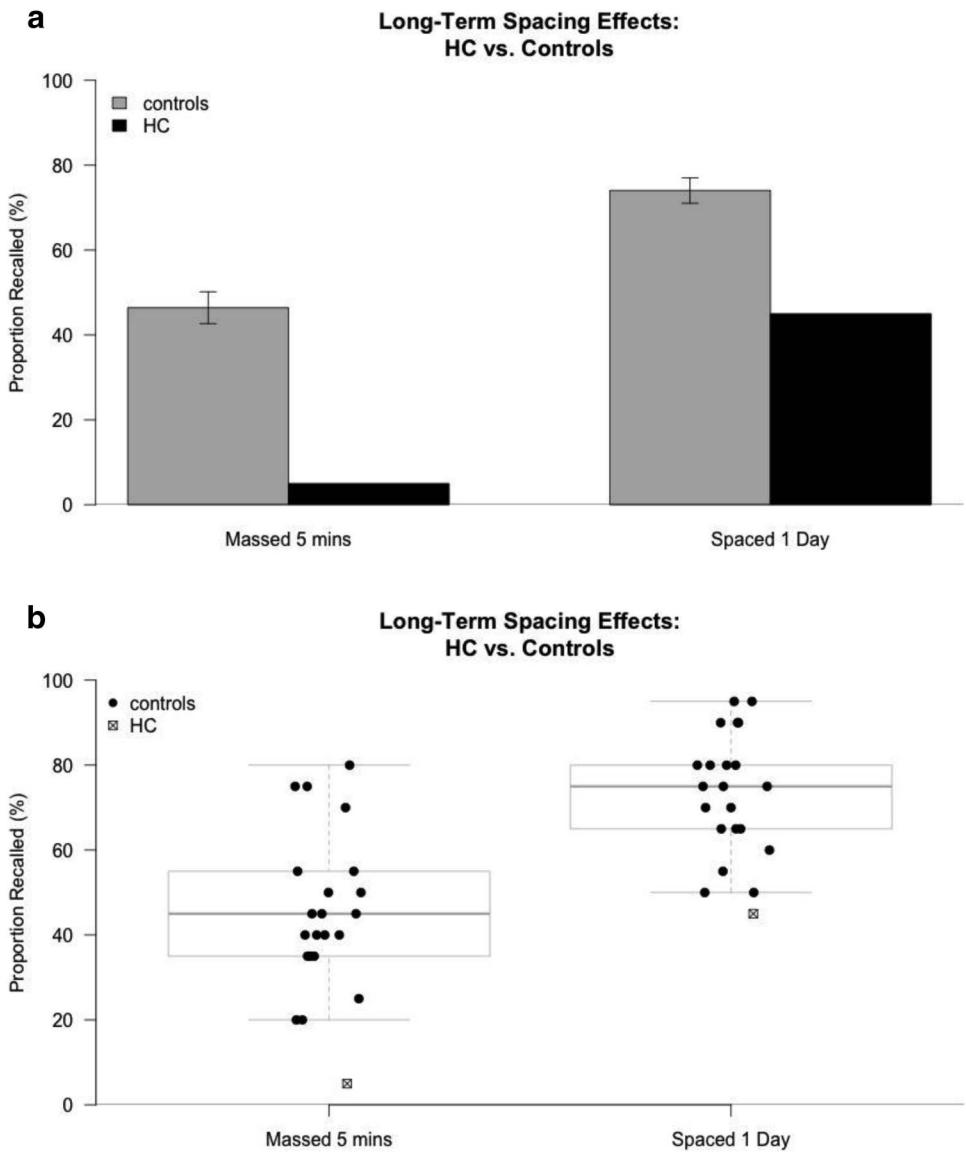

Figure 4. For both controls and H.C., word pairs that were reviewed 1 day after initial learning (spaced review) were remembered better on a final test 8 days later compared with word pairs that were reviewed immediately after initial learning (massed review; $p < .001$). (a) Bar graph of participants' averaged data. Error bars represent *SE*. (b) Dot plot of participants' raw data.

the control group. We ran a 2 (learner type: regular control, $n = 20$; slow control, $n = 1$) $\times$ 2 (lag: massed, spaced) between-within repeated measures ANOVA (on the control data only). There was a significant main effect of spacing ($p = .009$), but no significant main effect of learner type and no significant interaction. This suggests that the slowest control—a participant closest to H.C.'s learning style—observed a spacing effect comparable with controls. Second, we ran a correlation analysis examining the relationship between a control participant's spacing effect (i.e., final test spaced performance−final test massed performance) and a participant's rate of learning (i.e., mean number of trials to reach criterion in Session 1). The correlation was not significant ($r = -0.28$, $p = .23$). This result also confirms that participants of all learning ability levels show the spacing effect. Finally, we are aware of a study using the same paired associate software as this experiment, in which developmentally typical older adults took twice as many trials as young adults to reach criterion during learning (6.8 vs. 3.7, respectively), yet both groups showed an intact spacing effect (Simone, Bell, & Cepeda, 2013). Thus, older adults, who, like H.C., may take longer to learn new stimuli, also exhibit a spacing effect.

The results of Experiment 2 are the first to extend the findings of long-term spacing advantages for verbal material (e.g., Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008) to a patient with amnesia. It is noteworthy that H.C. took nearly four times as many trials per paired associate to learn to criterion compared with controls,

indicating that she may be using a different strategy to learn the material.

General Discussion

The experiments reported here provide evidence of a spacing effect in an individual with developmental amnesia over both short and long delays. Experiment 1 replicated the findings of Cermak et al. (1996) in patient H.C., whose recall performance increased by 20% between Lag 0 and Lag 24, despite the additional constraint of controlling RI across lists. These results were extended in Experiment 2, which included a longer and more ecologically relevant delay of 1 week. Despite H.C.'s lower overall memory performance, we found that her recall improved by 40% between massed and spaced repetitions 1 week later. Notably, H.C. demonstrated the typical spacing effect, as her performance was best for items repeated at the longest lags in both Experiments 1 and 2. To our knowledge, this is the first time a long-term spacing-effect paradigm has been implemented in an individual with amnesia.

Both Experiments 1 and 2 demonstrated a statistically significant spacing effect, but it is unclear which statistical method is optimal for comparing a single case to a group, especially across multiple conditions. The current study included several approaches: one that treated H.C.'s single score as a mean and adopted the same variance as the control group (Corballis, 2009), together with a conservative modified *t*-test approach for single cases (Crawford & Howell, 1998), and another in which H.C.'s data were resampled multiple times to create an artificial group. We acknowledge that each of these approaches has its weaknesses, which limit the conclusions that can be made from our results. Our first approach suffers from the assumption that H.C.'s data can be considered a mean, and that her variance would be the same as the variance of the control group. Our second approach suffers from a lack of independent observations and might be more akin to understanding the reliability of H.C.'s responses rather than creating artificial between-subjects variability in an "H.C. group." That being said, there is no perfect solution for conducting statistical analyses on a single case, yet there is benefit to studying an individual with a unique deficit such as H.C. As such, presenting converging results across multiple statistical methods strengthens the interpretation that H.C. demonstrates a true spacing effect. Even casual inspection of H.C.'s and controls' raw data clearly indicates that spacing is offering a benefit to memory (see Figures 2b and 4b). The current results build on previous findings in older adults (Balota et al., 1989; Balota, Duchek, Sergent-Marshall, & Roediger, 2006; Simone et al., 2013) and adult-onset amnesic patients (Cermak et al., 1996) to strengthen and broaden the appeal of the spacing effect as an effective and clinically meaningful memory intervention technique.

Less clear is the neural mechanism(s) responsible for producing the preserved spacing effect demonstrated in H.C. and adult-onset amnesic individuals (Cermak et al., 1996). Gene transcription and fMRI studies seem to favor a hippocampus-mediated mechanism for the spacing effect. Guzowski et al. (2006) have demonstrated that more CA1 neurons in the hippocampus transcribe the gene *Arc* (important for promoting memory consolidation) when rats have spaced exposure to an environment than when the exposure is massed, though it is unclear if this reflects the processes underlying the spacing effect or the resulting benefits to memory. A separate fMRI study by Brozinsky et al. (2005) demonstrated repetition suppression in the hippocampus during a memory task, as indicated by reduced bilateral hippocampal activity for shorter lags between repeated items at study compared with longer lags between items (see also Grill-Spector, Henson, & Martin, 2006). However, these findings appear to conflict with the current finding that an individual with impaired hippocampal function benefits from spaced repetition. One possibility is that H.C.'s residual hippocampal tissue is capable of supporting the spacing effect. Another possibility is that a spacing effect in H.C.'s performance was achieved via a neocortical compensatory mechanism, suggested by the substantially longer time that it took H.C. to reach criterion in Experiment 2 compared with controls (McClelland, McNaughton, & O'Reilly, 1995). Even if the spacing effect relies on residual hippocampal tissue, the very fact that what remains of this region is not sufficient to support episodic memory is further evidence of the robustness of the spacing effect.

It is also difficult to differentiate between popular competing spacing theories, such as encoding variability (Glenberg, 1979; Melton, 1970) and study-phase retrieval (Greene, 1989; Thios & D'Agostino, 1976), based on the current study design. According to the encoding variability view, space between repetitions is beneficial because it provides additional contextual cues to aid retrieval (Glenberg, 1979; Hintzman, 1974). Massed items are more likely to carry the same contextual cues at their first and second presentations, whereas spaced items—in which context has had more time to fluctuate—are more likely to carry different contextual cues at their first and second presentations. The greater the number of different retrieval cues associated with an item, the more likely it will be retrieved in a test scenario. Study-phase retrieval theory proposes that additional space between repetitions makes retrieval during study more effortful, as one attempts to reconstruct an item's first presentation upon seeing its second presentation. This reconstruction process is relatively easy for a massed item (having just seen the item's first presentation); however, for a spaced item, reconstruction is effortful. This effortful retrieval thereby strengthens the item's memory trace, making it more likely to be retrieved in a test scenario.

If we were to speculate as to how our results inform these theories, we could look to H.C.'s review data in Experiment 2. H.C. performed relatively poorly on each review (25% massed; 30% spaced), suggesting that it is unlikely that successful reconstruction took place. This appears to be problematic for the study-phase retrieval theory, as H.C. demonstrated a relatively intact spacing effect. In contrast, the encoding variability theory depends on variation in context between repeated items, so one might predict that H.C. would perform poorly based on her inability to encode episodic details. For example, when asked to make either a deep or shallow judgment about words during a levels-of-processing manipulation at encoding (Craik, 2002), H.C. did not benefit from deep encoding as did controls (Rosenbaum et al., 2011). This could mean that H.C. does not benefit from additional associations, and therefore might not benefit from variable contextual cues associated with spaced retrieval.

Demonstration of an intact spacing effect in H.C. in both Experiments 1 and 2 might be taken to suggest that her residual hippocampus is supporting the encoding of context, or that she is able to acquire new semantic information, which she appears to successfully accumulate over time (Rosenbaum et al., 2011), albeit

slowly with many repetitions (see Gardiner, Brandt, Baddeley, Vargha-Khadem, & Mishkin, 2008). This alternate semantic explanation might account for the additional trials required for H.C. to reach criterion in Experiment 2. It is also noteworthy that some of her errors in learning the stimuli to criterion in Experiment 2 were semantic errors (e.g., the target word was TOAST, and she responded with BREAD). These observations possibly reflect a slow-learning neocortical system (McClelland et al., 1995) that differs from the system responsible for the spacing effect that was the product of Experiment 1. Although the present experiments were not specifically designed to disentangle different spacing-effect mechanisms, future work involving neuroimaging or manipulation of context or retrievability among to-be-learned items and/or approach of testing in patients with clear functional dissociations in memory has the potential to advance theoretical models of the spacing effect.

A potential limitation of this study is that H.C. was presented with the same words in Experiment 1 and in Experiment 2 (which took place 1 year after Experiment 1). We believe that it was important to maintain strict control over the various criteria of our word lists, and that it was unlikely that H.C. retained the words over the course of a year based on a single exposure, especially given that the words were presented in a novel, random pairing in Experiment 2. Although we feel confident that our findings reflect a spacing effect, future attempts to replicate these results might include different word lists (which might necessarily require relaxed linguistic criteria).

In sum, in two experiments, we showed that an individual with developmental amnesia can experience significant gains in memory retrieval when repeated items in a study list are spaced rather than presented in immediate succession. This spacing effect held not only when RI was controlled and memory was tested following a delay of 30 min but also when a more ecologically valid delay of 1 week was imposed, which, until now, had not been demonstrated in an amnesic population. The application of a spacing strategy might be particularly effective in concert with other intervention techniques that have been shown to improve memory in adult-onset amnesia, including self-imagining (Grilli & Glisky, 2012) and errorless learning with external aids (Svoboda & Richards, 2009; Svoboda, Richards, Leach, & Mertens, 2012). Future work is also needed to determine if spaced repetition of real-world material is beneficial at even longer delays. For now, we believe that the gains experienced by H.C. thus far provide sufficient incentive for the adoption of the spacing effect as an effective method for treating episodic memory impairment, which is the unfortunate consequence of healthy aging and a wide range of neurological and psychiatric conditions.

References

Balota, D. A., Duchek, J. M., & Paullin, R. (1989). Age-related differences in the impact of spacing, lag, and retention interval. *Psychology and Aging, 4,* 3–9. doi:10.1037/0882-7974.4.1.3

Balota, D. A., Duchek, J. M., Sergent-Marshall, S. D., & Roediger, H. L. (2006). Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychology and Aging, 21,* 19–31. doi:10.1037/0882-7974.21.1.19

Brozinsky, C. J., Yonelinas, A. P., Kroll, N. E. A., & Ranganath, C. (2005). Lag-sensitive repetition suppression effects in the anterior parahippocampal gyrus. *Hippocampus, 15,* 557–561. doi:10.1002/hipo.20087

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U.S. history facts. *Applied Cognitive Psychology, 23,* 760–771. doi:10.1002/acp.1507

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology, 56,* 236–246. doi:10.1027/1618-3169.56.4.236

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132,* 354–380. doi:10.1037/0033-2909.132.3.354

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning a temporal ridgeline of optimal retention. *Psychological Science, 19,* 1095–1102. doi:10.1111/j.1467-9280.2008.02209.x

Cermak, L. S., Verfaellie, M., Lanzoni, S., Mather, M., & Chase, K. A. (1996). Effect of spaced repetitions on amnesia patients' recall and recognition performance. *Neuropsychology, 10,* 219–227. doi:10.1037/0894-4105.10.2.219

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 33,* 497–505. doi:10.1080/14640748108400805

Corballis, M. C. (2009). Comparing a single case with a control sample: Correction and further comment. *Neuropsychologia, 47,* 2696–2697. doi:10.1016/j.neuropsychologia.2009.04.012

Craik, F. I. M. (2002). Levels of processing: Past, present . . . and future? *Memory, 10,* 305–318. doi:10.1080/09658210244000135

Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *Clinical Neuropsychologist, 12,* 482–486. doi:10.1076/clin.12.4.482.7241

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology*. New York, NY: Dover.

Gardiner, J. M., Brandt, K. R., Baddeley, A. D., Vargha-Khadem, F., & Mishkin, M. (2008). Charting the acquisition of semantic knowledge in a case of developmental amnesia. *Neuropsychologia, 46,* 2865–2868. doi:10.1016/j.neuropsychologia.2008.05.021

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior, 15,* 1–16. doi:10.1016/S0022-5371(76)90002–5

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition, 7,* 95–112. doi:10.3758/BF03197590

Goverover, Y., Arango-Lasprilla, J. C., Hillary, F. G., Chiaravalloti, N., & Deluca, J. (2009). Application of the spacing effect to improve learning and memory for functional tasks in traumatic brain injury: A pilot study. *American Journal of Occupational Therapy, 63,* 543–548. doi:10.5014/ajot.63.5.543

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38,* 404–411. doi:10.1037/h0059831

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 371–377. doi:10.1037/0278–7393.15.3.371

Grilli, M. D., & Glisky, E. L. (2012). Imagining a better memory: Self-imagination in memory-impaired patients. *Clinical Psychological Science, 1,* 93–99. doi:10.1177/2167702612456464

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10,* 14–23. doi:10.1016/j.tics.2005.11.006

Guzowski, J. F., Miyashita, T., Chawla, M. K., Sanderson, J., Maes, L. I., Houston, F. P., . . . Barnes, C. A. (2006). Recent behavioral history modifies coupling between cell activity and Arc gene transcription in hippocampal CA1 neurons. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 103,* 1077–1082. doi:10.1073/pnas.0505519103

Hawley, K. S., Cherry, K. E., Boudreaux, E. O., & Jackson, E. M. (2008). A comparison of adjusted spaced retrieval versus a uniform expanded retrieval schedule for learning a name-face association in older adults with probable Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 30,* 639–649. doi:10.1080/13803390701595495

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test (WCST) Manual, revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Hintzman, D. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium* (pp. 77–99). New York, NY: Halsted Press Division of Wiley.

Hintzman, D., Block, R., & Summers, J. (1973). Modality tags and memory for repetitions: Locus of the spacing effect. *Journal of Verbal Learning & Verbal Behavior, 12,* 229–238. doi:10.1016/S0022-5371(73)80013-1

Hurley, N. C., Maguire, E. A., & Vargha-Khadem, F. (2011). Patient HC with developmental amnesia can construct future scenarios. *Neuropsychologia, 49,* 3620–3628. doi:10.1016/j.neuropsychologia.2011.09.015

Kwan, D., Carson, N., Addis, D. R., & Rosenbaum, R. S. (2010). Deficits in past remembering extend to future imagining in a case of developmental amnesia. *Neuropsychologia, 48,* 3179–3186. doi:10.1016/j.neuropsychologia.2010.06.011

Madigan, S. (1969). Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning & Verbal Behavior, 8,* 828–835. doi:10.1016/S0022-5371(69)80050-2

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102,* 419–457. doi:10.1037/0033-295X.102.3.419

Melton, A. W. (1970). The Situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior, 9,* 596–606.

Meyers, J. E., & Meyers, K. R. (1996). *Rey Complex Figure Test and Recognition Trial: Supplemental norms for children and adolescents*. Odessa, FL: PAR.

Olsen, R. K., Palombo, D., Rabin, J. S., Levine, B., Ryan, J. D., & Rosenbaum, R. S. (2013). Volumetric analysis of medial temporal lobe subregions in developmental amnesia using high-resolution MRI. *Hippocampus, 23,* 855–860.

Rabin, R. S., Carson, N., Gilboa, A., Stuss, D. T., & Rosenbaum, R. S. (2012). Early development of episodic memory is necessary for theory of mind only under some conditions: Evidence from developmental amnesia. *Frontiers in Psychology,* 3:588.

Rosenbaum, R. S., Carson, N., Abraham, N., Bowles, B., Kwan, D., Köhler, S., . . . Richards, B. (2011). Impaired event memory and recollection in a case of developmental amnesia. *Neurocase, 17,* 394–409. doi:10.1080/13554794.2010.532138

Shaughnessy, J. J., Zimmerman, J., & Underwood, B. J. (1972). Further evidence on the MP-DP effect in free-recall learning. *Journal of Verbal Learning and Verbal Behavior, 11,* 1–12. doi:10.1016/S0022-5371(72)80053–7

Simone, P. M., Bell, M., & Cepeda, N. J. (2013). Diminished but not forgotten: Effects of aging on magnitude of spacing effect benefits. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 68,* 674–680. doi:10.1093/geronb/gbs096

Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology, 25,* 763–767. doi:10.1002/acp.1747

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). New York, NY: Oxford University Press.

Svoboda, E., & Richards, B. (2009). Compensating for anterograde amnesia: A new training method that capitalizes on emerging smartphone technologies. *Journal of the International Neuropsychological Society, 15,* 629–638. doi:10.1017/S1355617709090791

Svoboda, E., Richards, B., Leach, L., & Mertens, V. (2012). PDA and smartphone use by individuals with moderate-to-severe memory impairment: Application of a theory-driven training programme. *Neuropsychological Rehabilitation, 22,* 408–427. doi:10.1080/09602011.2011.652498

Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior, 15,* 529–536.

Tombaugh, T. N., Kozak, J., & Rees, L. (1996). Normative data for the Controlled Oral Word Association Test. Personal Communication. In O. Spreen & E. Strauss (Eds.), *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed., p. 455). New York, NY: Oxford University Press.

Vargha-Khadem, F., Salmond, C. H., Watkins, K. E., Friston, K. J., Gadian, D. G., & Mishkin, M. (2003). Developmental amnesia: Effect of age at injury. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 100,* 10055–10060. doi:10.1073/pnas.1233756100

Received June 26, 2013
Revision received January 10, 2014
Accepted January 10, 2014 ■