Does Testing Improve Learning for College Students With Attention-Deficit/Hyperactivity Disorder?

Laura E. Knouse1, Katherine A. Rawson2, Kalif E. Vaughn2, and John Dunlosky2  
1Department of Psychology, University of Richmond and 2Department of Psychology, Kent State University

Abstract
Taking tests on to-be-learned material is one of the most powerful learning strategies available to students. We examined the magnitude and mechanisms of the testing effect in college students with (n = 25) and without (n = 75) attention-deficit/hyperactivity disorder by comparing the effect of practice testing versus a comparable amount of restudy on long-term recall. Participants learned two lists of 48 words representing eight categories—one via eight consecutive study trials and another via four alternating study and test trials—and took recall tests 2 days later. Both groups demonstrated a moderate testing effect (d = 0.50, 0.57), and testing improved memory by enhancing both relational and item-specific processing. Results support the use of test-enhanced learning to promote the academic achievement of college students with attention-deficit/hyperactivity disorder and the inclusion of self-testing strategies into skills-based interventions for this population.

Keywords
attention-deficit/hyperactivity disorder (ADHD), memory, cognitive, processes

Received 3/24/14; Revision accepted 11/20/14

Taking tests on to-be-learned material is one of the most powerful learning strategies available to students (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). Numerous studies have supported the efficacy of retrieval practice in enhancing long-term retention above comparable amounts of restudy alone (Roediger & Butler, 2011). Although this testing effect is one of the most robust and well-documented effects in the literature on cognitive psychology, surprisingly few researchers have explored the benefits of testing over restudy in clinical populations that might benefit from test-enhanced learning (Sumowski, Chiavarotti, & DeLuca, 2010; Sumowski, Wood, et al., 2010). If students who struggle in traditional learning environments benefit from testing, this technique may provide a basis for interventions to improve their achievement.

Attention-deficit/hyperactivity disorder (ADHD) is a prevalent developmental disorder characterized by impairing inattention or hyperactivity/impulsivity that persists into adulthood in approximately two thirds of cases (Barkley, Murphy, & Fischer, 2008). ADHD is associated with pervasive academic impairments. College students with ADHD use effective study strategies less consistently, earn lower grades, and are less likely to graduate compared with their peers (see DuPaul, Weyandt, O'Dell, & Varejao, 2009, for a review). Despite such academic impairment, few studies have focused on interventions for college students with ADHD and, to our knowledge, no studies have investigated the effectiveness of specific learning strategies in this population.

Given the impact of ADHD on academic functioning, we evaluated whether college students with ADHD can benefit from testing as much as non-ADHD students. In a previous study in which participants could use strategies of their choosing to learn word pairs, adults with ADHD were less likely to self-test and had poorer recall for the to-be-remembered items than did control participants.

Corresponding Author:
Laura E. Knouse, Department of Psychology, University of Richmond, 28 Westhampton Way, Richmond, VA 23173  
E-mail: lknouse@richmond.edu
(Knouse, Anastopoulos, & Dunlosky, 2012). However, the extent to which testing benefits the learning of ADHD students relative to non-ADHD peers remains unknown. To investigate this issue, we used a method from Zaromb and Roediger (2010) in which participants attempted to learn lists of words from several categories (e.g., six fruits, six metals). We compared the within-subject effect of testing versus a comparable amount of restudy on the long-term recall for college students with and without ADHD.

This method also allowed us to evaluate the hypothesis that taking tests enhances distinctive processing. Distinctive processing, or encoding differences between items (item-specific processing) in the context of encoding similarities among them (relational processing), results in robust learning because relational processing defines the memory search set while item-specific processing helps one to identify targets within the set (Hunt, 2013). For example, participants can use relational processing by mentally organizing words into their appropriate categories. Item-specific processing would involve identifying unique aspects of the word referents within a category (e.g., noting that apples are crunchy whereas bananas are soft). Zaromb and Roediger (2010) found that as participants had more practice tests, both relational and item-specific processes increased, thereby suggesting that testing improves learning by increasing the likelihood of distinctive processing.

The hypothesis that testing enhances distinctive processing is intriguing, given findings on relational processing in people with ADHD. A few studies have investigated the performance of adults with ADHD on the California Verbal Learning Test (Delis, Kramer, Kaplan, Ober, & Fridlund, 1985), in which participants learn a list of words that can be categorized. Compared with healthy control participants, adults with ADHD typically remember fewer words at free recall and—most important for our present purposes—they are less likely to cluster items by category during recall, which suggests poorer relational processing (Holdnack, Moberg, Arnold, Gur, & Gur, 1995; Roth et al., 2004; Seidman, Biederman, Weber, Hatch, & Farah, 1998). If adults with ADHD engage in less relational processing during learning, testing may help to ameliorate this deficit by boosting their relational processing.

In the current study, we compared the recall performance of college students with and without ADHD for items that were restudied or tested multiple times. We chose items that could be grouped into categories so that we could examine the effect of testing on measures of relational and item-specific processes (described later). Young adults without ADHD were sampled from a larger study, which enabled us to match the ADHD and non-ADHD groups on recall performance in the restudy-only condition so that any interaction between group and study condition (restudy vs. testing) would be meaningful and not scale dependent (Loftus, 1978; Wagenmakers, Krypotos, Criss, & Iverson, 2012). On the basis of outcomes of prior research (Zaromb & Roediger, 2010), we predicted that college students without ADHD would benefit from testing, which would enhance relational and item-specific processing. Of greatest interest was the extent to which students with ADHD show a similar benefit of testing and whether this benefit reflects distinctive processing. Prior research on relational processing for adults with ADHD, as described earlier, has suggested that reduced relational processing may undermine the benefits of testing for students with ADHD. Alternatively, testing may be sufficiently robust to enhance their performance to the same extent as it does for students without ADHD. Resolving these issues is crucial to develop interventions for students with ADHD on the basis of test-enhanced learning.

Method

Participants

Participants were 100 undergraduates from the University of Richmond (n = 21) and Kent State University (n = 79). Twenty-five participants were included in the ADHD group, and the remaining 75 were non-ADHD-matched-comparison students. We oversampled non-ADHD students to increase power to detect effects. For each participant with ADHD, we selected 3 non-ADHD participants from the larger study with similar level of recall for restudied items. If more than three matches were available, to the extent possible, we selected students with similar age, vocabulary test score, and grade point average (GPA). Despite efforts to match on these secondary variables while preserving equivalence in study-only recall, participants with ADHD were on average 1 year older (ADHD group: M = 20, SE = 0.3; non-ADHD group: M = 19, SE = 0.2), t(94) = 2.46, p = .016, d = 0.57, had higher vocabulary scores (ADHD group: M = 32 out of 40, SE = 0.8; non-ADHD group: M = 29, SE = 0.4), t(98) = 3.75, p < .001, d = 0.87, and lower self-reported GPA (ADHD group: M = 3.10, SE = 0.11; non-ADHD group: M = 3.32, SE = 0.06), t(95) = 1.84, p = .069, d = 0.44. To foreshadow, we consider matching further in the Discussion section.

Inclusion criteria. Participants were included in the ADHD group if they met the following criteria: (a) reported diagnosis of ADHD by a health-care professional, (b) received ADHD assessment and diagnosis within the past 5 years or received accommodations in college under the Americans With Disabilities Act,
(c) scored above the clinical cutoff (93rd percentile using age-based norms) on either the inattentive or combined symptom subscales of the Barkley Adult ADHD Rating Scale (BAARS; Barkley, 2011), (d) endorsed onset of symptoms before age 12 on the BAARS, and (e) endorsed current impairment as a result of ADHD symptoms in at least two domains. Participants from both groups were excluded from analysis if they endorsed past or current schizotypal disorders, bipolar disorder, seizure disorder, or traumatic brain injury.

Recruitment and screening. At the University of Richmond, participants responded to flyers and e-mail advertisements by completing a brief online screening questionnaire. The screening included BAARS and self-report items to address the inclusion and exclusion criteria. The screening questionnaire also assessed current medication status and past or current diagnoses of disorders other than ADHD. Participants who screened into the group with ADHD were scheduled for two in-person study sessions. All staff members who scheduled and interacted with participants during the experimental sessions were blind to participants’ ADHD status. Participants received $20.00 for completing the sessions.

Of the 70 participants who completed the screening, 46 reported an ADHD diagnosis and 26 met inclusion criteria. The most common reasons for rule out (not mutually exclusive) were onset of symptoms after age 12 (n = 10), impairment in only one domain (n = 6), presence of rule-out condition (n = 5), and ADHD symptoms below threshold (n = 4). Four eligible participants did not respond to attempts to schedule sessions and 1 completed only the first session, which resulted in 21 participants with ADHD completing at Richmond.

Students in a larger study (n = 234) at Kent State completed the study sessions in exchange for credit in their psychology course. All participants completed questionnaires that included measures identical to the online screening instruments used at Richmond, thereby enabling identification of Kent participants that met criteria for inclusion in the ADHD group (n = 4). Participants selected from the larger study for the non-ADHD group (n = 75) included participants who did not have an ADHD diagnosis and who were matched to the ADHD participants as described earlier.

Sample characteristics. Our sample was 74% female and 26% male with a mean age of 19.54 years (SD = 1.89), and the majority self-identified as Caucasian (79%); we did not measure socioeconomic status. On the basis of self-reported diagnoses, our sample of ADHD students showed typically observed elevated risk for other psychiatric disorders (Kessler et al., 2006). The ADHD group was more likely to report current depression (24%) compared with the non-ADHD group (8%), \( \chi^2(1, N = 100) = 14.55, p = .001 \), and to report a generalized anxiety disorder (ADHD group: 12%; non-ADHD group: 4%), \( \chi^2(1, N = 100) = 12.88, p = .002 \). The ADHD group more frequently reported being diagnosed with learning disorders (ADHD group: 20%; non-ADHD group: 1.3%), \( \chi^2(1, N = 100) = 11.58, p = .001 \). Specifically, 3 participants reported a reading disorder, 1 reported dysgraphia, and 1 reported an auditory processing disorder; 1 non-ADHD participant also reported a reading disorder. Differences in self-reported rates of current posttraumatic stress disorder (ADHD group: 8%; non-ADHD group: 1.3%), \( \chi^2(1, N = 100) = 2.86, p = .091 \), and bulimia (ADHD group: 4%; non-ADHD group: 0%), \( \chi^2(1, N = 100) = 3.30, p = .082 \), did not reach significance, and rates of self-reported current social anxiety disorder (ADHD group: 0%; non-ADHD group: 1.3%), \( \chi^2(1, N = 100) = 0.34, p = .562 \), obsessive-compulsive disorder (ADHD group: 4%; non-ADHD group: 2.7%), \( \chi^2(1, N = 100) = 0.12, p = .735 \), and alcohol-use disorders (0% both groups) did not differ.

Measures

BAARS. The BAARS (Barkley, 2011) is a nationally normed rating scale that assesses participants’ self-reported Diagnostic and Statistical Manual of Mental Disorders (4th ed., DSM–IV; American Psychiatric Association, 1994) inattentive and hyperactive-impulsive ADHD symptoms during the past 6 months on a 4-point scale ranging from 0 (never or rarely) to 4 (very often). The scale includes items that assess other DSM–IV ADHD criteria, including age of onset and domains of impairment. Internal consistency for the full 18-item scale (Cronbach’s \( \alpha = .92 \)) and for the Inattentive subscale (Cronbach’s \( \alpha = .91 \)) in the sample was excellent.

Shipley Institute of Living Scale–Revised. Participants completed the vocabulary test from the Shipley Institute of Living Scale–Revised (Zachary, 1986), which includes 40 multiple-choice questions in which a word is presented and participants are to select the best synonym from among four choices. The Shipley vocabulary test demonstrated acceptable split-half reliability in this sample (.75 with Spearman-Brown correction).

Memory task

Design. We used a within-subjects design for two conditions that involved either eight consecutive study trials or four alternating study and test trials. Assignment of word list to condition and order of condition were counterbalanced across participants.
Study-Test completed the questionnaires described earlier. They had not studied. After the recognition test, participants responded by indicating which words they had (vs. studied words and lures in a random order one at a time recognition test for each list. They were presented in counterbalanced order. Next, participants completed a final tests. They were reminded of the color and position of the lists and asked to complete a free recall test for each.

Materials. A total of 192 words were sampled from 16 categories (12 words per category) from the norms by Van Overschelde, Rawson, and Dunlosky (2004). Six words in each category composed the studied lists, and the remaining 6 words in each category were lures for the final recognition test. Words were divided into two lists of 48 words with 8 categories in each list.

Procedure. All procedures were approved by institutional review boards at both institutions. Participants completed two sessions scheduled 2 days apart. During Session 1, they completed 16 blocks of trials on a computer (8 in each condition) that involved either study or free recall of the word lists as per the design described earlier. During study trials, words were presented in random order (for 3.25 s each) with a 500-ms interstimulus interval. Each list was presented for study trials either in blue font on the left side of the screen or in red font on the right side of the screen (to aid participants in differentiating the lists during final free recall in Session 2), and these assignments were counterbalanced across participants. During test trials, participants were given 3 min to recall and type as many words as they could. The first session lasted approximately 1 hr.

Two days later, participants returned to complete the final tests. They were reminded of the color and position of the lists and asked to complete a free recall test for each in counterbalanced order. Next, participants completed a recognition test for each list. They were presented with studied words and lures in a random order one at a time and responded by indicating which words they had (vs. had not) studied. After the recognition test, participants completed the questionnaires described earlier.

Results

Analysis plan

After matching groups on study-only recall (as described earlier), we examined recall by group during the learning trials in the study-test condition. Next, we conducted 2 × 2 mixed-factor analyses of variance (ANOVA) to examine the effect of condition (restudy only vs. study-test) and group (ADHD vs. non-ADHD) on final recall and recognition performance and on measures of relational and distinctive processing.

Relational processing was measured using two metrics: the extent to which participants clustered items into categories during free recall (adjusted ratio of clustering, ARC; Roenker, Thompson, & Brown, 1971) and category access (computed as the percentage of categories for which at least one item was recalled).3 ARC scores range from –1 to 1; a score of 0 represents chance levels of clustering. Item-specific processing was measured using two metrics: the mean number of correct items recalled per category accessed and recognition performance (d′).

Free recall during learning

As expected, recall improved across practice test trials for both groups (ADHD group: 55, 54, 59, and 60%, SE = 3.1–4.8; non-ADHD group: 29, 49, 57, and 65%, SE = 1.6–2.2). A 4 (Test Trial) × 2 (Group: ADHD, non-ADHD) ANOVA showed a main effect of trial, F(3, 294) = 102.07, p < .001, η² = .51. The main effect for group was not significant, F(1, 98) = 0.34, p = .56, η² = .003. The interaction of test trial and group was significant, F(3, 294) = 3.58, p = .014, η² = .04. Participants with ADHD had slightly higher recall than did participants in the non-ADHD group on the first test trial, (98) = 1.85, p < .068, Cohen’s d = .43; performance on the remaining trials did not differ significantly (ts < 1.35, ds = 0.31, 0.09, –0.26). One explanation for this result is that the novelty of the task motivated participants in the ADHD group relative to the non-ADHD group during the initial trial but that this effect dissipated over time.

Free recall at final test

Of greatest interest, both the ADHD and the non-ADHD groups demonstrated the testing effect for free recall after a 2-day retention interval (see Fig. 1 for results). A 2 (Condition: Restudy Only, Study-Test) × 2 (Group: ADHD, Non-ADHD) ANOVA showed a main effect of encoding condition with better final free recall in the study-test condition than in the restudy-only condition, F(1, 98) = 21.42, p < .001, η² = .179. The main effect of group was
not significant, $F(1, 98) = 0.02, p = .89, 
\eta^2_p = .000$, nor was the interaction, $F(1, 98) = 0.23, p = .63, 
\eta^2_p = .002$. The testing effect was similar in magnitude for the ADHD ($d = 0.57$) and non-ADHD ($d = 0.50$) groups.

### Relational and item-specific processing

Relational processing as measured by ARC did not explain the benefits of testing on recall (see Table 1). A 2 (Condition: Restudy Only, Study-Test) × 2 (Group: ADHD, Non-ADHD) ANOVA yielded no main effects and no interaction ($F$s < 1, all $\eta^2_p$s < .01). Relational processing as measured by category access showed a main effect for encoding condition; more categories were accessed in the study-test condition than in the restudy-only condition, $F(1, 98) = 28.97, p < .001, 
\eta^2_p = .228$. The main effect of group and the interaction were not significant ($F$s < 1, $\eta^2_p$s < .01), thereby indicating that testing had similar effects on category access in both groups.

We examined item-specific processing as measured by items recalled per category accessed (see Table 1). A main effect of encoding condition indicated that students recalled more items per category in the study-test condition than in the restudy-only condition, $F(1, 98) = 19.91, p < .001, 
\eta^2_p = .170$. Neither the main effect of group nor the interaction were significant ($F < 1$, $\eta^2_p$s < .01, and $F = 2.28, 
\eta^2_p = .023$, respectively). Similarly, for recognition ($d$), the effect of encoding condition was significant, $F(1, 98) = 55.17, p < .001, 
\eta^2_p = .36$; the effect of group and the interaction were not ($F < 1$, $\eta^2_p = .004$, and $F < 1, 
\eta^2_p = .008$, respectively). Thus, both measures indicated that practice testing improved item-specific processing and that this effect was similar in magnitude regardless of ADHD status.

### Discussion

Young adult students with ADHD demonstrated test-enhanced learning as evidenced by superior long-term recall after study-test trials compared with study trials alone. Students with ADHD showed a testing benefit similar in magnitude to their non-ADHD peers. Results also indicate that the mechanisms underlying these benefits include improvements in relational and item-specific processing for both groups (see also Lipowski, Pyc, Dunlosky, & Rawson, 2013; Zaromb & Roediger, 2010). Although testing was not associated with increases in category clustering during recall, it was associated with greater category access (relational processing) and with more items recalled per category accessed and improved recognition (item-specific processing). We did not observe ADHD-related deficits in semantic clustering that have sometimes been reported in studies using the California Verbal Learning Test. Close inspection of those studies revealed that differences in semantic clustering have mostly arisen on measures that do not correct for number of words recalled. On corrected measures, between-groups differences are inconsistent (Holdnack et al., 1995) or unreliable (Roth et al., 2004). Thus, our results on the ARC, which corrects for words recalled, are not unexpected, given the prior literature. Most important, this study demonstrated that testing improved relational and item-specific processing for learners regardless of their ADHD status.

### Table 1. Measures of Distinctive Processing by Group and Encoding Condition

<table>
<thead>
<tr>
<th>Measure</th>
<th>Group</th>
<th>ADHD</th>
<th>Non-ADHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relational processing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restudy only</td>
<td>0.65 (0.06)</td>
<td>0.72 (0.04)</td>
<td></td>
</tr>
<tr>
<td>Study-test</td>
<td>0.72 (0.04)</td>
<td>0.72 (0.03)</td>
<td></td>
</tr>
<tr>
<td>Category access</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restudy only</td>
<td>4.84 (0.49)</td>
<td>4.77 (0.26)</td>
<td></td>
</tr>
<tr>
<td>Study-test</td>
<td>6.28 (0.27)</td>
<td>6.25 (0.17)</td>
<td></td>
</tr>
<tr>
<td>Item-specific processing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPCA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restudy only</td>
<td>2.88 (0.32)</td>
<td>3.03 (0.15)</td>
<td></td>
</tr>
<tr>
<td>Study-test</td>
<td>3.61 (0.28)</td>
<td>3.45 (0.13)</td>
<td></td>
</tr>
<tr>
<td>Recognition ($d$')</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restudy only</td>
<td>2.06 (0.22)</td>
<td>2.28 (0.12)</td>
<td></td>
</tr>
<tr>
<td>Study-test</td>
<td>2.88 (0.17)</td>
<td>2.92 (0.10)</td>
<td></td>
</tr>
<tr>
<td>Hit (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restudy only</td>
<td>83.58 (4.07)</td>
<td>83.00 (2.06)</td>
<td></td>
</tr>
<tr>
<td>Study-test</td>
<td>90.08 (2.25)</td>
<td>88.97 (1.78)</td>
<td></td>
</tr>
<tr>
<td>False alarm (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restudy only</td>
<td>24.08 (3.79)</td>
<td>16.11 (1.50)</td>
<td></td>
</tr>
<tr>
<td>Study-test</td>
<td>10.17 (2.07)</td>
<td>8.33 (0.89)</td>
<td></td>
</tr>
</tbody>
</table>

Note: The table presents means for each measure. Standard errors of the mean are shown in parentheses. ADHD = attention-deficit/hyperactivity disorder; ARC = adjusted ratio of clustering; IPCA = items per category accessed.
Roediger, 2009), and use of this strategy appears to be particularly effective. Many students do not use this strategy consistently under which adults with ADHD will use testing effects for educationally relevant materials and outcomes (e.g., Butler, Carpenter, Pashler, & Cepeda, 2009; Rawson, Dunlosky, & Sciarretti, 2013). For example, Butler (2010) presented undergraduates with lengthy expository texts for initial study followed by either restudy or practice short-answer tests. Practice testing dramatically improved performance on a comprehensive test administered 1 week later. Outcomes such as these support optimism that similar extensions would be observed for students with ADHD, but this remains an open question, given that complex materials may place a relatively heavier demand on the cognitive capacity of students with ADHD. It is fortunate that these previous studies provided methods that can be used to explore the extent to which learners with ADHD benefit from testing with more complex materials (cf. our extension of Zaromb & Roediger’s, 2010, method in the current study).

Another crucial step in translating testing effects from the laboratory to daily life is to establish the conditions under which adults with ADHD will use self-testing effectively. Many students do not use this strategy consistently when studying on their own (Karpicke, Butler, & Roediger, 2009), and use of this strategy appears to be even less common for students with ADHD (Knouse et al., 2012; Reaser, Prevatt, Petscher, & Proctor, 2007). ADHD-related memory deficits can often be attributed to self-regulation failures—in other words, difficulty translating intention into action in the contexts in which strategies are needed. We expect that task affordance, or the extent to which task parameters support strategy implementation, will have a powerful effect on use of self-testing for students with ADHD, as has been demonstrated in older adults (Bottirol, Dunlosky, Guerini, Cavallini, & Hertzog, 2010). For this reason, in future studies, researchers must empirically examine the level and type of task affordance that students with ADHD need to use and benefit from self-testing in their daily lives.

Moreover, researchers could examine whether and when strategy implementation differences emerge between learners with and without ADHD by testing variables related to structural, temporal, and motivational aspects of a task. For example, Rawson et al. (2015) examined the frequency and effectiveness with which undergraduates in a psychology course used spaced practice testing. They found that supervised in-laboratory practice sessions and unsupervised sessions that students completed on their own yielded comparable benefits in terms of exam performance and long-term retention. However, results may differ for students with ADHD, and studies using similar methods could be used to investigate how students with ADHD engage in self-testing and to identify the most important elements of a testing-based intervention tailored to the needs of this population.

Although further intervention research will be valuable for identifying conditions under which students with ADHD can best capitalize on test-enhanced learning, this does not preclude recommending that students with ADHD use test-enhanced learning strategies. Coaching in self-testing strategies can be provided through on-campus disability services offices and academic skills centers. Clinicians providing skills-based cognitive-behavioral therapy for ADHD (e.g., Safren et al., 2010) may find it worthwhile to add self-testing skills to manualized approaches. It may be particularly important to introduce students to user-friendly tools that support self-testing, such as “low-tech” paper flash cards, Web-based applications, and apps for mobile devices.

To summarize, the current study is the first to focus on the magnitude and mechanisms of the testing effect in people with ADHD. Our results replicate prior investigations on the role of relational and item-specific processing in the testing effect and provide a novel and important extension that supports the use of test-enhanced learning to improve the academic achievement of college students with ADHD. Future research will be needed to replicate our findings, to extend them to other subgroups of people with ADHD, and to determine the most effective...
ways to support self-testing in college students with ADHD.

Author Contributions
All authors contributed to the study concept, study design, and data collection. L. E. Knouse designed the screening protocol. K. A. Rawson and K. E. Vaughn designed the computer program and analyzed the data. L. E. Knouse and J. Dunlosky drafted the manuscript, and all authors provided critical revisions. All authors approved the final version of the manuscript for submission.

Acknowledgments
The authors wish to thank Emily J. Blevins for her assistance with data collection at the University of Richmond. Kalil E. Vaughn is now in the Department of Psychology at Williams College.

Declaration of Conflicting Interests
The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding
The research reported here was supported by a Collaborative Award from the James S. McDonnell Foundation 21st Century Science Initiative in Bridging Brain, Mind and Behavior. L. E. Knouse was supported by a University of Richmond Faculty Summer Research Grant during the completion of this study.

Notes
1. A few other studies with clinical samples involved practice testing but did not include a restudy comparison group (e.g., Balota, Duchek, Sergent-Marshall, & Roediger, 2006; Serrano, Latorre, Gatz, & Montanes, 2004).
2. A power analysis using G*Power 3.1.5 (Faul, Erdfelder, Buchner, & Lang, 2009) with power of .80 (α = .05) and assuming a repeated measures correlation of .50 indicated that 100 participants would provide sufficient sensitivity to detect an interaction (f) as small as 0.14.
3. For computation of both ARC and category access, responses were included in the analysis regardless of whether the item was correct or a categorical commission to minimize the extent to which these measures reflected item-specific memory (given that discrimination of target items from nontargets is largely based on item-specific information).

References


