Paper Title One for the Ages: The Pretesting Effect Is Robust Throughout Adulthood but Metacognitive Beliefs Differ

Author(s) Stephany Duany Rea, University of Texas at Austin; Veronica X. Yan, University of Texas at Austin; Faria Sana, Athabasca University

Session Title Cognitive and Metacognitive Processes to Improve Learning

Session Type Paper

Presentation Date 4/12/2024

Presentation Location Philadelphia, Pennsylvania

Descriptors Cognitive Processes/Development, Memory

Methodology Quantitative

Unit Division C - Learning and Instruction

DOI https://doi.org/10.3102/2101515

Theoretical Framework and Purpose

The Pretesting Effect

The pretesting effect refers to a counterintuitive finding that taking a test *before* learning new information in a subsequent lesson results in better memory recall for that information. The pretesting effect has been widely explored in the cognitive literature, and the evidence for its benefit is robust, even when participants generate incorrect guesses during the pretest phase (Grimaldi & Karpicke, 2012; Metcalfe, 2017; Yan et al., 2014). The practical implications of pretesting are expansive: it not only has implications for how new ideas and concepts might be taught in schools (see review by Carpenter et al., 2023), but it also has implications for life beyond formal education (e.g., learning on the job, managing financial and health information).

There are several theoretical accounts to explain the memory benefits of pretesting. One account emphasizes that pretests enhance attentional processing of pretested information by directing attention to key, and potentially important or testable information making it more noticeable when it does appear in the subsequent lesson (Carpenter et al., 2023). There is, however, an even greater learning benefit when learners attempt to respond to the pretest questions, even though they score below chance on it, compared to when they simply study the pretest questions with the correct responses highlighted (Sana et al., 2021). This suggests that the act of trying to generate answers, rather than simply reading questions or being directed to answers, plays a crucial role in the pretesting benefit. To explain these findings, theories relating to semantic activation in memory have been proposed, suggesting that presenting learners with pretest questions or prompts related to upcoming material allows them to retrieve related information or prior knowledge from memory, leading to richer and more elaborative encoding of to-be-learned information (Carpenter, 2011; Pyc & Rawson, 2010).

Aging and Memory

Previous studies on the pretesting effect have primarily focused on younger adults (i.e., undergraduate students). A separate literature however has focused on the relative benefits of errorless versus trial-and-error learning, which demonstrates that guessing (analogous to pretesting) when learning related word pairs is beneficial for both younger and older adults (Cyr & Anderson, 2012). Despite this finding, social comparisons and self-efficacy about memory can play a role in individuals' approaches to their own behaviors. For example, O'Brien and Hummert (2006) found that older adults performed better on a word recall test when exposed to a higher threat of comparison with younger adults, indicating that older adults perceive social comparison about memory performance as a concern, and Lineweaver et al. (2023) reported that older adults perceived cognitive memory strategies as more difficult when their self-efficacy beliefs about their memory were lower. In summary, it is not simply memory performance, but the intrapersonal beliefs about it as well that have implications for behavior engagement and subsequent performance. Thus, exploring the effect of a memory performance strategy, namely pretesting, in a group of individuals who may be at a higher risk of experiencing perceived aging threats can be important for theoretical frameworks around the strategy and for practical day-to-day tasks. To this end, the purpose of the current study is to examine the extent to which the pretesting effect is observed across younger and older adults, and whether individual differences in age, memory contentment, and strategy effectiveness judgments moderate the effect.

Research Questions

RQ1. Is the pretesting benefit moderated by age or memory contentment?
RQ2. How do participants judge the effectiveness of pretesting as a learning strategy?

Methods

Participants and Design

Three hundred participants were recruited via Prolific.co, with fifty participants being recruited in each of the following age groups: 18-29 years, 30-39 years, 40-49 years, 50-59 years, 60-69 years, 70+ years. We stratified the sampling by age in order to ensure an even distribution of ages (see Table S1 for descriptives). Participants could be recruited from any country as long as they were fluent English speakers. After removing participants who did not complete the main experiment or who did not fill in their age, we were left with data from 273 participants. The experiment employed a within-participants design. All participants studied 32-word pairs. For each participant, 16-word pairs were randomly assigned to the pretest condition, and 16 were randomly assigned to the study-only condition.

Materials

Word Pair Stimuli

The stimuli were word pairs and they consisted of a cue word and a target word (e.g. Whale: Mammal; Train: Caboose; based on 60 low-associate word pairs used in Kornell et al., 2009). The target words were selected to be weak associates of the cue word, with a forward association strength between 0.05 and 0.054. Consistent with this norming information, in the present study, we found that participants correctly guessed the target word only 5% of the time (SD = 5%). For each participant, a random set of 32 pairs were drawn from the larger pool of 60 words.

Multifactorial Memory Questionnaire: Contentment subscale

The multifactorial memory questionnaire (MMQ) was used to assess subjective reports about memory (scale developed for older adults by Troyer & Rich, 2002). We focused on one factor of this scale, contentment with one's memory (people's contentment or anxiety over their memory; "I have confidence in my ability to remember things", on a scale of 1 to 6 (strongly disagree/agree), $\alpha = .93$).

Procedure

Participants were directed to a Qualtrics survey to provide consent, then they were then given a link to the main experiment (hosted on Collector (https://github.com/gikeymarcia/Collector) an open-source PHP-based program designed to run psychology experiments). Once they completed the main experiment, they were given a verification code to enter back into Qualtrics to complete the Qualtrics survey (post-test judgments, memory contentment items, and demographics).

For the pre-testing phase, participants were told that they would see pairs of related words and that for some of the pairs they would be asked to first guess what the second word is before being shown the correct answer and that for the other pairs, they would simply see the pair without having to make a guess. They were also told that on the final test, their task would be to type in the correct answer, not their guess. Once participants finished reading the instructions, they clicked a button to begin and were then shown one-word pair at a time. They saw a total of 32 pairs. The order of the pairs was fully randomized for each participant. For the pairs randomized to the pretest condition, the cue was shown with a blank box presented for them to type in their response. This remained on screen for 8-sec after which the correct word pair response was presented in the center of the screen for 5-sec. For the pairs randomized to the study-only condition, the cue-target pair was presented in the center of the screen for 13 seconds. The total time on task was identical for both conditions (13-sec), but the duration that the correct answer appeared on the screen in the pretest condition was much shorter than in the study condition. See Figure S1 for an illustration of this Phase. After the presentation of the word pairs, participants were asked to make judgments of learning (i.e., "How likely are you to correctly recall this item on a later test? 0-100.").

After all 32 pairs had been presented, participants engaged in a 60-sec distractor task, followed by the self-paced test phase. They were presented with the 32 cues, one at a time, and instructed to recall the corresponding correct responses (not their guesses) that were presented in the pretest/study phase. Participants typed each response in a blank box next to the cue. Participants were then provided with a link to return to the Qualtrics survey to answer the final study questions.

Participants were reminded that for half of the word pairs, they simply studied the full pair and that for the other half, they were first asked to guess what the answer was before being shown the correct response. They were then asked to select from one of three options, which they thought led to better learning of the word pairs: studying the full pair led to better learning, making a guess first led to better learning, or it did not matter. Next, they rated their agreement on fourteen memory contentment scale items. Finally, they answered demographic questions about gender, age, race/ethnicity, and highest level of education.

Results

RQ1. Is The Pretesting Benefit Moderated By Age Or Memory Contentment?

We replicated the pretesting benefit. Final test performance was significantly higher for the word pairs in the pretest condition (M = .71, SD = .20) than those in the study condition (M = .64, SD = .22), $t(540.2) = 4.09$, $p < .001$, Cohen's $d = 0.35$. To analyze whether the pretesting benefit was moderated by our individual difference measures, we conducted nested linear mixed effects analyses in which both participant and item IDs were entered as random effects. Specifically, we conducted two analyses. In Model 1, we tested whether the accuracy of target word recall was predicted by condition (study vs. pretest), gender (male = -1, female = +1), age (standardized), and memory contentment score (standardized). In Model 2, we added two-way interactions between conditions and the individual difference (age, memory contentment). Model 2 was not a significantly better fit than Model 1, $\chi^2(2) = 1.44$, $p = .488$. The results of the analyses are detailed in Table S2. Model 1 reveals that the pretest condition led to higher accuracy than the study condition. Moreover, in general, women scored higher than men and those who were more content with their memories also scored higher than those who were less content with their memories. These main effects were also found in Model 2. The interaction terms from Model 2, however, are the most central to our research questions and neither were statistically significant. These results show that the pretesting effect was not moderated by age or by memory contentment.

RQ2. How do Participants Judge the Effectiveness of Pretesting as a Learning Strategy?

To examine whether the judgments varied by age, we categorized age into six age groups: 18-29, 30-39, 40-49, 50-59, 60-69, and 70+ years. A chi-square test-of-independence revealed that age and judgment were not independent, $\chi^2(10) = 24.69$, $p = .006$. Participants in their 30s or younger were most likely to think that guessing first was better for learning, participants in their 40s and 50s were most likely to think that studying only was better for learning, and participants over the age of 60 were relatively evenly split across the three options (see Figure S2). Finally, we categorized memory contentment into quartiles in order to analyze how the judgments varied by contentment. Results showed that memory contentment and judgments were independent ($\chi^2(10) = 24.69$, $p = .006$).

Conclusions and Significance

We replicated the pretesting effect across the span of adulthood (Cyr & Anderson, 2012). Moreover, the results demonstrate that this pretesting benefit is not significantly moderated by age or memory contentment. While much of the literature on pretesting has focused on younger adults (e.g., college-aged students), long-term recall is a beneficial skill to have or maintain throughout life (e.g., remembering medication information, daily tasks, school, etc.). It seems that the older adults who may be more susceptible to intrapersonal beliefs about memory performance also seem to benefit from pretesting given that we did not find memory contentment as a moderating factor of the pretesting effect. Additionally, the way in which different age groups perceive the effectiveness of studying strategies differs, with a mix of younger participants recognizing the value of guessing and mid-older adults finding studying-only as the best approach. The current study is one of the first to examine the boundary conditions of the pretesting effect demonstrating that the effect generalizes across adulthood and does not differ in age-related memory contentment, but there are differences in its perceived effectiveness that can impact whether individuals use the strategy or not.

Word count: 1983

References

Carpenter, S. K., King-Shepard, Q., & Nokes-Malach, T. (2023). The prequestion effect: Why it is useful to ask students questions before they learn. *Copyright and Other Legal Notices*, 74.

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1547. https://doi.org/10.1037/a0024140

Cyr, A. A., & Anderson, N. D. (2012). Trial-and-error learning improves source memory among young and older adults. *Psychology and aging, 27*(2), 429. https://doi.org/10.1037/a0025115

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding?. *Memory & cognition, 40*, 505-513. https://doi.org/10.3758/s13421-011-0174-0

Lineweaver, T. T., Fansler, S. D., Horhota, M., Crumley-Branyon, J. J., & Wright, M. K. (2023). Older adults' memory beliefs predict perceptions of memory strategy difficulty and effectiveness. *Aging, Neuropsychology, and Cognition, 30*(1), 66-77. https://doi.org/10.1080/13825585.2021.1962794

Metcalfe, J. (2017). Learning from errors. Annual review of psychology, 68, 465-489.

O'Brien, L. T., & Hummert, M. L. (2006). Memory performance of late middle–aged adults: Contrasting self–stereotyping and stereotype threat accounts of assimilation to age stereotypes. *Social Cognition, 24*(3), 38-358.https://doi.org/10.1521/soco.2006.24.3.338

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science, 330*(6002), 335-335. https://doi.org/10.1126/science.1191465

Sana, F., Yan, V. X., Clark, C. M., Bjork, E. L., & Bjork, R. A. (2021). Improving conceptual learning via pretests. *Journal of Experimental Psychology: Applied, 27*(2), 228–236. https://doi.org/10.1037/xap0000322

Troyer, A. K., & Rich, J. B. (2002). Psychometric properties of a new metamemory questionnaire for older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 57*(1), P19-P27. https://doi.org/10.1093/geronb/57.1.P19

Yan, V. X., Yu, Y., Garcia, M. A., & Bjork, R. A. (2014). Why does guessing incorrectly enhance, rather than impair, retention?. *Memory & Cognition, 42*(8), 1373-1383. https://doi.org/10.3758/s13421-014-0454-6

Supplemental Materials

Table S1

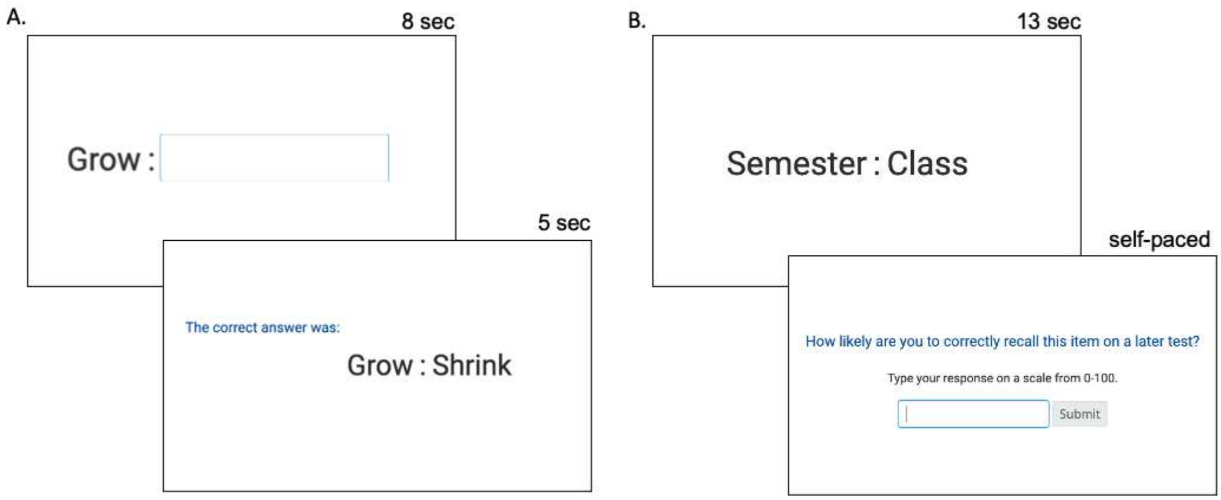
Participants' Descriptives

Demographic Variable	**Variable Subcategories**	
Age Mean (*SD)*	Range 18-82	48.26 (17.08)
Gender	Female	146
	Male	125
	Non-Binary	2
Race/Ethnicity	White	88%
	Asian	3%
	Black or African American	1%
	Hispanic, Latino, or Spanish Origin	6%
	Middle Eastern or North African	1%
	Other	2%

Degree	**20s**	**30s**	**40s**	**50s**	**60s**	**70s+**	**Total**
Less than HS diploma	2 (4%)	2 (4%)	0 (0%)	2 (4%)	0 (0%)	0 (0%)	6 (2%)
High school diploma (or GED)	5 (11%)	7 (15%)	4 (9%)	9 (18%)	5 (12%)	10 (24%)	40 (15%)
Some college, no degree	9 (19%)	8 (17%)	8 (18%)	7 (14%)	3 (7%)	8 (19%)	43 (16%)
Associate's degree	4 (9%)	1 (2%)	1 (2%)	1 (2%)	3 (7%)	1 (2%)	11 (4%)
Vocational/technical degree	3 (6%)	2 (4%)	4 (9%)	6 (12%)	7 (16%)	3 (7%)	25 (9%)
Bachelor's degree	15 (31%)	14 (30%)	13 (29%)	11 (22%)	15 (35%)	11 (26%)	79 (29%)
Master's degree	9 (19%)	12 (26%)	11 (24%)	11 (22%)	6 (14%)	7 (17%)	56 (21%)
Doctoral Degree (PhD)	0 (0%)	0 (0%)	2 (4%)	1 (2%)	1 (2%)	1 (2%)	5 (2%)
Professional Degree (M.D., D.D.S., J.D., etc.)	0 (0%)	1 (2%)	2 (4%)	1 (2%)	3 (7%)	1 (2%)	8 (3%)

Figure S1

Illustration of Study Phase Trials



Note. The left half (A) depicts a pretest condition trial. The right half (B) depicts a study condition trial followed by an item-based judgment of learning.

Table S2

Regression Summary Statistics: Predicting Test Accuracy

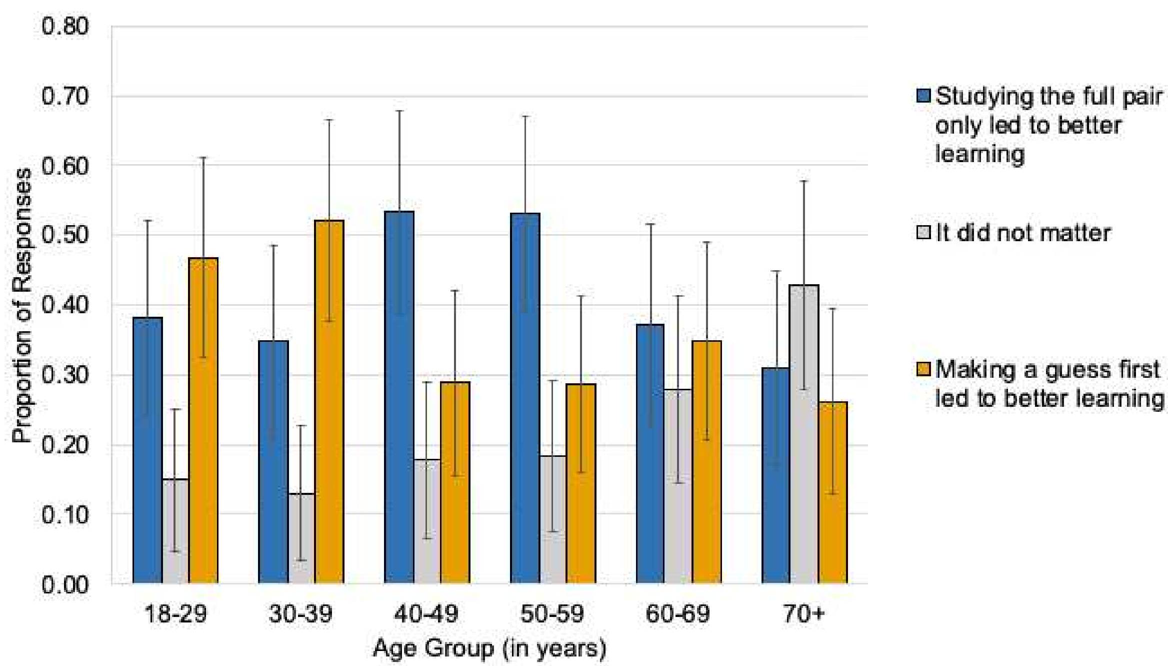
	Model 1				Model 2			
	Est.	*SE*	*t(df)*	*p*	*Est.*	*SE*	*t(df)*	*p*
Intercept	.637	.019	33.47 (128.2)	< .001	.637	.019	33.47 (128.0)	< .001
Condition	.070	.009	7.80 (8343)	< .001	.073	.009	7.81 (8341)	< .001
Age	-.010	.011	-0.75 (266.8)	.452	-.004	.012	-0.31 (360.3)	.759
Gender	.047	.011	4.16 (266.8)	< .001	.047	.011	4.16 (266.8)	< .001
Memory Contentment	.038	.011	3.45 (266.7)	< .001	.040	.012	3.29 (361.4)	.001

Condition × Age	-.009	.009	-1.05 (8341)	.295
Condition × Memory Contentment	-.005	.009	-0.50 (8341)	.614

Note. The reference level for condition is the study condition. Age and memory contentment score were both standardized. Gender was dummy-coded such that -1 = male, and +1 = female. Participant ID and item ID were entered as fully crossed random effects.

Figure S2

Metacognitive Judgments of Strategy Effectiveness By Age Group



Note. Error bars represent 95% confidence interval around the proportion, calculated as 1.96 × $\sqrt{(\rho(1-\rho)/n)}$.