The Testing Effect Does Not Uniformly Affect Source Memory

Gene A. Brewer, M.S. Arlo Clark-Foos, M.S. Richard L. Marsh, Ph.D. Jason Hicks*, Ph.D. Joseph T. Meeks, M.S.

Department of Psychology The University of Georgia, Athens, GA

Department of Psychology Louisiana State University*, Baton Rouge, LA

Introduction

Broadly construed, the testing effect is the empirical finding that information that has been tested earlier (usually by free recall) leads to more durable memory during a later testing phase (e.g., Roediger & Karpicke, 2006).

Recently, Chan and McDermott (2007) proposed that the testing effect increases aspects of recollective experience, namely, they found an increase in list discrimination performance and an increase in remember as compared with know responses on the final test.

Chan and McDermott used a two list paradigm in which each list was either tested or not tested (see Figure 1).

According to the source-monitoring framework, temporal information and recollective aspects of earlier experiences represent qualitative details that could be stored in memory during the initial test phase. The goal of our work was to ascertain whether the testing effect would extend to other, less salient characteristics that were present during the original encoding.

On the one hand, temporal information from the recall test may be imbued in items as they are recalled by participants. Other source characteristics that are less relevant to free recall may not, and they may not be restrengthened as well. If so, the testing effect should extend to non-temporal characteristics as well.

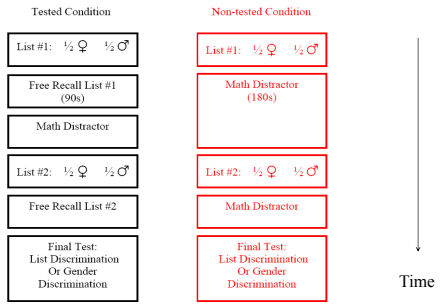
To test these competing hypotheses, we directly replicated Chan and McDermott's paradigm with two critical changes.

First, items during original learning of two lists were spoken by either a male or female speaker (i.e., had gender information).

Second, we sought to replicate the list discrimination (temporal) advantage in one condition, but critically, tested gender information in another condition.

If the testing effect extends to all manner of qualitative characteristics, then we should obtain comparable increases in both list and gender discrimination judgments subsequent to testing as compared with not having an earlier free recall task.

Figure 1. General Procedure



Experiment 1, 2, & 3 Procedures

Experiment 1

- 4 orthogonally crossed conditions:
  - 2 received free recall testing whereas 2 did not
  - 2 were tested on list membership whereas 2 were queried on gender
- All groups studied 30 items in each list (half from each gender)
- Timing was equated in all four groups using distractor tasks
- The final list or gender discrimination test contained all 60 studied items

Experiment 2

- 4 orthogonally crossed conditions:
  - 2 received free recall testing whereas 2 did not
  - 2 were tested on list membership whereas 2 were queried on gender
- Both free recall condition specified gender after each recalled item

Experiment 3

- 2 conditions tested:
  - Immediate free recall
  - Immediate free recall also specifying within-list location
- Only list discriminations were tested on the final test

Experiment 1, 2, & 3 Results

Experiment 1

List discrimination judgments wholly replicated Chan and McDermott's (2007) finding that free recall testing after each list improved performance on the final test as compared with not being tested after each list.

By contrast, there was no difference in gender discrimination in the recall versus the no-recall conditions despite performance being statistically above chance.

One hypothesis is that free recall testing imbues items with extra temporal details at the time of recall that subsequently aids list discrimination judgments on the final test. By such an account, the gender of the speaker is not necessary during free recall and may not be accessed or strengthened; and consequently, memory for this information is not affected by immediate testing.

Experiment 2

Gender discrimination is improved following immediate free recall testing that involves a strengthening of the gender characteristic.

Thus, a testing effect can be obtained on a non-temporal dimension as long as the immediate test is designed to strengthen or reinforce that attribute. By contrast, this gender-recall manipulation eliminated the testing advantage for list membership observed by Chan and McDermott and that we obtained in Experiment 1.

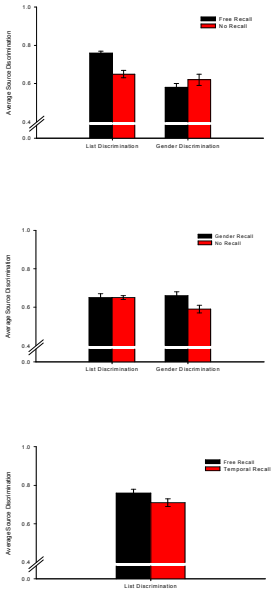
We believe that focusing people on gender information during the immediate recall test reduced the encoding of temporal characteristics that otherwise would have been stored during the test.

Experiment 3

Requiring within-list temporal judgments on the immediate recall test attenuated the testing effect on final list discrimination judgments.

Nevertheless, standard free recall resulted in better list judgments than did the within-list manipulation. Of course, the magnitude of these effects is being judged relative to the two free recall results observed in Experiment 1, and we realize that this is a cross-experimental comparison.

Figure 2. Discrimination Performance across three experiments



Discussion

Together, these three experiments converge on the notion that the testing effect is not a monolithic or a unitary influence on subsequent memory performance. Researchers may believe that testing ultimately strengthens recollective details, or qualitative characteristics as specified in the source-monitoring framework. We do not disagree with this general line of reasoning, but we do believe that the exact memorial details affected by testing will be a function of the nature of the free recall test *and* what dimension is ultimately tested on a later occasion. Strengthening gender information or increasing within-list positions tended to reduce between-list discriminations. In this regard, the results from this study both qualify the effects of immediate testing, and also draw a strong parallel to the transfer-appropriate processing literature. Concerning the latter, the most readily available information on a given memory test will be a consequence of the manner in which it was studied or reinforced prior to that test. Obviously, other dimensions and procedural variations on our experiments need to be tested, but we believe that these results can be used profitably to advance both theoretical and applied educational principles.

Acknowledgments

We would like to thank Hollis Johnson and Jason Kinney for their dedicated assistance in collecting the data for the three experiments presented in this poster.