

Appendix A

Example of a completed QALMRI summary

<p>Ecological validity of the testing effect: The use of daily quizzes in introductory Psychology (Batsell, Perry, Hanley, and Hostetter, 2017)</p> <p>Full article: http://journals.sagepub.com/doi/full/10.1177/0098628316677492</p>	
<p>The “testing effect” refers to the finding that people better remember information after being tested on that material, as compared to only studying. Testing has also been shown to enhance memory for study material that did not appear on the initial test (referred to as retrieval-induced facilitation), demonstrating more general benefits of testing as a way to enhance memory and learning. Although there is much evidence to demonstrate the testing effect in the laboratory, there is little evidence of the testing effect in real-life, classroom settings. The current study examines the testing effect in a naturalistic setting by comparing memory performance for Introductory Psychology material in one class that studied the material and received brief daily quizzes to another that studied the material but did not receive daily quizzes.</p>	
<p>Question:</p> <p>What was the broad question being asked by this research project?</p> <p>What was the specific question being asked by this research project?</p>	<p>What kinds of teaching techniques will improve student learning in a classroom?</p> <p>Will daily quizzes enhance retention of assigned study material?</p>
<p>Alternatives:</p> <p>What was the author’s hypothesis?</p> <p>What were the alternative hypotheses?</p>	<p>H1: Daily quizzes will improve memory for study material whether the material had appeared in a quiz or not.</p> <p>H2: Daily quizzes will only improve memory for study material that appeared in the quizzes.</p> <p>H3: Daily quizzes will not improve memory for study material</p>
<p>Logic:</p> <p>What was the logic of the hypotheses?</p> <p>If the author’s hypothesis is true, what should happen?</p>	<p><i>If H1, then a quizzed group will perform better than a study-only group on a memory test for all the studied material.</i></p> <p><i>If H2, then a quizzed group will perform better than a study-only group on a memory test, but only for the material that appeared on the tests.</i></p>

	<p><i>If H3, then a quizzed group will not perform better than a study-only group on a memory test for the studied material.</i></p>
<p>Methods:</p> <p>What were the methods?</p>	<p>General overview: The study used a 2x3 factorial design (quasi-experimental). Students enrolled in two Intro Psychology courses took part in the study. The participants were all assigned textbook readings which consisted of material not taught during the lectures. One class received daily quizzes (21 total), while the other did not. Memory was tested three times throughout the term using 15 multiple-choice questions. Test questions were either identical to the quiz questions (identical), similar in content to the quiz questions (similar), or questions that did not appear in the quizzes (new).</p> <p>Independent Variable A: Class (study-only and quiz)</p> <ul style="list-style-type: none"> • One class received daily quizzes while the other did not. • This variable is between-subjects and quasi-experimental (students were not randomly assigned to classes). <p>Independent Variable B: Question-Type (identical, similar and new)</p> <ul style="list-style-type: none"> • Memory test questions were either identical to those used in the quizzes (identical), similar in content to those used in the quizzes (similar), or did not appear in the quizzes (new). • This variable is within-subjects and experimental (questions were randomly assigned to each condition). <p>Dependent Variable: Accuracy averaged across three memory tests.</p>
<p>Results:</p> <p>What were the important results?</p>	<p>There were two results of importance:</p> <ol style="list-style-type: none"> 1. The Class x Question-Type ANOVA and follow-up <i>t</i>-tests show that the quiz group outperformed the study-only group on all three question types. However, this difference was greatest for identical questions (21.8%), then similar questions (17.6%), and smallest for new questions (12.7%). 2. ANOVAs with follow-up <i>t</i>-tests were also run separately on the quiz and study-only groups. The results for the study group showed no differences in performance across the identical (58.4%), similar (62.8%), and new (60.4%) question types. The results for the quiz group however, showed that performance

	<p>for the identical (80.2%) and similar (80.4%) questions was significantly better than performance for the new questions (73.1%).</p>
<p>Inferences:</p> <p>What inferences about the hypotheses and questions can be made based on the results?</p>	<p>Summary: The results of the experiment show enhanced retention of studied material when participants were quizzed daily regardless of whether the study material was actually presented during the quizzes or not. The authors conclude that they successfully replicated the testing effect, previously shown in laboratory settings, in a naturalistic classroom setting. Furthermore, they conclude that the testing benefit generalizes non-quizzed material and therefore instructors do not need to quiz all of the study material to gain the testing effect benefit.</p> <p>These results are consistent with the original hypothesis (H1), and suggest that periodic quizzes can enhance the retention of assigned material (specific question) and testing could be used as an effective teaching technique to improve student learning in a classroom setting (big question).</p> <p>Thinking critically: The authors ran multiple statistical tests on the same set of data without controlling for Type I error inflation. This potentially could impact the interpretation of their results. For example, the interaction between Class and Question Type resulted in a “significant” <i>p</i>-value of .047. From this result, the authors suggested that the testing effect varied across question types. Applying any error correction for the number of statistical tests would render that result non-significant.</p> <p>The use of two different classes, across two different semesters, with different instructors could potentially produce other confounds. As the authors note, they cannot rule out the possibility that some characteristics of the class or the instructor motivated students more in the quiz class than the study class, which lead to better performance.</p> <p>Additionally, students were not randomly assigned to each of the classes. Therefore, some characteristic of the course (e.g., time of day, time of year, instructor) might cause certain types of students to self-select for one class over the other, which could lead to differences in performance.</p>