



## Literature Review on Low-Stakes Testing:

### Introduction:

In the realm of education, assessments serve as vital tools for measuring students' understanding and progress. While high-stakes assessments, with their significant impact on grades and advancement, often take the spotlight, a quieter yet equally important form of evaluation exists – low-stakes testing. This literature review is going to cover scholarly work examining the relationship between low stakes testing and course performance. There will also be discussion regarding different forms of low stakes testing as well as factors that affect students' test-taking performance. By comprehending the nuances of low-stakes testing, educators, policymakers, and students can harness its potential to enhance learning outcomes, improve pedagogical approaches, and create a well-rounded assessment framework.

### Purpose:

The Vice President of Instruction (VP) submitted a research request for the Office of Institutional Effectiveness, Research and Planning (OIERP) to conduct a literature review. Through an in-depth analysis of scholarly literature, we seek to answer key questions surrounding low-stakes testing, such as:

- What are the different types of low stakes testing?
- Does low stakes testing lead to increased learning?
- How effective is low stakes testing by discipline as well as aggregated?
- How is low stakes testing related to course success and persistence?
- How is low stakes testing related to disproportionate impact?

### Findings:

Note, there was no specific research found regarding disproportionate impact in the realm of low stakes testing. There was also little research in regard to low stakes testing by discipline. However, the researcher focused on the different forms of low stakes testing as well as the variables influencing students' performance. The intention was to initiate a comprehensive dialogue across the campus about the integration of this methodology into the classroom. Following this methodology, the answers to the research questions above are listed below:

- Low-stakes testing encompasses various forms such as quizzes, rapid-response tasks, and assignments. While additional types exist, limited research has been conducted on them.
- Evidence suggests that low-stakes testing enhances memory retention by engaging the active retrieval of information from memory, as opposed to relying solely on passive studying.
- Low stakes testing has been shown to be an effective pedagogical tool within education and psychology disciplines, but there is not enough research to conclude that low stakes testing is effective over all disciplines.
- Research indicates that low stakes testing not only boosts performance on subsequent exams but also elevates the likelihood of successfully completing the course.
- The answer to the final question remains ambiguous due to the limited availability of research regarding the effects of low stakes testing on disproportionately impacted groups.

## Review of the testing effect

The testing effect, also known as retrieval practice, is a cognitive phenomenon that challenges conventional notions of learning. It centers around the idea that actively retrieving information from memory, through testing or quizzes, significantly enhances long-term retention compared to more passive studying methods. Sotola & Crede (2021) confirmed this when they took a look at the numerous theoretical explanations for the finding of the testing effect and found that the process of *retrieval*, when students are tested, is a crucial component. One explanation for this effect is that repeated testing forces students to reencode the information they have learned, whereas restudying materials requires them to only reproduce the encoding of the learned knowledge (Karpicke & Roediger, 2008 as cited in Yang et al., 2021). Rather than merely re-reading or reviewing materials, engaging in retrieval practice requires students to actively recall and reconstruct the information, reinforcing neural connections and facilitating durable learning. To elaborate, Iwamoto et al. (2017) cites Roediger & Butler (2011b) when they identified five points regarding the testing effect:

- Testing produces better retention relative to passive studying methods (e.g., re-reading material).
- Testing repeatedly is more beneficial than taking a single test.
- The testing effect can be seen when no feedback (providing correct answers) is given after a test, however, providing feedback yields greater benefits.
- Some time is required between being exposed to material and testing for retrieval practice to be beneficial.
- The benefits of the testing effect are not constrained to learning a specific response, rather, can be generalized to different contexts.

Previous research done in regard to the testing effect has been focused on different types of memory tests, a wide range of materials, across different schedules of studying and testing, and with or without feedback. However, Putnam & Roediger (2017) wanted to investigate the influence of response mode on the testing effect. The authors cited Kellogg (2007) who theorized that because speaking is less practiced than writing, and therefore requires fewer cognitive resources, that speaking would produce better recall. Putnam & Roediger (2017) conducted a study using different response modes (typed, spoken, or covert) with both an experimental and control group. In summary, they found that using different response modes did not have a significant impact on the testing effect in paired-associative learning. These studies suggest that the testing effect is an important tool in students' education regardless of modality (i.e., quizzes, assignments, response type, etc.). Acknowledging the existence of the testing effect, the question becomes, "Is there an optimal way to utilize this effect?"

## Review of the different types of low stakes testing

In the ever-evolving landscape of education, educators continuously seek innovative approaches to improve students' learning outcomes and foster a deeper understanding of the material. One such pedagogical tool that has gained significant attention is low stakes testing. Low stakes testing refers to assessments that carry minimal weight in determining students' final grades but are administered frequently throughout the course. These quizzes/assignments are designed to be less intimidating, allowing students to engage with the material regularly and receive prompt feedback on their performance. By encouraging active retrieval and reinforcing learning, low-stakes testing offers a distinct approach to promoting long-term retention and application of knowledge. There is an array of low-stakes assessments, but for the purpose of this literature review, the discussion will focus on quizzes, quick-response activities, and assignments.

### Quizzes

Quizzing is the most common form of low stakes testing as it is the most versatile, it is quick and informal, and has the inherent ability to strike a balance between assessment and learning enhancement. Quizzing requires students to actively

recall information from memory rather than passively reviewing notes or textbooks, forcing the brain to retrieve information from its storage, which strengthens the neural pathways associated with that information. Research studies have shown the efficacy of frequent quizzing in improving learning outcomes, contributing to the popularity of this assessment method among educators.

According to Rausch & McKenna (2020) there are many benefits to low stakes quizzing. The benefits include better retention, increased learning in subsequent modules, corrective feedback, improved study habits, and metacognitive benefits. They also suggest specific strategies that can increase the effectiveness of utilizing low stakes quizzes. This includes implementing desirable difficulties such as forgetting, spacing, interleaving, overlearning, stimulus variability, and the inclusion of the Generation Effect.

Furthermore, Kenney & Bailey (2021) wanted to specifically examine the claim of better retention of information by applying retrieval practice through daily reviews within the classroom setting. A group of students ( $N = 47$ ) enrolled in a cognitive psychology course participated in these daily reviews conducted at the start of each class. These reviews comprised 2-4 questions designed to encourage students to recall and practice the content discussed in the previous week's lectures. Upon reaching the end of the semester, students undertook a comprehensive final exam that included material from daily reviews, unit exams, both, or neither. Their study confirmed earlier findings indicating that retrieval practice has a positive impact on memory enhancement. Specifically, they observed that students exhibited significantly enhanced performance in responding to questions that had been covered both in daily reviews and unit exams compared to daily reviews only. However, there were no discernible performance differences among items exclusively covered in either daily reviews, unit exams, or those not covered in either.

Additionally, Sotola & Crede (2021) conducted a meta-analysis that delves into the connection between the regular implementation of low stakes quizzes within the classroom setting and the academic accomplishments of students in those particular classes. Drawing data from 52 distinct samples extracted from actual classes ( $N = 7864$ ), their findings propose a moderate correlation of  $d = .42$  between the utilization of quizzes and academic achievement. This correlation is even more pronounced within psychology classes ( $d = .47$ ), particularly when quiz performance contributes to the final class grades ( $d = .51$ ). Furthermore, their investigation uncovers a robust correlation between quiz performance and overall academic achievement ( $k = 19$ ,  $N = 3814$ ,  $r = .57$ ). This underscores that quiz performance is a notably reliable predictor of subsequent examination success. Another significant finding is the substantial surge in the likelihood of passing a class associated with the utilization of quizzes ( $k = 5$ ,  $N = 1004$ , odds ratio = 2.566). Highlighting the odds ratio is a straightforward way to emphasize the statistical significance of this finding. An odds ratio above 1 means a greater chance of success, like passing a class. With an odds ratio of 2.566, the increase in passing a class is substantial.

### Quick-response activities

Kahoot! is a popular pedagogical tool that is widely used in educational settings to engage students and promote interactive learning. It is an online platform that allows educators to create and deliver interactive quizzes, surveys, discussions, and games in a gamified format. Kahoot! is designed to make learning more fun, participatory, and dynamic, while also assessing students' knowledge and understanding of the subject matter. It has gained popularity among educators for its ability to facilitate formative assessment, encourage participation, and create a lively learning atmosphere.

Mdlalose et al. (2021) used Kahoot! as part of a Bachelor of Education program to administer quizzes on various topics to assess students' conceptual understanding. The authors collected data through semi-structured interviews along with the quantitative results from each quiz. The study's principal discoveries revealed that Kahoot! has a noteworthy impact on improving students' academic achievements, drive, and involvement in the context of remote education. Consistent utilization of Kahoot! as a game-centered educational platform seemed to heighten students' comprehensive academic

progress, as indicated by the outcomes achieved in subsequent quiz assessments. Moreover, Kahoot!'s implementation offered avenues for students to actively participate and cooperate as a united community of learners.

Correspondingly, Iwamoto et al. (2017) conducted a quasi-experimental research study in a Psychology class with two sections (a control and experimental group). The groups received the same syllabi, lectures, viewed the same videos and presentation slides, and conversed about the same topics during in-class group discussions. The only difference was in the last 10 minutes of the class. For the experimental group, they were asked to stop conversating and take a Kahoot! quiz while the control group conversated until the end of the class. Both of the groups took the high-stakes final exam the same day. When asking the students "What helped you in class prepare for the exam?", the difference between the control and experimental group were quite different as shown below.

Figure 1. Control Group: What helped you in class prepare for the exam? (n = 23 responses)

**Table: 1**  
**Control Group: What helped you in class prepare for the exam? (n = 23 responses)**

Word Used	Times Mentioned
PowerPoint	11
Notes	9
Study Guide	8
Videos	6
Lectures	3
Discussion	2
Labs	2
Textbooks	2

Note. From Iwamoto, D. H., Hargis, J., Vuong, K., & Taitano, E. J. (2017). Analyzing the efficacy of the testing effect using Kahoot™ on student performance. *Turkish Online Journal of Distance Education*, 80–90.

Figure 2. Experimental Group: What helped you in class prepare for the exam? (n = 24 responses)

**Table: 2**  
**Experimental Group: What helped you in class prepare for the exam? (n = 24 responses)**

Word Used	Times Mentioned
Kahoot	17
Notes	5
Review	4
Quizzes	3
PowerPoints	2

Note. From Iwamoto, D. H., Hargis, J., Vuong, K., & Taitano, E. J. (2017). Analyzing the efficacy of the testing effect using Kahoot™ on student performance. *Turkish Online Journal of Distance Education*, 80–90.

Similarly, to Mdlalose et al (2021), Iwamoto et al. (2017) found that the use of Kahoot! had a significant effect on the test scores beyond chance. Students in the control group mentioned that the PowerPoint (n=11), notes (n=9), and study guide (n=8) helped the most in preparing them for the exam. Conversely, in the experimental group, Kahoot (n=17) helped the most in preparing for the exam. Incorporating games within the classroom environment has proven to be a successful instructional strategy for enhancing academic achievements. This approach seems to harmonize well with the prevailing societal trend favoring mobile applications and video games. Observably, students frequently employ their mobile devices for communication, leisure, and educational purposes on campus. A significant majority of students demonstrate a strong proficiency in utilizing technology as a tool for learning.

**Assignments**

Note, there was minimal research on low stakes assignments, but one article was found by Stewart-Mailhiot (2014) which looks specifically at the benefits, types, and importance of low stakes research assignments. The focus of this article is on utilizing low stakes assignments to enhance research skills in students. The reason for this focus stems from the common experience among undergraduate students in conducting research. The author references Burton and Chadwick (2000) when they found that “94% of students surveyed had an assignment that required locating information in “sources beyond the course textbook, “with 66% of respondents being assigned a research paper”. Stewart-Mailhiot (2014) cites Elbow (1997) regarding five benefits when integrating low stakes writing assignments into curriculum as listed below:

- Low stakes writing helps students involve themselves more in the ideas or subject matter of the course
- When students do high stakes writing they often struggle in nonproductive ways and produce terrible and tangled prose
- Low stakes writing improves the quality of students’ high stakes writing
- Low stakes writing gives us a better view of how students are understanding the course material
- Probably the main practical benefit of frequent low stakes assignments is to force students to keep up with the assigned readings every week

Stewart-Mailhiot (2014) expands on Elbow’s benefits of low stakes research assignments by giving examples along with a rationale of the different kinds of research assignments as shown in the table below:

Figure 3. Examples of Low Stakes Research Assignments

Low Stakes Research Assignment	Rationale
Select one topic that was discussed during lecture. Develop and write out a list of questions or possible research topics related to it.	Students often struggle with selecting a topic. This gives students an opportunity to practice developing and narrowing a topic.
Highlight only those citations on the assigned bibliography that are citations to articles.	Students often have difficulty distinguishing between citations for books, chapters, and articles. This can help them develop that skill and prepare them for citing sources correctly in their own work.
Working with today’s class reading, determine how many sources it references and try to find out how many times it has been cited in other sources (books or article).	Students often see citations as a requirement for avoiding plagiarism, without understanding the value of citations as part of the ongoing conversation taking place within scholarly literature. This exercise can help clarify this connection.
Locate one article each week on the main theme of the course (e.g., Poverty). You are required to use a different database each week and include a brief written description of the database contents/focus (subject coverage, type of publications, ease of use, etc.).	Students often rely on general databases such as Academic Search Complete or ProQuest Research Library. By requiring students to explore other databases, they will become more aware of the breadth of subject specific resources available for future research projects.

Note. From Stewart-Mailhiot, A. (2014). Developing Research Skills with Low Stakes Assignments. *Communications in Information Literacy*, 8(1), 32–40.

An identifying characteristic of these assignments is that they are tasks that carry minimal, if any, weight in determining the overall course grade (low stakes). By structuring several of these tasks or revisiting a specific task in various scenarios, educators can offer students numerous chances to hone the research skills necessary for significant course projects. This enables students to practice these skills and receive constructive input within a relaxed learning setting.

## Review of factors that affect students' test-taking performance

The process of test-taking is a crucial aspect of modern education, serving as a measure of students' understanding, retention, and application of learned material. However, it is widely acknowledged that a multitude of factors intersect to shape an individual's performance during examinations. These factors extend beyond mere subject knowledge and delve into the intricate interplay of psychological, cognitive, environmental, and even socio-economic elements. Recognizing and understanding these diverse factors that influence students' test-taking performance is of paramount importance in fostering a fair and supportive educational environment. For the scope of this section, the researcher focused on motivation/test-taking effort and overconfidence.

### Motivation/Test-taking effort

Motivation during testing refers to the level of drive, interest, and determination that a student experiences while taking an exam or test. It encompasses the psychological and emotional factors that influence a student's willingness to engage with the test, give their best effort, and persist through challenges during the testing process. Motivation during testing can greatly impact a student's performance and the overall outcome of the test.

In the examination of the influence of motivation during test-taking on performance, the analysis conducted by Wise and DeMars (2005) revealed a consistent trend: students who exhibit higher motivation tend to achieve superior results compared to their less motivated counterparts. This suggests that the influence of high ability might be concealed by diminished motivation (as cited by Barry et al, 2021). Correspondingly, Mdlalose et al. (2021) utilized Kahoot! in their study as an attempt to increase student motivation during testing. They found along with previous research that the use of Kahoot! serves to enhance student learning experiences, motivation, and engagement.

Some researchers also discussed test-taking effort and its relationship to motivation. Whereas motivation focuses on the drive to succeed, according to the expectancy-value theory of achievement motivation, test-taking effort is influenced by the individual's belief in their performance (expectancy of success) and the value they associate with the assessment task (task value). The task value includes components such as the desire to perform well, subjective interest, utility, and perceived resource requirement for the task. Research indicates that when assessment results have no personal consequences, students tend to assign less subjective value to the task, resulting in reduced effort and an underestimation of performance. Barry et al. (2010) investigated the trends of test-taking effort and suggested that there are distinct patterns of this effort as listed below:

- **Type 1** may show variations in effort between less demanding non-cognitive tests and more demanding cognitive tests.
- **Type 2** might involve an initial surge in test-taking effort followed by a gradual decline throughout the test session.
- **Type 3** could exhibit a consistent effort pattern across the tests, maintaining either a high, moderate, or low level throughout.

Additionally, Rios (2021) suggests four approaches that can be used to increase students' subjective value of assessments which in-turn can increase test-taking effort. These include increasing test relevance, modifying assessment design, promising feedback, and providing external incentives.

- **Increasing test relevance:** This approach emphasizes increasing the personal relevance of assessment results to students through the way test instructions are framed. This means that rather than raising the stakes, researchers propose making the assessment results more meaningful to students by emphasizing how the results contribute to improving various aspects of education, like classroom instruction, curriculum development, and the reputation of the educational institution.

- **Modifying assessment design:** This approach suggests practical changes to assessment design that can positively impact students' test-taking efforts and overall engagement in the testing process.
  - Avoid lengthy item stems and minimize response options to reduce the perception of mental taxation.
  - Limit the number of open-ended items to maintain students' engagement.
  - Employ computer adaptive testing (CAT), where the difficulty level of items is tailored to individual test-takers' abilities, preventing frustration or disengagement due to item difficulty.
  - Make the assessment content more appealing to students by aligning it with their interests.
  - Experiment with creative approaches like including students' names in item stems or incorporating game design elements in the assessment context.
- **Promising feedback:** This strategy involves providing students with performance feedback related to their test performance, improvement areas, and the utilization of assessment results. It emphasizes the importance of quality and timely feedback, as well as the need for careful planning and allocation of resources to implement the strategy successfully.
- **Providing external incentives:** This strategy involves providing external incentives, like rewards, to students for their test performance. It draws from previous research showing that rewards can enhance academic performance and suggests that these findings can apply to test-taking efforts as well.

Upon accounting for all moderating factors, the external incentives intervention type emerged as the treatment with the most substantial effect on test-taking effort and motivation. Additionally, this particular intervention type, along with the test relevance intervention, demonstrated the most significant influence on test performance. Understanding the multifaceted interplay between intrinsic motivation, external incentives, and strategic assessment design offers a promising avenue for fostering enhanced test-taking effort and ultimately optimizing educational outcomes.

## Overconfidence

Overconfidence in test-taking refers to a cognitive bias where individuals overestimate their own abilities, knowledge, or preparedness for an assessment, leading them to have a higher level of confidence in their performance than what their actual outcomes reflect. Kenney & Bailey (2021) explain that a potential reason behind students' overconfidence lies in their inability to precisely oversee their learning process, a crucial aspect of metacognition. For example, when students highlight text, it doesn't prompt them to engage in metacognitive monitoring to ascertain the depth of their understanding of the information.

Kenney & Bailey (2021) found that students exhibited reduced overconfidence levels when information was introduced through a daily review task and an examination. The application of retrieval practice furnished students with increased insights into their learning progress, regardless of the thoroughness of their grasp of the material. This, in turn, contributed to an enhancement in their metacognitive perception. As students adeptly monitor their learning process, their confidence assessments are likely to become more precise. Additionally, Sotola & Crede (2021) found that students with lower academic performance are likely to experience the greatest advantages from the incorporation of class quizzes. Students with weak academic performance might exhibit a heightened tendency for overconfidence in their skills. The implementation of class quizzes, however, could prove efficacious in mitigating this overconfidence, ultimately encouraging students to adopt appropriate learning behaviors like studying and seeking assistance from instructors.

### Conclusion:

In the studies that were reviewed, low-stakes testing occupies a significant place within the educational landscape. Its distinctive characteristics and purposes offer educators and researchers valuable insights into the assessment process. As an integral component of the broader assessment framework, low-stakes testing continues to shape the educational journey of students while maintaining a distinct role in evaluating comprehension and fostering growth. Throughout this discussion, various aspects of low-stakes testing have been explored, including its benefits, challenges, and the strategies that could be used to enhance its effectiveness. The key findings are shown below:

- A significant finding is the marked improvement in the likelihood of passing a class associated with the use of quizzes. An odds ratio of 2.566 indicates that taking quizzes boosts a student's chance of passing a class by more than 2.5 times compared to not taking quizzes. This highlights a significant advantage for quiz users.
- Another significant finding is the incorporation of games within the classroom environment. This has proven to be a successful instructional strategy for enhancing academic achievements. The contrast in study preferences between the control group (with PowerPoint as the most mentioned resource) and the experimental group (with Kahoot as the most mentioned) underscores the potential benefits of using engaging and interactive methods for preparing students for low-stakes assessments.
- Another significant finding is that students with lower academic performance can gain significant advantages from the incorporation of class quizzes. While these students might display overconfidence in their capabilities, the introduction of quizzes can play a role in reducing this overconfidence, thereby motivating them to embrace more beneficial learning practices, such as engaging in focused study and seeking guidance from instructors.

Numerous studies have explored the effectiveness of low stakes testing, but it is crucial to investigate its impact on student groups that may be disproportionately impacted, as this is essential in truly assessing its overall effectiveness. Furthermore, it is crucial to examine the influence of low stakes testing in various academic fields. While this literature review concentrated on Psychology and Education, numerous other disciplines warrant similar exploration. By considering these factors, educational institutions can make informed decisions about when and how to implement low-stakes testing, ensuring that it aligns with their objectives and contributes positively to students' learning experiences.



## References

- Barry, C. L., Horst, S. J., Finney, S. J., Brown, A. R., & Kopp, J. P. (2010). Do examinees have similar test-taking effort? A high-stakes question for low-stakes testing. *International Journal of Testing*, 10(4), 342–363.
- Iwamoto, D. H., Hargis, J., Vuong, K., & Taitano, E. J. (2017). Analyzing the efficacy of the testing effect using Kahoot™ on student performance. *Turkish Online Journal of Distance Education*, 80–90.
- Kenney, K. L., & Bailey, H. (2021). Low-stakes quizzes improve learning and reduce overconfidence in college students. *Journal of the Scholarship of Teaching and Learning*, 21(2).
- Mdlalose, N., Ramaila, S., & Ramnarain, U. (2021). Using Kahoot! as a formative assessment tool in science teacher education. *International Journal of Higher Education*, 11(2), 43.
- Putnam, A. L., & Roediger, H. L. (2012). Does response mode affect amount recalled or the magnitude of the testing effect? *Memory & Cognition*, 41(1), 36–48.
- Rausch, T., & McKenna, K. (2020). Low Stakes Quizzing: A Tool for Practice Not Assessment. *American Association for Adult and Continuing Education*, 2020
- Rios, J. (2021). Improving test-taking effort in low-stakes group-based Educational Testing: A meta-analysis of interventions. *Applied Measurement in Education*, 34(2), 85–106.
- Sotola, L. K., & Crede, M. (2021). Regarding Class Quizzes: a Meta-analytic Synthesis of Studies on the Relationship Between Frequent Low-Stakes Testing and Class Performance. *Educational Psychology Review*, 33, 407-426.
- Stewart-Mailhiot, A. (2014). Developing Research Skills with Low Stakes Assignments. *Communications in Information Literacy*, 8(1), 32–40.
- Yang, A. C. M., Chen, I. Y. L., Flanagan, B., & Ogata, H. (2021). Automatic Generation of Cloze Items for Repeated Testing to Improve Reading Comprehension. *Educational Technology & Society*, 24 (3), 147–158.