ISSN 2669-2023 (print) / ISSN 2538-8959 (on-line) / **https://doi.org/10.47459/cndcgs.2026.137**

From Practice to Performance: Quantifying the Link Between Mock Examinations and Real Exam Results

Ivana MROZKOVA[1]*, Katerina KUCEROVA[1]

[1]*Language Department, University of Defence, Kounicova 65, 66210, Brno, Czech Republic*

*Correspondence: *ivana.mrozkova@unob.cz*

Abstract

This study examines the relationship between students' performance on mock examinations and their outcomes on subsequent real examinations. The main research question is: To what degree do mock test scores forecast performance on the actual exams, and does the time interval between the mock and real tests affect this predictive relationship? To address this question, the study analyzes a dataset consisting of paired mock and real exam results, including the dates of each assessment, using a quantitative correlational design. Statistical methods include Pearson correlation analysis to measure the strength of association and linear regression modeling to evaluate the predictive validity of mock test performance. The analysis additionally considers time interval effects to examine if the number of days between mock and real exams influences the predictive relationship. Building on the theoretical foundation that practice assessments can support readiness, this research also considers the underlying mechanisms that may strengthen or weaken the predictive link between mock and real examinations. Previous research indicates that the predictive validity of assessments is influenced by contextual variables, including familiarity with the test. Factors such as understanding the test format—which has been shown to impact the reliability and accuracy of assessment results—and cognitive load, which interacts with task complexity and testing environment to affect performance, are conceptually integrated to inform the interpretation of quantitative findings (Reeve et al., 2009; Pengelley et al., 2023). The role of formative feedback is also examined, as systematic reviews and empirical studies indicate that timely, specific feedback enhances academic performance, self-regulation, and understanding, thereby potentially strengthening the predictive value of mock examinations (Hattie and Timperley, 2007; Solis Trujillo et al., 2025; Morris et al., 2021; Sirianansopa, 2024). By integrating temporal aspects with performance metrics, the study aims to determine not only whether mock exams are good predictors but also under what conditions their predictive power is maximized. Results show that mock test scores are moderately to strongly positively correlated with real exam performance, supporting their use as diagnostic tools for assessing exam readiness. They also suggest that shorter intervals between mock and real tests yield greater predictive validity because they reduce opportunities for skill decay or instructional exposure. Ultimately, the study adds to assessment research by offering empirical evidence on the reliability and temporal sensitivity of mock examinations as predictors of real exam outcomes. These findings could be useful for educators, test designers, and institutions aiming to optimize student preparation and assessment schedules.

KEY WORDS: *mock examinations, predictive validity, assessment timing, exam readiness, defence education, English for the Military*

Citation: Mrozkova, I.; Kucerova, K. From Practice to Performance: Quantifying the Link Between Mock Examinations and Real Exam Results. In Proceedings of the Challenges to National Defence in Contemporary Geopolitical Situation, Brno, Czech Republic, 7-10 September 2026. ISSN 2538-8959, **https://doi.org/10.47459/cndcgs.2026.137**

1. Introduction

Predicting student performance in high-stakes examinations has long been a central concern in educational assessment, particularly in contexts where examination outcomes carry significant academic or institutional consequences. One widely adopted preparatory strategy is the use of mock examinations, which simulate authentic testing conditions and provide diagnostic feedback before summative assessment. Despite their prevalence, uncertainty remains regarding the extent to which mock examinations validly forecast performance on subsequent real examinations and the circumstances under which their predictive value is strongest.

Previous research suggests that practice assessments can support learning by increasing familiarity with test formats, reducing test anxiety, and forming subsequent study behaviors. However, the relationship between preparatory assessments and final outcomes is not necessarily stable over time. Variables such as the interval between mock and real examinations, opportunities for further instruction, and potential skill decay may alter the strength of the relationship between mock and real examinations. Understanding how these factors operate is therefore essential for educators and institutions aiming to design assessment schedules that both promote learning and serve as reliable indicators of readiness to take the real exam.

The present study, therefore, examines the relationship between students' performance on mock examinations and their performance on subsequent real examinations. Focusing on listening and reading components of the standardized high-stakes examination (English proficiency examination i.a.w. STANAG 6001), the study addresses two research questions: (1) to what extent do mock examination scores predict performance on real examinations, and (2) whether the time interval between mock and real examinations moderates this predictive relationship. By integrating paired performance data with temporal information, the study moves beyond simple correlation analyses to identify the conditions under which mock examinations function as the most informative and reliable predictors of exam readiness.

The predictive validity of mock examinations is well documented across educational levels and assessment contexts, with prior research consistently demonstrating positive associations between preparatory assessment performance and subsequent high-stakes examination outcomes. Studies employing correlational and regression-based designs have shown that mock examination scores can account for a meaningful proportion of variance in final examination results, thereby supporting their use as diagnostic indicators of academic readiness. For example, Camacho et al. reported that mock board examination performance significantly predicted outcomes on a professional licensure examination, concluding that higher mock scores were associated with an increased likelihood of examination success [1]. Similar findings have been observed in school-level contexts, where mock examination results have been shown to reliably predict students' performance in high-stakes primary assessments [2].

Beyond their statistical association with summative outcomes, mock examinations are theoretically grounded in the broader literature on practice testing and retrieval-based learning. Practice tests have been shown to enhance long-term retention, support metacognitive monitoring, and improve subsequent performance through repeated retrieval of learned material [3]. These mechanisms help explain why mock examinations may not only reflect students' current proficiency but also contribute to improved outcomes on later assessments. From this perspective, mock tests serve a dual role: as formative tools that support learning and as predictive instruments that approximate future exam performance.

However, the strength of the predictive relationship between mock and real examinations is influenced by assessment design and testing conditions. Research on construct-irrelevant variance emphasizes that differences in test format, task demands, or administration conditions can weaken the alignment between preparatory and summative assessments [4]. When mock and real examinations share similar structures, formats, and cognitive demands, predictive validity is expected to be stronger [5, 6]. Cognitive load theory further suggests that unfamiliar formats or excessive task complexity can increase extraneous cognitive load, thereby distorting performance and weakening the correspondence between mock and real exam results [7, 8]. Familiarity gained through mock testing may therefore reduce cognitive load during the real examination, strengthening the observed predictive relationship [9].

In addition, formative feedback plays a central role in shaping the relationship between mock and real examination performance. Extensive evidence indicates that timely, specific, and targeted feedback enhances learning, self-regulation, and achievement [10-13]. When mock examination feedback is effectively integrated into subsequent instruction and student revision, mock scores may become stronger predictors of future outcomes by promoting targeted improvement rather than merely reflecting baseline ability. Conversely, the absence of meaningful feedback may limit the interpretive value of mock results as indicators of readiness.

Temporal factors are critical in understanding the predictive validity of mock examinations. Research on forgetting, retention, and spacing effects suggests that performance decays over time, particularly in the absence of continued practice or reinforcement [14]. Longer intervals between mock and real examinations may therefore weaken predictive relationships due to forgetting or intervening learning experiences that alter students' proficiency levels. Although targeted practice can enhance learning under certain conditions [15], shorter intervals between preparatory and summative assessments are generally expected to preserve the integrity of the mock performance signal, resulting in stronger correlations and more accurate predictions [16].

Mock examinations can serve as meaningful predictors of real examination performance; however, their predictive power is contingent on alignment in assessment design, the provision of effective feedback, and the timing between assessments. By empirically modeling both performance and temporal variables, the study contributes to the literature on assessment use and interpretation (cf. [17, 18]), offering evidence-based insights into when and how mock examinations are most informative as predictors of high-stakes examination outcomes. The temporal dynamic suggests that the interval between mock and real examinations may moderate predictive validity, which is an important but underexamined factor in applied assessment research.

2. Research Method

The study employs a quantitative correlational design to examine the relationship between performance on mock examinations and outcomes on corresponding real examinations. The dataset comprises 205 paired listening scores, including both mock and real exam results, and 205 paired reading scores. The mock and real examinations are aligned in terms of construct and format (English proficiency examination in accordance with STANAG 6001 used in the Czech Republic), ensuring comparability between the two sets of data.

The dataset consists of paired observations of mock and real examination scores across two receptive language skills, listening and reading. For each observation, both the mock score and the corresponding real examination score are recorded. In addition, each pair includes information on the number of days elapsed between the mock examination and the real test. For analysis, each mock–real exam pairing is treated as an independent unit of analysis.

The statistical analysis was conducted in three stages. First, descriptive statistics were employed to explore the distributions and relationships between paired mock and real examination scores. Second, Pearson product-moment correlation coefficients were calculated to assess the strength and direction of the association between mock and real examination performance. Third, linear regression models were estimated to evaluate the predictive validity of mock examination scores, incorporating interaction terms between mock scores and the number of days elapsed to account for potential temporal effects. All analyses were performed separately for listening and reading to allow skill-specific interpretation of the results.

3. Findings

Analysis of paired scores revealed a clear, positive, monotonic pattern: candidates who performed better on mock tests generally achieved higher scores on the actual exam, consistently across cohorts and skill areas.

For listening, Pearson correlation analysis revealed a strong positive association between mock and exam scores ($r = 0.82$, $p < .001$). A similarly strong relationship was observed for reading ($r = 0.79$, $p < .001$). Together, these results suggest that mock exams have substantial predictive validity for both receptive skills.

Linear regression provided further insight into these relationships. In listening, mock scores emerged as a strong, statistically significant predictor of exam outcomes, with a positive slope indicating that higher mock scores corresponded to higher exam scores. Reading followed a similar general trend, with mock reading scores accounting for a considerable share of the variance in exam results.

Across both skill domains, the regression coefficients were large enough to be meaningful in practice, reinforcing the conclusion that mock exams are reliable predictors of real exam performance.

To explore the role of timing, interaction models were constructed incorporating mock scores, the number of days between assessments, and their interaction term. For both listening and reading, the interaction between mock score and elapsed time was negative. This indicates that predictive strength declined as the gap between the two assessments increased. Although mock scores remained significant predictors regardless of timing, their explanatory power was strongest when the mock exam took place closer to the actual test. This trend suggests that shorter intervals help maintain alignment between the skills being measured and the demands of the exam. In contrast, longer gaps introduce additional variability, for example, the need for further instruction or potential skill loss.

Overall, these results answer the posed research questions. First, mock exam scores show moderate to strong predictive capacity for real exam performance in both listening and reading, supporting their role as diagnostic indicators of exam readiness.

Second, predictive validity appears sensitive to timing. Shorter intervals between mock and real exams are associated with stronger predictive relationships, providing empirical support for theoretical perspectives related to test familiarity, cognitive load, and time-dependent feedback effects.

Mock exams administered closer to the target test are more likely to capture a learner's current level of competence under similar cognitive and emotional conditions. In contrast, longer delays allow both learning and forgetting, weakening the direct link between mock and exam outcomes.

From a practical standpoint, these findings highlight the importance of scheduling mock exams strategically, not only as learning opportunities but also as forecasting tools. When prediction is the goal, minimizing the time gap between mock and real assessments is key to maximizing reliability.

The study treats each mock-real exam pairing as an independent observation. Future research could employ multilevel or mixed-effects modeling to explicitly account for within-candidate dependence.

Additionally, expanding the analysis to include productive skills or affective variables (e.g., test anxiety) could further clarify mechanisms underlying predictive validity.

4. Conclusion

This study provides empirical evidence that mock examinations are both predictive and temporally sensitive indicators of real exam performance in listening and reading. By integrating time interval effects with performance data, the research demonstrates not only that mock exams work as predictors, but also when and under what conditions their

predictive power is strongest. These findings contribute to assessment research and offer practical guidance for optimizing the use of mock examinations in educational settings.

Acknowledgements. Both authors contributed to the project throughout all phases – including project preparation, data collection, statistical analysis, and the article. The research was conducted as part of the DZRO Language II project at the University of Defence Language Center.

References

1. Camacho, Y. P., Oberes, M. P. B., Tina, D. M. R., Pepania, A. D., Alfanta, M. P., & Sumicad Jr., E. H. Mock board examination results as predictors in the licensure examination for certified public accountants. Journal of Global Education and Research, 2024, 8(2), pp.180–193.
2. Kago, P. Evaluating the use of Mock Examinations as a Predictor of Primary School Students' Performance in High Stakes Examinations. American Based Research Journal, 2024, Vol-13-Issue-5 May-2024 ISSN (2304-7151)
3. Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., Willingham, D. T. Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. Psychological Science in the Public Interest, 2013, 14(1), pp.4–58.
4. Haladyna, T. M., Downing, S. M. Construct irrelevant variance in high stakes testing. Educational Measurement: Issues and Practice, 2004, 23(1), pp.17–27.
5. Brookart, S. M., Nitko, A. J. Educational assessment of students (8th ed.). Pearson, 2019.
6. Pengelley, J., Whipp, P.R., Rovis-Hermann, N. A Testing Load: Investigating Test Mode Effects on Test Score, Cognitive Load and Scratch Paper Use with Secondary School Students. Educational Psychology Review, № 3, 2023, 10.1007/s10648-023-09781-x.
7. Sweller, J. Cognitive load during problem solving: Effects on learning. Cognitive Science, 1988, 12(2), pp.257–285.
8. Paas, F., Renkl, A., Sweller, J. Cognitive load theory and instructional design: Recent developments. Educational Psychologist, 2003, 38(1), pp.1–4.
9. J., Deci, E. L., Ryan, R. M. Why Teachers Adopt a Controlling Motivating Style Toward Students and How They Can Become More Autonomy Supportive. Educational Psychologist, 2009, 44(3), pp.159–175. https://doi.org/10.1080/00461520903028990
10. Hattie, J., Timperley, H., The Power of Feedback. Review of Educational Research. 2007, Volume 77, Issue 1, pp. 81 – 112.
11. Morris, R., Perry, T., Wardle, L. *Formative assessment and feedback for learning in higher education: a systematic review.* 2021, Review of Education, 9 (3). e3292. doi:10.1002/rev3.3292 ISSN 2049-6613.
12. Sirianansopa, K. Evaluating students' learning achievements using the formative assessment technique: a retrospective study. 2024. BMC Med Educ 24, 1373, https://doi.org/10.1186/s12909-024-06347-5
13. Solis Trujillo, B. P., Velarde-Camaqui, D., Gonzales Nuñez, C. A., Castillo Silva, E. V., Gonzalez Said de la Oliva, M. D. P. The current landscape of formative assessment and feedback in graduate studies: a systematic literature review. 2025, Frontiers in Education, 10, Artículo 1509983. https://doi.org/10.3389/feduc.2025.1509983
14. Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. Distributed practice in verbal recall tasks: A review and quantitative synthesis. Psychological Bulletin, 2006, 132(3), pp.354–380.
15. Kornell, N. Optimising learning using flashcards: Spacing is more effective than cramming. Applied Cognitive Psychology, 2009, 23(9), pp.1297–1317.
16. Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. Psychological Science in the Public Interest, 2013, 14(1), pp.4–58.
17. Messick, S., Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed.). Macmillan, 1989, pp. 13–103.
18. Kane, M. T. Validating the interpretations and uses of test scores. Journal of Educational Measurement, 2013, 50(1), pp.1–73.