Special Issue

Through-Year Assessments: Advancements and Implications for Hybrid Interim-Summative Testing

Introduction to the Special Issue

Through-Year Assessments: Advancements and Implications for Hybrid Interim-Summative Testing

Duanli Yan
Editor

The use of computerized adaptive testing (CAT) in large-scale assessments has gained widespread adoption due to its ability to optimize measurement and reduce the burden on both students and test administrators. In particular, CAT's efficiency has been central to its application in high-stakes assessments, where precision and fairness are critical (van der Linden & Glas, 2010). As the educational landscape evolves to meet the diverse needs of students and teachers, adaptive testing is increasingly being integrated into large-scale assessments to provide more timely, actionable feedback and better alignment with curriculum goals (van der Linden & Glas, 2010). These innovations have led to the development of new assessment models and new assessment systems.

One promising development in assessment is through-course summative assessment (TCSA), which combines scores from multiple tests administered throughout the academic year, instead of relying solely on an end-of-year exam. TCSA aims to provide more nuanced and continuous feedback, thus offering a more comprehensive picture of student progress. Thus, the adaptive through-year assessment system using TCSA represents a significant evolution of traditional summative assessments by providing ongoing formative feedback. The expected benefits of TCSA and through-year assessments are substantial, including finer-grained feedback (Preston & Moore, 2010), increased validity through the inclusion of performance tasks (Bennett et al., 2011), and better alignment with the curriculum (Wilson & Sloane, 2000). However, these models face numerous challenges, such as technical difficulties with score aggregation and increased complexity in designing the assessments themselves (Wise, 2011). In this context, a key innovation is the Through-Year Computerized Adaptive Test (TY-CAT), which promises to address many of these challenges by integrating the advantages of both adaptive testing and through-course models.

This special issue presents a collection of four papers that explore key challenges and innovations in the design and implementation of adaptive testing within through-course summative assessments (TCSA) and through-year assessments (TYA). These papers discuss various technical and policy considerations, offering insights into the potential of adaptive testing to enhance both summative and interim assessments.

The first paper in this issue by Gianopulos provides a literature review focusing on the benefits and challenges of through-course summative assessments (TCSA). TCSA models combine scores from assessments administered at different points throughout the school year, offering benefits such as more granular feedback, reduced measurement error, and greater alignment with curricular goals (Preston & Moore, 2010; Bennett et al., 2011). The author of this paper argue that a through-year computerized adaptive test (TY-CAT) could address many of the technical challenges inherent in traditional TCSA models. By leveraging the flexibility of adaptive testing, TY-CAT could provide more accurate and timely measurements while reducing the burden on students and teachers. This paper sets the stage for further exploration into adaptive testing's potential for continuous, high-quality assessment throughout the school year.

The second paper by Schneider et al shifts focus to the development and policy considerations involved in designing a multiple-administration adaptive test, i.e., a through-year assessment. This paper provides a comprehensive overview of the challenges faced by stakeholders in creating a through-year assessment system that integrates formative and summative assessments into a unified framework. Using a prototype design and score reports, the authors demonstrate how through-year assessments can support both teachers in understanding student progress and states in meeting accountability requirements. The implications of this work are significant for educational policymakers, as it suggests a pathway for developing assessments that not only meet accountability standards but also provide actionable feedback that enhances instructional practice.

The third paper by Gianopulos, et al investigates a critical aspect of adaptive testing: the composition of item banks. Specifically, the authors examine the impact of different item distributions—uniform versus bell-shaped—on measurement outcomes within a hybrid interim-summative assessment model. Using simulations for Grade 4 and Grade 6 mathematics, the study explores how item bank size and distribution affect measurement precision, accuracy, and item exposure rates. This research highlights the trade-offs involved in designing adaptive tests with sufficient precision and fairness across diverse student populations. It underscores the importance of carefully managing item bank composition to ensure robust and reliable assessment outcomes while minimizing potential biases in item exposure.

Finally, the fourth paper by Lee, et al presents an innovative hybrid interim-summative adaptive assessment design that dynamically routes students to off-grade items based on their ability estimates. This design, evaluated through simulations in Grade 4 and Grade 6 mathematics, explores the feasibility of transitioning students to off-grade assessments without compromising the integrity of proficiency determinations. By incorporating both on-grade and off-grade testing, this design aims to support a more comprehensive understanding of student abilities, especially for students performing at the extremes of the achievement distribution. The paper's findings suggest that this adaptive approach could meet federal requirements while also addressing the instructional needs of diverse student populations.

The papers in this special issue underscore the potential of adaptive testing to improve educational assessments by providing more personalized, accurate, and timely measures of student performance. The significance of this research lies not only in advancing the technical aspects of

adaptive testing but also in its implications for policy, test development, and educational practice. By addressing critical issues related to test design, item pool composition, and the integration of formative and summative assessment models, this work contributes to the ongoing evolution of assessment systems that better serve the needs of students, educators, and policymakers alike.

The advancements presented in this special issue provide a clearer path forward for the development of assessments that are not only more adaptable and efficient but also better aligned with the needs of both students and educators. As we move toward more personalized and data-driven educational systems, these innovations are crucial in ensuring that assessments can provide accurate, actionable insights at multiple points throughout the academic year.

References

Bennett, R. E., Kane, M., & Bridgeman, B. (2011). *Theory of action and validity argument in the context of through-course summative assessment*. Center for K–12 Assessment & Performance Management at ETS.

Gianopulos. G. (2025). A literature review of through-course summative assessment models: The case for an adaptive through-year assessment (2025). *Journal of Computerized Adaptive Testing, 12(1)*, 4-34. *CrossRef*

Gianopulos. G., Lee, J., Lim, S., Niu, L., Lee, S., and Choi, S. (2025). The impact of item bank size and item bank distribution on student ability estimates for a hybrid interim-summative CAT. *Journal of Computerized Adaptive Testing,12*(1), 54-87. *CrossRef*

Lee. J., Lim, S., Schneider, M.C., Gianopulos, G., Niu, L., Lee., S., & Choi, S.W. (2025). The impact of item bank transition rules on student ability estimates and achievement level classifications, *Journal of Computerized Adaptive Testing, 12(1)*, 88-122. *CrossRef*

Preston, J., & Moore, J. E. (2010). *An introduction to through-course assessment*. North Carolina Department of Public Instruction.

Schneider, M.C., Choi, S.W., & Lewis, D. (2025). Design considerations and reporting solutions for a multiple administrations adaptive testing system. *Journal of Computerized Adaptive Testing,12(1)*, 35-53. *CrossRef*

van der Linden, W. J. and Glas, C.A,W. (2010). *Elements of adaptive testing*. Springer. New York.

Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education, 13(2),* 181–208. *CrossRef*

Wise, L. L. (2011). *Picking up the pieces: Aggregating results from through-course assessments*. Center for K – 12 Assessment and Performance Management at ETS.

Author Address

Duanli Yuan. Email: dyan08550@gmail.com

Citation