

HIGHLIGHTS

Human in the Loop

Human-centered AI accelerates discovery of knowledge from digital-based large-scale educational assessments.



AUTHORS

Hongwen Guo,
Principal Research
Scientist at ETS

Matthew S. Johnson,
Principal Research
Director at ETS

Luis Saldivia,
Strategic Advisor at ETS

Michelle Worthington,
Assessment
Development Manager

Kadriye Ercikan,
Senior Vice President of
Global Research at ETS,
President/CEO of ETS
Canada Inc.; President of
the International Testing
Commission (ITC)

The problem

- Traditional large-scale assessments (LSAs) typically provide scores but don't explain how or why students arrived at their performance levels.
- Teachers need more than just scores to effectively address gaps in students' learning, especially for low-performing students.
- Performance data alone often falls short in capturing how a student engages with a test, such as how they navigate tasks, use digital tools or manage their time.
- Fine-grained process/log data tracks each step a student takes while interacting with a digital assessment platform. This level of detail paints a holistic picture of how students at different proficiency levels engage with assessments and solve tasks. However, the large volume and complexity of data from LSAs make it challenging for human experts to manage and analyze effectively.

The solution

ETS scientists developed a human-centered AI (HAI) framework that combines the data from student responses with previously-untapped process data.

1. The new HAI framework combines integrated data with **traditional performance data**, such as student responses, with **sequential process data**:



2. The HAI generates “engagement profiles” that help to provide deeper insights into students’ test-taking strategies, time management and cognitive processes.
3. These profiles help reveal why students performed at their current level, offering a richer context for understanding performance beyond a performance score.

HUMAN-CENTERED AI ARCHITECTURE

	Log data		
	Human expert	Algorithm	
	<div>Step 1. Data preprocessing</div> <div>Literature & Human Knowledge</div>	<div>Step 2. Knowledge discovery</div> <div>Human Expertise: Profile creation and definition</div>	
		<div>Step 3. Scaling up</div> <div>Human-in-the-loop annotation for challenging instances</div>	
		<ul style="list-style-type: none"> • Data visualization • Feature creation • Data manipulation <ul style="list-style-type: none"> • discretization • navigation state • padding, etc. 	<ul style="list-style-type: none"> • Autoencoder <ul style="list-style-type: none"> • dimension reduction • sequential information • Clustering • Typical instances
		<ul style="list-style-type: none"> • Active learning: <ul style="list-style-type: none"> • Query • Model updates 	

The HAI framework helps shed light on students’ knowledge gaps, motivation, time management, affective states and other factors that may affect learning.

The research findings are described in The Journal of Measurement and Evaluation in Education and Psychology.*

“The application of HAI is intended to assist and amplify (rather than displace) human expertise in understanding students’ knowledge, skills, and abilities (KSAs), beyond a sole focus on students’ core academic performance measured by large-scale assessments”

Why is this research unique?

The new human-centered framework offers a level of detail from large-scale assessments that was previously only achievable at a small scale.

These insights have the potential to transform how we, as a society, support each student’s learning journey.

NEXT STEPS

- Piloting the new HAI framework with teachers to develop a practical tool for real-world classrooms.
- Using the framework to bridge between large-scale assessments (like NAEP, PISA or PISA for Schools) and classroom settings to support teaching and learning globally.

“Students’ engagement profiles with the visualizations, combined with other complementary information about the students, ... would help educators to prepare meaningful conversations with students who have different profiles for further interventions.”

*Guo, H., Johnson, M., Saldivia, L., Worthington, M., et al. (2024). Human-Centered AI for Discovering Student Engagement Profiles on Large-Scale Educational Assessments. Journal of Measurement and Evaluation in Education and Psychology, 15 (Special Issue), 282-301. <https://doi.org/10.21031/epod.1532846>

About the authors

Hongwen Guo is a principal research scientist at ETS, specializing in statistical modeling and machine learning of multi-source data in measurement. She has led 40 research projects, published nearly 70 articles, and presented at numerous national and international conferences. Guo has also contributed to major educational measurement programs like the SAT® and PSAT®, earning the Bradley Hanson Award in 2019 for her innovative work on writing process modeling.

Matthew S. Johnson is a principal research director at ETS, specializing in statistical methods for education and psychology, with a focus on item response theory. He returned to ETS in 2018 after a career in academia, including roles at Teachers College, Columbia University, and Baruch College. Johnson has served as co-editor of the Journal of Educational and Behavioral Statistics and authored the ETS guidelines on the responsible use of AI in assessments.

Luis Saldivia is a Strategic Advisor at ETS, specializing in assessment development, including the National Assessment of Educational Progress (NAEP), the Advanced Placement (AP) Statistics Test, the Graduate Record Examination (GRE) Mathematics Subject Test, the College-Level Examination Program (CLEP) College Algebra Test, and the Examen de Admisión a Estudios de Postgrado (EXADEP). He also led the ETS team that helped translate the Common Core State Standards-Mathematics (CCSS-M) for assessment purposes for the Partnership for Assessment of Readiness for College and Careers (PARCC), and more recently the translation of the NAEP Mathematical Practices for assessment purposes.

Michelle Worthington is an assessment development manager, specializing in assessment innovation with a focus on mathematics assessment design for K-12 and higher education. She has led cross-functional teams to design and implement new interactive components into assessments, such as graphing tools and simulations.

Kadriye Ercikan is the Senior Vice President of Global Research at ETS and President/CEO of ETS Canada Inc. She is also the President of the International Testing Commission (ITC) and the International Academy of Education. Ercikan oversees ETS's foundational and applied research and the delivery of large-scale assessments like the National Assessment of Educational Progress (NAEP), and the Programme for International Assessment of Adult Competencies. Her research addresses validity, and fairness of issues in assessments, AI applications and adaptivity. Ercikan has authored six books, four journal special issues and over 100 publications, she currently serves as the NCME Book Series Editor (2021-2026).