International Conference on Teaching and Learning Proceeding
Faculty of Education and Teacher Training – Universitas Terbuka
UTCC, South Tangerang, Banten, May 13th, 2026
Volume 4, pg. 16 – 35
ISSN: 3046-594X

FROM ASSESSMENT TO LEARNING: INTEGRATING ITEM ANALYSIS AND LEARNING ANALYTICS FOR DATA-DRIVEN INSTRUCTION IN DISTANCE EDUCATION

Amalia Sapriati[1]*, Kartono[2], Maya Pusptasari[3], Fitria Amastini[4]
[1]Pendidikan Guru Pendidikan Anak Usia Dini, Universitas Terbuka, Bandung, Indonesia
[2,3,4]Universitas Terbuka, Bandung, Indonesia
email: lia@ecampus.ut.ac.id

Abstract: Distance education has changed assessment from reporting static scores to data-rich environments, yet item analysis is still mostly used for test evaluation rather than instructional improvement. This project aims to reposition item analysis as a core component of learning analytics for data-driven online teaching and learning. A thorough literature assessment of a topic framework for synthesizing empirical and theoretical work on item analysis, learning analytics and technology-enhanced learning guided by PRISMA across major databases. The synthesis yields three important findings: (2) validity-focused approaches like Item Response Theory make fairness and interpretability better for different online learners; and (3) formative assessment and adaptive feedback mechanisms turn analytics into actionable instructional interventions that boost engagement and learning results. Current methodologies have limitations in terms of data quality, evidence of causation, interpretability, and challenges in ethical data governance. This work proposes to incorporate item analysis in a learning analytics framework to improve instructional decision-making by combining psychometrics and behavioral analytics. The outcomes of the assessment, especially the item analysis, can be used for the quality assessment of the test, instructional design, personalization and equitable learning in distance education

Keywords: item analysis, learning analytics, distance education, formative assessment, adaptive learning.

:
Accepted: 12 May 2026 Approved: 12 June 2026 Published: 29 Juni 2026

INTRODUCTION

The role of evaluation has radically changed with the rapid development of remote education and online learning settings. Classical test theory (CTT) has always been the main method of evaluation. The analysis of the items was based on statistical indices such as the difficulty of the items (p-values) and discrimination indices acquired from the students' answers. These methods offer useful but limited insight into item quality and test reliability (Lahza et al., 2022). Today's online and mixed learning environments create rich longitudinal data (log files, navigation patterns, and time on task) that can provide a more full picture of student behavior and performance. The behavioral data along with the standard item responses enable

Volume 4, pg. 16 – 35
ISSN: 3046-594X

for a more detailed analysis of the item characteristics and to disclose the learning patterns that were not detected in the traditional study (Lahza et al., 2022).

This shift has been hastened by the combination of learning analytics and evaluation. Current conceptualizations of assessment recognize three interconnected domains: analytics of assessment, analytics for assessment and validity of measurement. These domains emphasize the importance of data in understanding both the outcome of assessment processes and the assessment processes themselves (Lahza et al., 2022). Meanwhile, the use of advanced measurement models such as Item Response Theory (IRT) is used to investigate the functioning of items at different levels of learner ability and to test for measurement invariance, a concern in heterogeneous populations of distance learning (Horzum & Uyanık, 2015). Besides, the research has demonstrated the importance of evaluation to determine the efficacy of blended educational models and the requirement for technology to quantify the impact in both online and offline settings (Giannousi et al., , 2014).

Furthermore, evaluation methods have increasingly included non-cognitive attributes such as self-efficacy and grit, as they can predict performance in online learning environments (Aventurado, 2019; Young & Archer, 2023). These changes are consistent with the broader trend towards formative, adaptive and feedback-rich assessment that allows continuous learning and not just summative evaluation (Chen et al., 2021; Torres et al., 2021). The advent of online testing has also created problems of equity and access. Inclusive and adaptable assessment practices are essential to reliably measure under varied contexts due to the different backgrounds and digital preparedness of learners (Mehran et al., 2017). These patterns are a glimpse of the path of assessment in distant education from a narrow concentration on score-based evaluation to a multi-dimensional data-rich environment.

The CTT-based measures reflect the overall performance and do not account for the underlying learning process or behavioral predisposition inherent in the student replies. Thus, important diagnostic information like as length on item, revision patterns or navigation behaviours are often not taken into account (Lahza et al., 2022). This limitation restricts instructors' understanding of the reasons for students' issues and the ways they use to cope with school. Also standard item statistics may hide problems of idea validity and measurement equivalence. However, using CTT alone may lead to misrepresentations of the constructs to be measured (Horzum & Uyanık, 2015). IRT-based approaches that evaluate item performance across different ability levels and demographic groupings have addressed some of these problems . Item analysis alone (Giannousi et al., 2014), however, cannot provide information about the systemic quality of the assessment, i.e. its alignment with learning outcomes and instructional coherence, especially in complex blended learning contexts.

A further major limitation is that learning analytics are under-utilized; the behavior data generated by online assessments can be used to calibrate items, design examinations, and develop instructional strategies. However, such data are not embedded into analysis and so

Volume 4, pg. 16 – 35
ISSN: 3046-594X

chances for proactive feedback and adaptive learning are missed (Lahza et al., 2022). Most notably, traditional item analysis provides less information for instructional enhancement. It refers to the change of test items without direction on teaching or learning (Chen et al., 2021; Torres et al., 2021).

There are questions of bias and fairness as well. Online assessments may be affected by access, digital skills and testing conditions. However, item analysis alone is not sufficient to compensate for these shortcomings until validity-oriented and analytics-informed approaches are adopted (Horzum & Uyanık, 2015; Mehran et al., 2017). Some scholars have regarded item analysis as a significant diagnostic tool, but a more comprehensive analytical framework has been used (Lahza et al., 2022; Horzum & Uyanık, 2015).

The more precise the educational data, the more important it is to translate exam results into study tactics. The data of the online evaluation comprise the final grades, the students' participation, the way they engage and their behavior. Such numbers enable teachers to grasp their students' knowledge, learning methods and deficiencies (Lahza et al., 2022; Chen et al., 2021; Torres, 2021).

Validity-based assessment approaches such as IRT produce more valid and fair exam results by examining the same ideas for all student groups (Horzum & Uyanık, 2015). Behavioral data can also be combined with item-level performance to better understand how people learn and where they get confused (Lahza et al., 2022). Formative review helps with this transition. Frequent, rapid assessments boost online learning and engagement. Hence, assessment-for-learning practices are increasingly implemented (Chen et al., 2021). Rubrics and self-assessment tools promote self-reflection, skill development, and learning from assessment data (Torres et al., 2021). Indicators of non-cognitive traits such as motivation, grit, preparedness, and SRL can be useful to evaluate a student's success. These parameters can be used to tailor treatments and adaptive learning routes (Mehran et al., 2017; Zhao et al., 2013; Young & Archer, 2023), especially in the online education where students have different learning styles.

Evaluation-instruction gaps have been a problem for distance education. If the evaluations do not fit the learning goals, students may consider the tests irrelevant and unfair, which then affects their motivation and performance in their studies. The imbalance makes it harder to generalize exam results to education.

Formative input helps close the gap. Research shows that timely and regular feedback increases participation and performance in online courses. However, without feedback children get bored and give up (Chen et al., 2021). Assessment data can be utilized in instructional design to develop personalized solutions for each student and to increase the success of a course (Torres et al., 2021).

Volume 4, pg. 16 – 35
ISSN: 3046-594X

This mix is augmented with learning analytics that assists to identify at-risk learners early and make data-driven schooling enhancements. Item-level data and behavioral analytics enable teachers to improve responsive and adaptive learning environments (Lahza et al., 2022; Horzum & Uyanık, 2015). These strategies must also be fair, accessible and open to the requirements of different learners.

Many challenges exist to using assessment data to improve teaching. Data-driven methods may be useful. Testing data and the purpose of what should be learned are not commonly linked and hence it is difficult to use the data to enhance the teaching. The use of classical item statistics restricts the possibility of complete investigation, particularly in the case of online learning environments with big data sets (Horzum & Uyanık, 2015).

It is hard to do data analysis without student activity. According to Lahza et al. (2022), teachers may not be aware of the causes of students' learning problems without behavioral data. Groups with fairness difficulties and measurement errors online demand advanced analytical methods (Horzum & Uyanık, 2015).

Data quality, technology availability, and instructor competence could make data-driven projects impracticable. Digital literacy and access may affect data collection and interpretation and should be considered when organizing the review (Mehran et al., 2017; Aventurado, 2019). Time constraints and lack of teacher professional development make it difficult to improve instructors' teaching approaches with data (Chen et al., 2021; Torres, 2021). The ethical use of learning analytics in schools, thus, should be studied regarding data security and governance ethics (Lahza et al., 2022; Horzum & Uyanık, 2015).

While technology improves conditions, education can spring from data-driven programs. The integration of learning analytics, validity-oriented measurement and formative assessment provides a strong foundation for improving learning processes (Horzum and Uyanık, 2015; Chen et al., 2021; Lahza et al., 2022).

Studies on Item Response Theory (IRT) give reliable assessments for different types of learners. Learning analytics technologies enhance the instructional design through the evaluation of behavioral and performance data (Horzum & Uyanık, 2015; Lahza et al., 2022). Adaptive feedback and self-assessment improve engagement and learning (Chen et al., 2021; Torres et al., 2021; Hemachandran, 2022).

More targeted and successful interventions can be made by assessing non-cognitive traits such as self-regulated learning and readiness (Zhao et al., 2013; Young & Archer, 2023; Mehran, 2017). AI based systems and hybrid learning models allow flexible and data driven training. (Giannousi et al., 2014; Chootongchai and Songkram, 2018).

Item analysis and learning analytics are well established, but missing in instructional design. Test evaluation and learning analytics are used for behaviour prediction but are rarely

Volume 4, pg. 16 – 35
ISSN: 3046-594X

combined for pedagogical intervention. The paper aims to reposition item analysis as an important part of learning analytics. It argues that item-level data can provide cognitive insights on data-driven online learning sessions. The objective of this project is to combine assessment with instruction and to construct adaptive data-driven learning environments by converting assessment data into usable learning insights. The research supports data-based education by demonstrating item analysis as a strategic tool for improving pedagogy in distant education.

METHOD

Review of protocol and technique

In this work, a systematic literature review (SLR) design is used to systematically synthesis empirical and methodological distance education item analysis into learning-oriented insights. Given the multitude of data sources, analytical approaches and teaching scenarios in this area, SLR is superior to narrative reviews because of its openness, reproducibility and scientific rigor. This study was included pre-specified review method according to PRISMA guidelines including systematic search strategies, eligibility criteria, and data extraction and synthesis. How can item analysis and learning analytics improve instructional design and outcomes in online education? References Horzum & Uyanık, 2015 Chen et al., 2021 Torres et al., 2021 Lahza et al., 2022 Mehran et al., 2017).

Database and search strategy

Search in the largest bibliographic database Scopus®. The database indexes major research in educational technology, psychometric and learning analytics including item analysis, online assessment and data-driven pedagogy (Horzum & Uyanık, 2015; Lahza et al., 2022; Chen et al., 2021; Torres et al., 2021) . A two-stage Boolean searching technique was developed using assessment and item analysis, learning analytics and instructional design, and distance/online learning contexts. Search string example:

("item analysis" OR "test item analysis" OR "assessment data" OR "assessment analytics") AND ("learning" OR "instruction" OR "teaching innovation" OR "instructional improvement" OR "learning outcome") AND ("distance education" OR "online learning" OR "e-learning" OR "distance learning")

Querying the database has a distinct syntax. The search method was carried out iteratively, and it was refined according to keywords and relevant papers that were discovered during screening (Chen et al., 2021; Horzum & Uyanık, 2015; Lahza et al., 2022; Torres et al., 2021). Search limited to: Publishing years: 2010-2026 (to show the evolution of learning analytics) English journal papers and conference papers. Search strings, database sources and retrieval dates were provided following PRISMA criteria to ensure the results were reproducible. PRISMA (Figure 1) was followed to select studies in four steps:

1. Search the databases for any records.
2. Screening: Eliminate duplicate and non-relevant titles/abstracts
3. Review of complete text or criterion for inclusion of credentials

International Conference on Teaching and Learning Proceeding
Faculty of Education and Teacher Training – Universitas Terbuka
UTCC, South Tangerang, Banten, May 13th, 2026
Volume 4, pg. 16 – 35
ISSN: 3046-594X

Selection screening was performed by one method, and several reviewers were utilized where available to improve dependability. Disputes were settled by talking. Additional relevant works were found using snowballing (backward and forward citation tracking) such as Lahza et al. (2022), Horzum & Uyanık (2015), Chen et al. (2021) and Torres et al. (2021).

Eligibility Criteria

Studies were included if they met criteria: (1) conducted in distance, online, blended, or technology-enhanced learning environments, (2) analyzed assessment data beyond final scores, including item analysis, IRT, log-data, or formative assessment, (3) employed empirical or validation designs (experimental, quasi-experimental, correlational, case study, or psychometric validation), (4) reported instructional implications, such as feedback, course improvement, or measurement validity, (5) published in peer-reviewed venues. These are consistent with the study's intention to correlate assessment data with instructional progress (Lahza et al., 2022; Horzum & Uyanık, 2015; Chen et al., 2021; Torres et al., 2021; Mehran et al., 2017). In this study, the excluded studies were: (1) focused on traditional classroom assessment, (2) did not address assessment data, measurement models, or learning analytics, (3) were non-empirical opinion papers without systematic evidence, (4) lacked sufficient methodological transparency or accessible full text.

The evaluation notes the multiplicity of methods. Traditional item analysis (CTT) is necessary but not sufficient for diagnosis. The research is focused on IRT, learning analytics and behavioral data, however classical methods are still valuable (Lahza et al., 2022; Horzum & Uyanık, 2015; Chen et al., 2021).

Quality evaluation/Data abstraction

Standardize data extraction for uniformity. The following variables were obtained: study context (online, hybrid, discipline), data (test responses, logs, click streams), methods (IRT, GRM, CTT), main results and implications for learning, validity and fairness, and limitations. Quality assessment focused on methodological rigor, validity evidence, transparency in analysis, instructional relevance. Heterogeneous online learners have been researched for measurement fairness and population variety (Horzum & Uyanık, 2015; Mehran et al., 2017).

Data Analysis

The study employed different analytical methods, such as CTT includes item difficulty and discrimination indices (Giannousi, 2014; Lahza, 2022); IRT covers model fit, item functioning and invariance (Horzum & Uyanık, 2015; Zhao et al., 2013); Log data, response time and navigation behavior is employed in learning analytics (Lahza et al., 2022; Torres, 2021; Consorti, 2015); Formative assessment: participation, feedback and learning outcomes (Chen et al., 2021; Torres et al., 2021); SRL, grit and readiness are non-cognitive analytics (Mehran et al., 2017; Young & Archer, 2023; Zhao, 2013). Many of the research have used mixed-method designs with a combination of quantitative and qualitative data for the purpose of interpretation (Chen et al., 2021; Horzum & Uyanık, 2015; Torres et al., 2021).

International Conference on Teaching and Learning Proceeding
Faculty of Education and Teacher Training – Universitas Terbuka
UTCC, South Tangerang, Banten, May 13th, 2026
Volume 4, pg. 16 – 35
ISSN: 3046-594X

Thematic synthesis (themes)

The findings were considered in four headings: (1) IRT Measurement Validity/Invariance. (2) Learning analytics and behavior insights, (3) Assessment and feedback throughout learning, (4) Equity of learners. Cross-study mapping revealed converging evidence and context-specific differences in data types, analytical techniques, and instructional effects (Lahza et al., 2022; Horzum & Uyanık, 2015; Chen et al., 2021; Torres et al., 2021; Zhao et al., 2013; Young & Archer, 2023). Where it was not possible to do quantitative comparison, findings were presented by direction of effect and level of evidence in absence of meta-analysis.

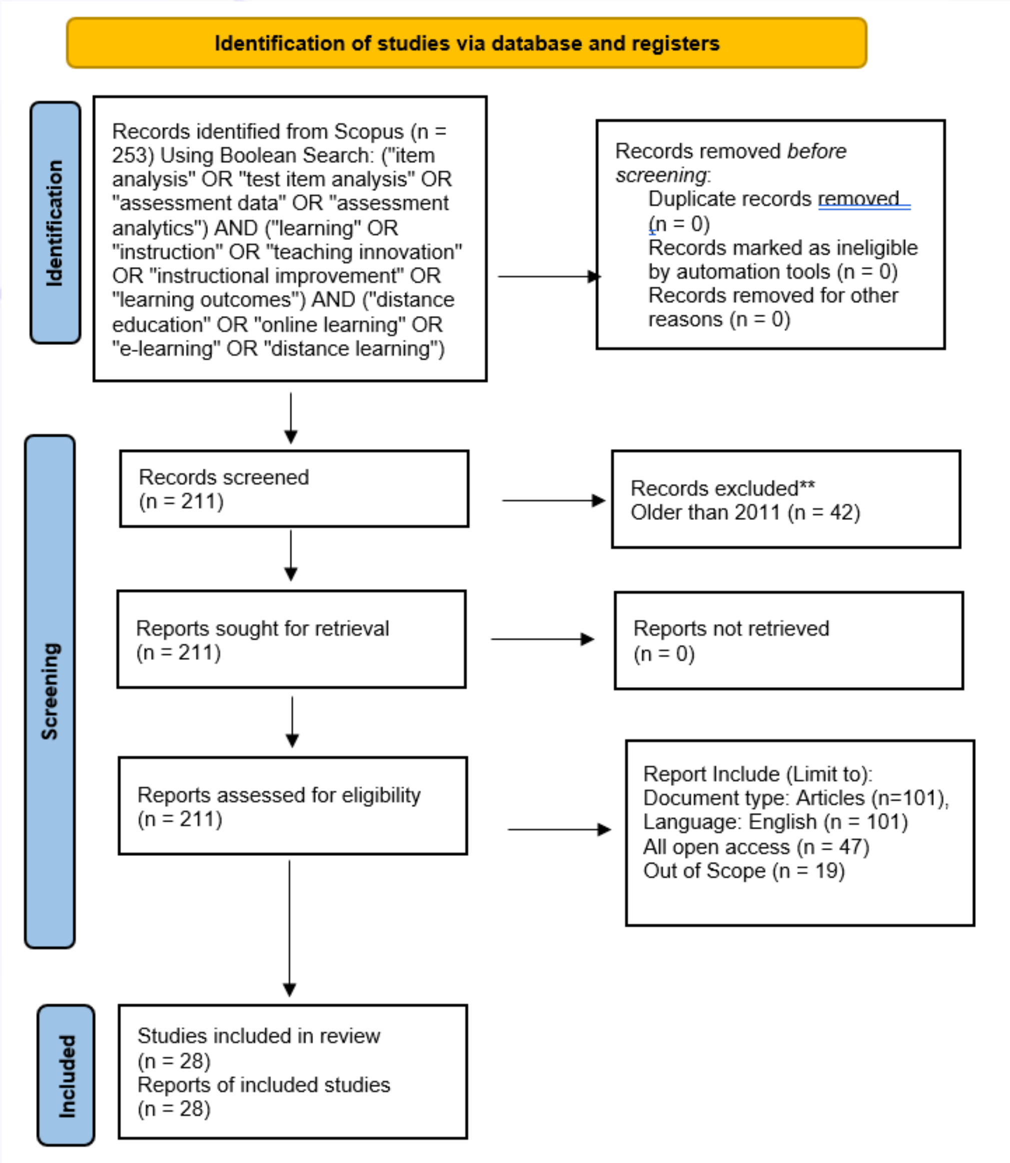

Figure 1. The PRISMA flow diagrams Synthesis Options in Research

Volume 4, pg. 16 – 35
ISSN: 3046-594X

Ethics, reliability, validity.

The review emphasises robustness. Procedures are documented to be transparent and replicable. Triangulating evidence conducted from multiple sources. The analysis assessed bias, quality of data and methodological limitations. The ethics of online education includes ethical use of learning analytics data, privacy and fairness (Horzum & Uyanık, 2015; Lahza et al., 2022; Mehran et al., 2017).

Linking Methods to Instructional Outcomes

The present study differs from the other SLRs in that it combines methodological synthesis with instructional design. The analytical framework includes item analysis (IRT / CTT), behavior data of learning analytics, formative assessment and learner characteristics. This integration turns assessment data into learning insights to deliver data-driven online education (Lahza et al., 2022; Chen et al., 2021; Torres et al., 2021; Zhao et al., 2013; Young & Archer, 2023; Mehran et al., 2017).

RESULT AND DISCUSSION

Items Observation and Evaluation Information

Table 1 gives a summary of 12 notable studies that indicate how the handling of data has improved in the evaluation of online and distant learning, mainly in the field of item analysis. Six papers are about tools, 3 about diagnosis and validation and 3 about assisting learners.

Table 1. Item Analysis and Assessment Data

No	Reference	Study Focus	Main Contribution
1	Čorović et al. (2016)	Item difficulty in e-learning	Demonstrates that item analysis can measure learning improvement based on test performance
2	Horzum & Uyanik (2015)	IRT in online learning	Establishes validity and reliability of instruments through item analysis in distance learning contexts
3	Widyasari et al. (2022)	Item analysis in e-learning	Uses item difficulty to evaluate the effects of learning interventions
4	Giannousi et al. (2014)	Test validity in learning	Applies item analysis to ensure instrument quality in blended learning
5	Consorti et al. (2015)	Item analysis & competencies	Uses item analysis for diagnosing student competencies
6	Kathiresan et al. (2024)	Attitude scale development	Applies item analysis for item selection and instrument validation
7	Lahza et al. (2023)	Beyond item analysis	Integrates item analysis with behavioral data for deeper insights
8	Bayazit et al. (2023)	Formative assessment analytics	Analyzes assessment data to understand student learning behavior

Volume 4, pg. 16 – 35
ISSN: 3046-594X

No	Reference	Study Focus	Main Contribution
9	Castleberry et al. (2023)	Tag-based assessment	Utilizes assessment data for instructional design and feedback
10	Lin & Lin (2025)	Assessment data visualization	Uses assessment data to support needs-based curriculum design
11	Yuniarti et al. (2024)	OBE-based assessment system	Integrates assessment data to determine competency levels and learning recommendations
12	Jia (2024)	Big data in education	Leverages assessment data for personalized learning

The common technique to see the difficulty and the differentiation of the test is by item analysis . Čorović et al. (2016) , Widyasari et al. (2022) , Giannousi et al. (2014) , Kathiresan et al. (2024) and Consorti et al. (2015) . The findings of these studies suggest that item-level metrics are still applicable in e-learning and mixed learning environments. However, evaluative language could also indicate that there is still dependent on Classical Test Theory (Lahza et al., 2022). In this mode of thought, test scores are considered as completion marks, not as resources for lesson planning.
Diagnostic and confirmation use in two trials (n = 3; 25%): Horzum & Uyanık (2015), Consorti et al. (2015) and Kathiresan et al. (2024). The investigations demonstrate idea validity, measurement invariance, and competency diagnosis using item analysis. Horzum and Uyanık (2015) state that IRT allows access to estimate data for distinct student groups. Results demonstrate that online learning should be evaluated in a truth-oriented fashion.

It is obvious that the last 25% of studies (n=3) by Lahza et al. (2023), Bayazit, Castleberry, Lin & Lin, Yuniarti, and Jia are concerned with using review data for learning and analytics. When you put item analysis, behavioral data, visualization, and big data analytics together, you obtain huge conceptual successes in smaller projects. Lahza et al. (2023) fill this gap using item-level measurements and log-data analytics. Traces of behavior and reactions can reveal more than simply accuracy. In contrast, Bayazit et al. (2023) and Castleberry et al. (2023) demonstrate the utilization of test findings to enhance informal feedback and lesson planning.

Of 12 quantitative studies, 3 (25%) utilized test data in the preparation of lessons. This reveals large gaps in the available data and how it is used in the classroom. Evaluation and measurement were incorporated in 9 of the 12 research (75%). In actual life, evaluation and teaching are often separated temporally. One reason for this disparity is because item analysis data is not used sufficiently to aid people in learning.

Similarly, the evolution of data fusion and display was reviewed by Yuniarti et al. (2024) and Lin and Lin (2025). The students and the organization of the course are presented in the research using evaluation data. Jia (2024) investigates evaluation-based customizing of large data. These techniques are limited and so rather focus on behavior or global data rather than on brain analysis at the item level.

Volume 4, pg. 16 – 35
ISSN: 3046-594X

When item analysis is used, it is more useful than judging items (Table 1). The item level assessment has been proved valid in previous investigations (Čorović et al., 2016; Giannousi et al., 2014; Widyasari et al., 2022). In recent studies (Lahza et al., 2023; Bayazit, 2023; Jia, 2024), these indications have been showing to be employed as an aid to instruct. All the research in Table 1 have employed item analysis, but mostly for exams (50%) rather than learning activities (25%). Testing data becomes informative knowledge, not only tests.

Learning analytics as Data-driven learning

Table 2 shows item-level cognitive evaluations are less popular compared to large-scale and behavioral learning analytics (10 research). Trends were noted for predictive analytics (5/10 studies; 50%), data integration and system-level decision support (3/10 studies; 30%) and assessment-informed or hybrid analytics (2/10 studies; 20%). Five research projects are using machine learning, neural networks and data mining to predict student progress. These are Ben Said, Junejo, Lin, Hemachandran and Wen (2024). Studies predict student learning from clickstream, LMS and evaluation data. In 2025, Junejo et al. used the neural network design, and in 2024, Ben Said et al. included the behavioral and assessment input. Ecosystems are becoming more and more predictive. Such algorithms may be biased toward outcome prediction and not beneficial to the classroom for meaningful educational design.

Table 2. Learning Analytics and Data-Driven Learning

No	Reference	Study Focus	Main Contribution
1	Jia (2024)	Big data & learning analytics	Uses assessment and behavioral data to model learning needs and personalize learning
2	Ben Said et al. (2024)	Student performance prediction	Integrates clickstream and assessment data to predict learning outcomes in online learning
3	Junejo et al. (2025)	Predictive learning analytics	Applies neural network models to predict performance using LMS and assessment data
4	Liu & Wang (2026)	Data-driven quality assurance	Utilizes learning and assessment data for AI-based evaluation of educational quality
5	Bayazit et al. (2023)	Formative assessment analytics	Analyzes student behavior in online assessments using learning analytics
6	Lahza et al. (2023)	Advanced assessment analytics	Combines item analysis and log data to generate deeper learning insights
7	Lin et al. (2024)	Automated grading & analytics	Develops AI models to analyze peer assessment data and predict learning outcomes
8	Yuniarti et al. (2024)	Data-driven assessment systems	Integrates multiple assessment data types to determine competency levels and learning recommendations
9	Hemachandran et al. (2022)	AI in learning	Uses assessment data to build predictive models and AI-based learning support systems

Volume 4, pg. 16 – 35
ISSN: 3046-594X

No	Reference	Study Focus	Main Contribution
10	Wen et al. (2024)	Big data & learning evaluation	Integrates data mining and AI to generate data-driven insights for learning evaluation

Jia (2024), Liu & Wang (2026), and Yuniarti et al. (2024) describe methods to integrate rating, behavioral and institutional data for customisation and quality improvement. Liu &Wang, 2026 AI for school evaluationJia, 2024 Big data for student learning simulationEvaluation in connection to the quantity of ability and learning concepts (Yuniarti et al., 2024) The transition from isolated analytics to integrated data ecosystems is still on and this cluster is still more towards system or aggregate level signals than item level signals.

Of the 10 projects, only Bayazit and Lahza link assessment data and learning analytics directly to instructional activities. Lahza et al. (2023) performed an item analysis and log data study on student learning, while Bayazit et al. (2023) conducted a developmental evaluation. These little research projects focus on one big methodological improvement: not using item level data in learning analytics with cognitive and behavioral data.

This is because 80% of the research is based on behavioural or system generated data (clickstream, LMS logs or AI outputs) against 20% cognitive data from items or tests . This bias compounds the issue that learning analytics only assesses behavior, not cognitive evidence from item analysis.

As shown in Table 2, 7/10 (70%) of the study uses AI and machine learning approaches, such as neural networks, automatic grading systems, and predictive modeling frameworks. These approaches provide more scalability and accuracy in future predictions but are more difficult to consume and to learn. Just 3 of 10 (30%) pedagogically oriented analytics research connect data to teaching decisions.

Table 2 illustrates that all the listed studies report that learning analytics is data and computationally expensive. In all study, data are examined in the quantitative distribution but only 20-30% teach with it. This is in support of the study statement that data-to-learning conversion is not enough.

The gap is even bigger in Table 1. Learning analytics are computationally advanced, but not related to the data from the cognitive assessments (Table 2). Item analysis is commonly utilized, but not appropriately applied for teaching (Table 1). The approach of integration proposed is based on the assessment-heavy assessment (Table 1) and the behavior-heavy analytics (Table 2). Most learning analytics research is predictive and behavioral (see table 2). Rarely do they use item level cognitive data integration and instructional design translation. Item analysis needs to be integrated into the practice of learning analytics, to make the data environment more egalitarian, interpretable and meaningful for instructors.

Develop tools for distance learning.

Table 3 lists 10 works on technology-enhanced learning in online education, adaptive systems and data informed instructional design. Quantitative trends were detected in adaptive and personalized learning systems (4/10 studies; 40%), formative and feedback-oriented education (3/10 studies; 30%), and learner-centered readiness and emotional features (3/10 studies; 30%).

The most frequent category (n=4, 40%) is formed by the papers of Consorti et al. (2015), Čorović et al. (2016), Lin et al. (2024) and Li et al. (2025) about personalized and adaptive learning systems. Use data-driven solutions to align student material, pace and learning route. LexMeter adjusts the difficulty of words to the students' achievement (Consorti et al., 2015). (Čorović et al., 2016) extend the technique to adaptive e-learning systems . Lin et al. (2024) and Li (2025) show the potential of AI and learning analytics to automate personalization. This cluster is about flexibility of the system. The methods are based on behavioral data, and may not be based on item-level brain diagnosis.

Table 3. Technology-Enhanced and Adaptive Learning in Distance Education

No	Reference	Study Focus	Main Contribution
1	Consorti et al. (2015)	Adaptive learning system	Development of LexMeter for adapting lexical competence in medical education
2	Giannousi et al. (2014)	Blended learning assessment	Evaluation of blended learning effectiveness based on assessment outcomes
3	Chen et al. (2021)	Formative online assessment	Impact of formative assessment on engagement and learning outcomes
4	Torres et al. (2021)	CAD online learning	Use of rubrics and self-assessment in CAD-based design learning
5	Drijvers et al. (2021)	Digital formative feedback	Implementation of technology-based feedback in mathematics learning
6	Čorović et al. (2016)	Adaptive e-learning	Development of adaptive learning systems based on user data
7	Lin et al. (2024)	AI-based personalization	Use of AI for personalizing learning based on assessment data
8	Li et al. (2025)	Learning analytics framework	Integration of analytics for adaptive instructional design
9	Shulruf et al. (2019)	Readiness & performance	Relationship between learning readiness and performance in online learning
10	Aventurado (2019)	Self-efficacy in e-learning	Role of self-efficacy in the success of online learning

Formative evaluation and feedback-based learning are the subject of three studies by Chen, Torres and Drijvers. Our results suggest that frequent low-stakes assessments with quick feedback enhance student learning. Online course feedback is an incentive and achievement (Chen et al., 2021) Self-assessment rubrics enhance CAD skills (Torres) Digital feedback offers conceptual information (Drijvers et al., 2020). The statistics indicate that 30% of the study connects evaluation with instruction. This links assessment and instruction.

Last Cluster: Giannousi (2014), Shulruf (2019), Aventurado. The topics are integrated learning, student readiness and emotion. Giannousi et al (2014) investigated mixed learning. Self-efficacy and preparation influence student performance (Shulruf et al., 2019; Aventurado, 2019). More than 30% of the research is on non-cognitive and environmental aspects that enhance technology aided learning. For data-based systems to be fair and meaningful, they must be able to assess student aptitude.

As shown in Table 2, four studies (40%) used AI-driven analytics to analyze blended learning and its students but with a more equal depiction of technology and pedagogical views. Around 60% of the research projects used technology-based or AI-assisted approaches such as adaptive systems, automated feedback and analytics frameworks (Consorti et al., 2015; Čorović et al., 2016; Lin et al., 2024; Li et al., 2025; Drijvers et al., 2021; Chen et al., 2021; Torres, 2021; Lin, 2024). Quantitative study 30% Uses data from measures with adaptive, personalized learning systems Most (70%) employ contextual learning elements or aspects of the adaptive system that are not examined. It features sophisticated analytics and assessment but no adaptive educational design.

Table 3 links evaluation, analytics and instructional design with Tables 1 and 2. Non-teaching assessments are dominated by item-level evaluations (Table 1). Table 2: Growth of Behavior-driven Analytics without Cognitive Augmentation. As shown in Table 3, technology-based learning environment training is personalized and flexible but does not include evaluative insights.

As presented in Table 3, online learning is becoming more flexible, personalized and feedback-rich. Wrong quantification: 100% of research show tech increases learning; 30% of research show classes altered using eval data. Adaptive pedagogy, analytics and item analysis are treated as part of the learning-oriented system (Consorti et al., 2015; Giannousi et al., 2014; Chen et al., 2021; Torres et al., 2021; Drijvers et al., 2021; Čorović et al., 2016; Lin et al., 2024; Li et al., 2025; Shulruf et al., 2019; Aventurado, 2019).

Evaluation, Analytics and Instruction in Distance Education Integration

Table 4 shows 10 papers related to distance education, learning analytics and evaluation of instructional design. validity of assessment and measurement (2/10 studies, 20%), formative and feedback-oriented learning (2/10 studies, 20%), learner characteristics and non-cognitive aspects (4/10 studies, 40%), and adaptive and analytics-oriented instructional systems (2/10 studies, 20%).

Volume 4, pg. 16 – 35
ISSN: 3046-594X

Table 4. Integrated Assessment–Analytics–Instruction Frameworks in Distance Education

No	Reference	Study Focus	Key Contribution
1	Lahza et al. (2022)	Item analysis + log data	Integration of response and log data to improve item diagnosis and instructional decision-making
2	Horzum & Uyanık (2015)	IRT validation	Validation of online instruments using IRT and invariance analysis
3	Chen et al. (2021)	Formative analytics	Use of formative assessment to improve engagement and learning outcomes
4	Torres et al. (2021)	Formative CAD learning	Integration of online rubrics and self-assessment in design learning
5	Mehran et al. (2017)	Online learning readiness	Influence of learning readiness on the interpretation of assessment data
6	Zhao et al. (2013)	Self-regulated learning	Role of SRL in online learning success
7	Young & Archer (2023)	Grit in distance learning	Validation of the grit construct in distance learning contexts
8	Aventurado (2019)	Self-efficacy	Influence of self-efficacy on e-learning success
9	Consorti et al. (2015)	Adaptive learning systems	Data-driven adaptive systems for personalized learning
10	Li et al. (2025)	Learning analytics framework	Integrated analytics framework for educational decision-making

The first dimension (n=2; 20%) concerns reliability in evaluation and measurement (Lahza et al., 2022; Horzum & Uyanık, 2015). Lahza et al. (2022) did an in-depth item analysis based on log and response data. They stress that behavioral traces can improve the discrimination and complexity of the items. The validity of online concept and measurement tools was studied by Horzum and Uyanık (2015) using IRT approach. In this work this study applies data driven learning measures. It demands a dispassionate, factual inquiry.

Second dimensional training (n=2; 20 per cent) is formative assessment (Chen and Torres et al., 2021). Regular formative evaluation improves student motivation and accomplishment (Chen et al., 2021) and rubric-based self-assessment boosts online design skills (Torres et al., 2021). Formal integration is mentioned only in 20% of the .” These studies help us to interpret test scores in terms of student learning.

The third and largest dimension is the study of Mehran et al. (2017), Zhao et al. (2013), Young & Archer (2023) and Aventurado (2019) (n=4; 40%). This stusy looks at student characteristics and non-cognitive variables All studies show that preparation, SRL, grit and self-efficacy influence the interpretation of assessment and analytics results. These are the subject of 40% of the study and a vital component for fair and successful data-driven assessments. Teachers focus on the context of the students.

Volume 4, pg. 16 – 35
ISSN: 3046-594X

The last dimension (n = 2; 20 %) is comprised of adaptive, analytics driven education systems by Li et al. (2025) and Consorti (2015). LexMeter is similar to the adaptable modules for personalized learning introduced by Cosorti et al. (2015). Li et al. (2025) suggested an analytics-driven approach for evaluation and training. 20% The lack of fully integrated evaluation, analytics and teaching systems is insufficiently acknowledged.

The 20% deep quantitative research covers measurement, analytics, and student learning (Lahza et al., 2022; Li, 2021). The other 80% is evaluation or formative approaches or learning qualities but they are not connected to each other. This is evident in tables 1 (assessments), 2 (analytics) and 3 (technology-based instructional innovations).

A structural gap is generated by linking four tables together. Only commodities that have aged are analyzed. More complex learning analytics does not mean tests are valid. Finally, technology-enhanced learning provides flexibility but not assessment (see Table 3). As demonstrated in Table 4, most of the research (20% of them attain an overall strategy) do not include clearly defined components.

Fabulous mismatches First, 40% of noncognitive and environmental aspects are crucial, therefore data driven solutions have to be more than technology. They are confronted with the concerns of diversity and equality among learners (Mehran et al., 2017; Zhao, 2013; Young & Archer, 2023; Aventurado, 2019). Second, 20 percent of assessments were adaptive systems. This exemplifies the paucity of evaluation data for initiatives to customize (Consorti et al., 2015; Li, 2021). Third, the missing integration across dimensions leads to a gap between proof and action: data collected and processed are not always used to improve instructional design (Chen et al., 2021; Torres, 2021).

Learning analytics, assessment and technology aided learning have improved but they need a learning oriented framework (Table 4). The literature indicates that for the success of distance education, a holistic model taking into account item level analysis, behavioral analytics, learner characteristics, and adaptive instructional design should be considered (Lahza et al., 2022; Horzum & Uyanık, 2015; Chen et al., 2021; Torres et al., 2021; Mehran et al., 2017; Zhao et al., 2013; Young & Archer, 2023; Aventurado, 2019; Consorti et al., 2015; Li et al., 2025).

Discussion

This study indicates that teachers need to change their evaluation from item level to integrated and analytics based online education assessment. Item analysis, learning analytics and technology-enhanced learning still are different disciplines. We report on a variety of applications of learning analytics to item analysis, the use of assessment data to drive lesson design, data driven methods to improve learning outcomes and the merits and cons of the combination.

Learning analytics pipelines are commodities, not quality control finals. Difficulty and differentiation are common but are not generally indicative of educational quality (Tables 1, 2). Learning analytics conceptualizes non-psychometric behavioral characteristics as Detailed diagnostic tools. Behavioral data such as time-on-task, revisits and navigation patterns as well as item level counts can be utilized to predict student progress (Lahza et al., 2022). The combination permits professors to differentiate between confusion of ideas and sloppy exam taking. Item analysis is not test quality, it is learning. Research (Giannousi et al., 2014) has shown that the learning environment and instructional design affect the assessment scores in blended settings.

Learning analytics has to move from diagnosis to analytics based manufacturing of items. Teachers can utilize indicators from log instead of p-values or discriminating indices to detect mental bottlenecks and adapt questions. Studies on formative assessment show that iterative learning (planning and feedback) is fun and effective (Chen et al., 2021; Torres, 2021). Analysis of the items would impact on the order, directions and layout of the questions. In adaptive learning research (Consorti et al., 2015; Čorović et al., 2016), item attributes are used to inform content dynamically and to connect psychometrics to specific learning paths (see Table 3).

Measurement models with a validity focus that increase assessment-analytics links are also important. Empirical and theoretical studies suggest that IRT and invariance testing are more transparent and fair to test outcomes for different groups of students (Horzum & Uyanık, 2015). These models are consistent with Learning analytics patterns being more related to differences in learning and less related to measurement error or access. This has implications for online learning with heterogeneous students. Table 2 shows that validity is rarely tested in analytics research. This becomes a major obstacle in fair data-driven decision making (Li et al., 2025).

Lesson Planning Empirical Theory Update Test Data The statistics on analytics and evaluation in Tables 3 and 4 indicate that the tutorials are flexible, iterative and learner-centred. (Chen et al., 2021; Torres, 2021) Research on formative assessment has shown that regular low-stakes exercises that provide students instant feedback are associated with increased engagement and academic success. And we can feed extra material to the classroom on the fly if students are taking too long, or making the same mistake. Behavioral indications of Learning analytics can be used for fast therapy (Lahza et al., 2022). They customize student instruction. Customizing lessons needs learner data and analytics. Students' response to adaptive therapies is affected by non-cognitive characteristics such as readiness, self-efficacy, and others (Mehran et al., 2017; Shulruf, 2019; Aventurado, 2019). These elements offer a good tutorial update and reactive based on the analytics. For example, one third of the technology-enhanced learning study covers student preparedness and impact (see Table 3). For best custom they are highlighted data driven, learning is more tailored to the requirements of students. As shown in Tables 2 and 3, analytics-based formative evaluation and feedback are better than summative evaluation in terms of engagement and performance (Chen et al., 2021; Torres, 2021). Adaptive systems are more efficient because they adapt the difficulty of work to the learner's aptitude. This is the case for medical education and design education (Consorti et al., 2015; Čorović et al., 2016). The outcomes enable data-based diagnosis and adjustment of feedback and instructional contents.

But the analysis also discloses the great inadequacies of the present methodologies. 1. Statistical quality and representativeness remain challenging to achieve. For example, Mehran et al. (2017), Li et al. (2025) offered behavioral data as a stronger predictor for access or digital abilities difficulties than learning problems. Which could bias the outcomes. Second, the majority of the study is correlational, and it is hard to understand data-driven solutions (Chen et al., 2021; Torres, 2021). Third, complex analytics models are difficult to understand and handle particularly if people require clear instructions (Li et al., 2025; Lahza, 2022). Some fields or organizations may be limited in use due to the probability of limited discoveries (Horzum & Uyanık, 2015; Lahza et al., 2022).

There are a variety of factors that imply there are gaps in the research. Future multi-site research is needed to demonstrate the sustainability of analytics-based education. In the study, identical and equal results were obtained with the use of item analysis, Learning analytics and validity models (Lahza et al., 2022; Horzum & Uyanık, 2015; Li et al., 2021). Data-driven therapeutic strategies are based on quasi-experiments or RCTs (Chen et al., 2021; Torres, 2021). Further, analytics (Mehran et al., 2017; Zhao, 2013; Young & Archer, 2022) can be used to assess preparation, self-regulated learning, and non-cognitive characteristics through instructional fairness.

At the conference it was agreed that the technology and the human monitoring have to be integrated. AI enabled analytics can provide tailored learning at scale but must be checked by specialists to ensure it is ethical and useful to learning. In complex and high stakes fields such as technical and medical education, the combination of automated analytics with instructor knowledge is desirable (Consorti et al., 2015; Čorović et al., 2016; Torres et al., 2021).

Remote learning is an exciting breakthrough in item analysis, learning analytics and adaptive instructional design. Data may drive decisions that personalize learning, improve education and advance equity. For this we need to investigate and tackle the methodological, ethical and practical problems.

CONCLUSION

This study evaluated the transition of remote education from item-based assessments to learning-centered, integrated frameworks. Item analysis, a key tool in validating tests, apparently is not enough. Rather, learning analytics, validity-oriented measurement, and formative assessment provide more meaningful insight into student and instructor performance. Core but related analytics ecosystem component: item analysis. A major finding. When combined with behavioral data such as log files, reaction times and interaction patterns, item level statistics demonstrate how people learn, think and engage. The integration enables teachers select specific remedial/adaptive feedback and customisation of the course. IRT and measurement invariance provide fair and understandable test results for many online students. The study highlighted formative, feedback-rich evaluation and instruction. Data-driven tests, especially those that offer continuous feedback loops and adaptable learning paths, keep students engaged, help them manage themselves better and boost their performance in online education. Grading is a tool for learning, not a goal. But there are some challenges that need to be addressed before data-driven evaluation can be completely realized. Complexity analytics model quality, readiness, ethics and understanding still are concerns. There is a need for causal research to determine the benefits of analytics-based education, as evidenced by the common correlational studies. The findings of this study indicate how item analysis and learning analytics can change testing and learning. Data and specifically item analysis can be used to evaluate the quality of the exam, improve student support, make lessons relevant and ensure fair and successful online learning. Researchers should develop standard frameworks, establish causality and leverage ethical analytics.

REFERENCES

Angell, K. (2023). Privacy Audit of Public Access Computers and Networks at a Public College Library. *Information Technology and Libraries*, *42*(3). https://doi.org/10.5860/ital.v42i3.16233

Aventurado, P. J. G. (2019). E-Learning Self-Efficacy of Operating Room Nurses of a Selected Hospital in Cebu, Philippines. *Iafor Journal of Education*, *7*(2), 51–68. https://doi.org/10.22492/ije.7.2.03

Chen, Z., Jiao, J., & Hu, K. (2021). Formative Assessment as an Online Instruction Intervention. *International Journal of Distance Education Technologies*, *19*(1), 50–65. https://doi.org/10.4018/ijdet.20210101.oa1

Chootongchai, S., & Songkram, N. (2018). Design and Development of SECI and Moodle Online Learning Systems to Enhance Thinking and Innovation Skills for Higher Education Learners. *International Journal of Emerging Technologies in Learning (Ijet)*, *13*(03), 154. https://doi.org/10.3991/ijet.v13i03.7991

Consorti, F., Basili, S., Proietti, M., Toscano, E., & Lenzi, A. (2015). LexMeter: Validation of an Automated System for the Assessment of Lexical Competence of Medical Students as a Prerequisite for the Development of an Adaptive E-Learning System. *Frontiers in Ict*, *2*. https://doi.org/10.3389/fict.2015.00002

Čorović, S., Mahnič-Kalamiza, S., & Miklavčič, D. (2016). Education on Electrical Phenomena Involved in Electroporation-Based Therapies and Treatments: A Blended Learning Approach. *Biomedical Engineering Online*, *15*(1). https://doi.org/10.1186/s12938-016-0152-7

Drijvers, P., Thurm, D., Vandervieren, E., Klinger, M., Moons, F., Ree, H. v. d., Mol, A., Barzel, B., & Doorman, M. (2021). Distance Mathematics Teaching in Flanders, Germany, and the Netherlands During COVID-19 Lockdown. *Educational Studies in Mathematics*, *108*(1–2), 35–64. https://doi.org/10.1007/s10649-021-10094-5

Giannousi, M., Vernadakis, N., Derri, V., Antoniou, P., & Kioumourtzoglou, E. (2014). A Comparison of Student Knowledge Between Traditional and Blended Instruction In a Physical Education in Early Childhood Course. *Turkish Online Journal of Distance Education*, *15*(1). https://doi.org/10.17718/tojde.99593

Hemachandran, K., Verma, P., Pareek, P., Arora, N., Kumar, K. V. R., Ahanger, T. A., Pise, A. A., & Ratna, R. (2022). Artificial Intelligence: A Universal Virtual Tool to Augment Tutoring in Higher Education. *Computational Intelligence and Neuroscience*, *2022*, 1–8. https://doi.org/10.1155/2022/1410448

Horzum, M. B., & Uyanık, G. K. (2015). An Item Response Theory Analysis of the Community of Inquiry Scale. *The International Review of Research in Open and Distributed Learning*, *16*(2). https://doi.org/10.19173/irrodl.v16i2.2052

Lahza, H., Smith, T. G., & Khosravi, H. (2022). Beyond Item Analysis: Connecting Student Behaviour and Performance Using E-assessment Logs. *British Journal of Educational Technology*, *54*(1), 335–354. https://doi.org/10.1111/bjet.13270

Li, J., Peng, H., Shumei, Z., Peng, Q., Wu, W., Lv, L. F., Yue, L., Zhong, H. J., & Huang, W. (2025). Development of an Interpretable Machine Learning Model for Frailty Risk Prediction in Older Adult Care Institutions: A Mixed-Methods, Cross-Sectional Study in China. *BMJ Open*, *15*(7), e095460. https://doi.org/10.1136/bmjopen-2024-095460

Lin, Z., Yan, H., & Zhao, L. (2024). Exploring an Effective Automated Grading Model With Reliability Detection for Large-scale Online Peer Assessment. *Journal of Computer Assisted Learning*, *40*(4), 1535–1551. https://doi.org/10.1111/jcal.12970

Mehran, P., Alizadeh, M., Koguchi, I., & Takemura, H. (2017). Are Japanese Digital Natives Ready for Learning English Online? A Preliminary Case Study at Osaka University. *International Journal of Educational Technology in Higher Education*, *14*(1). https://doi.org/10.1186/s41239-017-0047-0

Menyani, N., Boumehdi, A., & Jaadi, O. E. (2022). Online Assessment in the Digital Era: Moroccan EFL University Students' Experiences, Perceptions and Challenges. *Iafor Journal of Education*, *10*(1), 193–210. https://doi.org/10.22492/ije.10.1.10

Radiopoulou, D. (2026). Integrating Digital Technologies Into STEM Physics for Adult Learners: A Comparative Study in Second Chance Schools. *Computers*, *15*(2), 94. https://doi.org/10.3390/computers15020094

Shulruf, B., Velan, G. M., Forster, L., O'Sullivan, A. J., Harris, P., & Taylor, S. (2019). Insights Into Student Assessment Outcomes in Rural Clinical Campuses. *BMC Medical Education*, *19*(1). https://doi.org/10.1186/s12909-019-1828-z

Tan, Y., Ito, T., & Okada, K. (2023). Bayesian Penalization for Explanatory Cognitive Diagnostic Models: Covariate DINA Model and Covariate LCDM With the Lasso Prior. *Behaviormetrika*, *51*(1), 237–257. https://doi.org/10.1007/s41237-023-00219-w

Torres, M. J. A., Company, P., Contero, M., & Camba, J. D. (2021). CAD Training for Digital Product Quality: A Formative Approach With Computer-Based Adaptable Resources for Self-Assessment. *International Journal of Technology and Design Education*, *32*(2), 1393–1411. https://doi.org/10.1007/s10798-020-09651-5

Xiao-xia, J. (2024). Practice and Exploration of Ideological and Political Education in Colleges and Universities in the Context of Big Data. *Applied Mathematics and Nonlinear Sciences*, *9*(1). https://doi.org/10.2478/amns-2024-3242

Young, K. A., & Archer, E. (2023). Validating the Grit-S Scale Among Postgraduate Students in a South African Distance Education Institution. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1229433

Zhang, H., & Liu, Q. (2019). Online Learning Method for Drift and Imbalance Problem in Client Credit Assessment. *Symmetry*, *11*(7), 890. https://doi.org/10.3390/sym11070890

Zhao, H., Chen, L., & Panda, S. (2013). Self-regulated Learning Ability Of C Hinese Distance Learners. *British Journal of Educational Technology*, *45*(5), 941–958. https://doi.org/10.1111/bjet.12118