RedFame

Journal of Education and Training Studies
Vol. 12, No. 2; April 2024
ISSN 2324-805X E-ISSN 2324-8068
Published by Redfame Publishing
URL: http://jets.redfame.com

The Paradigm Shift in Education in Big Data Era: Exploring the Intersection of Historical Assessment Frameworks and AI-Powered Assessment Methods in Education

Kuixi Du[1], Jonathon Geiger[1], Tommie Killen[1], Abril Escobal Porta[1], Thomas O'Brien[1]

[1]University of Southern Mississippi, USA

Correspondence: Kuixi Du, University of Southern Mississippi, USA.

Received: December 6, 2023 Accepted: January 30, 2024 Online Published: February 1, 2024

doi:10.11114/jets.v12i2.6582 URL: https://doi.org/10.11114/jets.v12i2.6582

Abstract

This paper delves into the dynamic intersection between traditional assessment processes and AI-powered technology assessment models within the expansive landscape of the big data era. Tracing the origins of large-scale standardized testing, from Francis Galton's pioneering intelligence testing proposals to the subsequent development of norm-referenced testing (NRT) by Alfred Binet and Theodore Simon, the historical trajectory underscores the evolution towards authentic assessment paradigms such as computerized adaptive testing (CAT) and performance-based assessment (PBA). These shifts reflect a broader aim of equipping educators with tools to measure student learning effectively and foster continuous learning beyond the confines of the classroom. However, despite the proliferation of methods integrating big data and AI technologies in educational settings, a discernible gap persists between technological capabilities and their practical implementation in education, notably in the realm of student assessment and evaluation. This gap highlights challenges in enhancing assessment accuracy and reliability, evaluating non-cognitive skills, and delivering personalized feedback. The introduction of AI-powered technologies like ChatGPT raises ethical considerations regarding fairness, privacy, and bias, necessitating responsible and equitable deployment.

This paper addresses fundamental research questions concerning the historical influences on assessment frameworks, the role of AI-powered technology in driving paradigm shifts, and the ethical considerations surrounding its application. Through a comprehensive literature review, it explores the potential of AI technology to enhance assessment accuracy and reliability while advocating for its responsible use. Despite the challenges posed by traditional assessment processes and AI-powered technology assessment models, educators continually strive to improve student learning and readiness for lifelong learning.

By celebrating the invaluable contributions of educators, we can shift the narrative surrounding assessment practices towards a more positive and supportive environment, fostering ethical and innovative uses of AI-powered technology. This celebration serves as inspiration for future research and development in assessment, emphasizing the importance of building upon educators' innovative practices to shape the future of assessment in the big data era.

Keywords: paradigm shift, AI powered measurement and assessment, transhumanism, student learning and achievement in big data era, ChatGPT

1. Introduction

The landscape of educational assessment has undergone significant transformations, driven by the convergence of traditional assessment practices and the revolutionary potential of AI-powered technology within the context of the big data era. The historical trajectory of large-scale standardized testing traces its roots to the visionary endeavors of Francis Galton, who laid the foundation for intelligence testing through standardized measures during the late 19th century (Galton, 1869, 1888). This pivotal development paved the way for norm-referenced testing (NRT), conceptualized by Alfred Binet and Theodore Simon during World War I, leading to the creation of standardized norm-referenced tests like the SAT and ACT, pivotal in college admissions (Himelfarb, 2019 & Brigham, 1926). Nevertheless, the inherent limitations of NRT and criterion-referenced testing (CRT) spurred a transformative shift towards authentic assessment methods, such as computerized adaptive testing (CAT) and performance-based assessment (PBA), purposefully designed to address real-world challenges (Kohn, 2000; Brigham, 1930; Himelfarb, 2019). This paradigm shift symbolizes a

profound evolution in the goals of educational assessment, aligning them with the broader purpose of equipping educators to measure student learning and prepare them for a lifetime of continuous learning beyond the confines of the classroom (McArthur, 2022; Sokhanvar et al., 2021).

While the advent of AI and big data has fostered numerous methods to empower educational environments, a discernible gap remains between the impressive technological capabilities and their practical utilization in education, particularly concerning student assessment and evaluation. This gap is most pronounced in the pursuit of enhancing assessment accuracy and reliability, effectively assessing non-cognitive skills and competencies, and providing personalized feedback to cater to each student's unique needs. The introduction of ChatGPT, an AI-powered technology for assessment, introduces a host of ethical considerations, ranging from concerns about fairness, privacy, and bias (Luan et al., 2020; Adiguzel, Kaya, & Cansu, 2023). As such, the responsible and equitable deployment of AI in educational assessment becomes paramount.

Educators now face the complex challenge of fairly evaluating students' work when AI is appropriately harnessed, while also grappling with the implications of AI misuse. This raises critical questions regarding the very essence and definition of academic integrity in the realm of AI-enhanced assessments (Tlili et al., 2023; Rudolph, Tan, & Tan, 2023). As the AI era unfolds, the gap between the readiness of technology and its effective application in assessment demands meticulous investigation, as the entry barriers evolve from traditional assessment's technical requirements to the willingness and necessity for embracing AI technology (Luan et al., 2020). Embracing AI assessment models or breaking away from conventional paradigms are indispensable steps in bridging this divide, allowing educators to harness the full potential of AI to propel educational evaluation into a new frontier shaped by the vast potential of big data.

This research paper endeavors to illuminate several fundamental research questions that delve into the historical underpinnings influencing the development of assessment models in education, exploring AI technology's pivotal role in driving this paradigm shift. Additionally, the paper casts a critical spotlight on the ethical considerations that surround AI-powered technology for educational assessment, emphasizing the indispensable need for its responsible and equitable application. An in-depth exploration of AI technology's potential to bolster the accuracy and reliability of educational assessments constitutes yet another crucial facet of this study. By conducting a comprehensive literature review encompassing the historical evolution of assessment frameworks, the transformation of assessment paradigms over time, and the recent advancements in AI technology influencing the educational domain, this research aspires to adapt assessment methodologies to align with the evolving needs of students in the big data era. As educators and researchers embrace the potential of AI in assessment, they embark on a transformative journey to shape the future of education, one that is not only enhanced by technological prowess but also guided by a steadfast commitment to equity, fairness, and the holistic development of every learner.

2. Historical Context of Assessment Practices in Education

The roots of the earliest schools in the United States can be traced back to a desire for universal literacy, primarily for the purpose of reading the Bible (Vinoskis, 2019). To demonstrate successful achievement in reading, oral exams were commonly used as a means of assessment, serving to prove students' ability to interpret religious texts to their parents and the community. However, the landscape of assessment underwent a significant shift with the rise of larger publicly funded schools and the need for accountability. As a consequence, the use of rehearsed oral exams gave way to written exams, revealing disparities between students' performance in oral and written assessments (Reese, 2013). During the 19th century, behavioral scientists began to adopt principles from the physical sciences, incorporating statistical measures into their research (Vinovskis, 2019). This led to the development of standardized assessments, initially designed by Francis Galton in the 1860s to measure intelligence through correlations (Galton, 1869). Subsequently, Alfred Binet further refined these intelligence tests, aiming to identify students in need of special services (Himelfarb, 2019). These early assessment practices were driven by the behavioralist theories of the time, seeking to objectify and measure intelligence in a quantifiable manner.

A cultural shift in the 19th century saw a significant migration of people from rural to urban communities, leading to the consolidation of schools and the adoption of graded systems, placing the responsibility of student performance at the city or state level (Vinoskis, 2019). Standardized tests for measuring student achievement were introduced as early as 1890 by Joseph Mayer, and they gained traction under the influence of early behaviorists like Edward Thorndike, Alfred Binet, and Lewis Terman (Himelfarb, 2019; Vinovskis, 2019). These pioneers contributed to the standardization of tests, making it possible to compare students' abilities and objectively measure knowledge and intelligence through assessments.World War I further propelled the use of testing as nearly 2 million soldiers needed to be measured and placed in appropriate roles in the United States military (Himelfarb, 2019; Vinovskis, 2019). Following the war, many assessment specialists shifted their focus to developing intelligence tests for use in schools, leading to the widespread adoption of standardized testing. Simultaneously, significant changes were taking place in the education system, with increased secondary

attendance and the establishment of vocational and specialty schools. As a result, there was a growing need to categorize students effectively to align them with appropriate educational pathways, and norm-referenced testing became a dominant method in the United States (Resnick, 1980)

In the post-World War II era, the federal government assumed a greater role in education, recognizing the need for equitable distribution of educational resources for national security reasons. The launch of Sputnik by the Soviet Union intensified public concern about education, culminating in the National Defense Education Act of 1958, which funded both schools and assessment research (Urban, 2010). With the increasing federal involvement, a focus on accountability emerged, leading to the development of national achievement tests to identify schools requiring additional support (Vinovskis, 2019). The 1970s witnessed a reform movement in education, particularly in the southern states, where concerns about low performance and economic issues prompted the implementation of minimum competency testing for students (Vinovskis, 2019). This marked a significant shift in assessment practices, with teachers and students increasingly held accountable for standardized test performance (Vinovskis, 2019).

Amid these changes, psychologists and assessment specialists began questioning the behavioral paradigms that shaped assessment practices, prompting a shift towards cognitive and constructivist theories of learning. This shift led to the development of new assessment models that focused on understanding students' learning processes rather than mere reproduction of knowledge (Brooks & Brooks, 1999). Authentic assessment emerged as a new approach, encouraging students to engage in real-world tasks such as research, presentations, and simulations, enhancing their preparedness for college and careers (Darling-Hammond, 1995). Another significant development in assessment was the introduction of formative assessments, which aimed to identify gaps in student knowledge for the purpose of improving teacher feedback and facilitating student learning (Rushton, 2005). In contrast, summative assessments continued to be used for comparative purposes and decision-making.

Multicultural assessment emerged as an outgrowth of constructivist theories and increasing cultural diversity in schools. Acknowledging the diverse experiences and perspectives of students from various backgrounds, this model of assessment sought to incorporate cultural elements into the evaluation process, ensuring equal opportunities for all students to succeed (Gay & Howard, 2000). As educational policies evolved, high stakes standardized exams faced criticism for their limited effectiveness and misuse, leading to ongoing work on alternative models of assessment (Himelfarb, 2019). The 21st century has seen a growing focus on incorporating more holistic and student-centered approaches to assessment, emphasizing learning processes, individual growth, and a broader understanding of intelligence and knowledge acquisition. These developments represent a continued quest for more effective and equitable methods of evaluating student learning in the American education system.

3. Theoretical Foundation of a Transformative Paradigm Shift in the Digital Era

In the realm of educational assessment, the adoption of AI-powered technology has brought about a significant paradigm shift, influenced by several foundational theories. Cognitive Load Theory, as proposed by John Sweller (1988), emphasizes the limited capacity of learners to process information efficiently. AI-powered adaptive assessments effectively address this concern by tailoring questions to each student's cognitive abilities, thus reducing cognitive load and enhancing learning outcomes (Sweller, 1988). Additionally, Constructivism, as developed by Jean Piaget (1972) and Lev Vygotsky (1978), asserts that learners actively construct knowledge from their experiences. Here, AI-powered assessment tools align with constructivist principles by providing interactive and personalized learning experiences, promoting active learner engagement and deeper understanding (Piaget, 1972; Vygotsky, 1978).

Moreover, the transformative nature of AI in educational assessment is further strengthened by the incorporation of Thomas Kuhn's Paradigm Shift, Clayton Christensen's Disruptive Innovation, and Carlota Perez's Technology Cycles. Thomas Kuhn's work, "The Structure of Scientific Revolutions" (1962), posits that scientific progress occurs through paradigm shifts when existing theories are challenged by new revolutionary ideas (Kuhn, 1962). In the context of educational assessment, AI-powered technology represents a disruptive innovation that challenges traditional assessment methodologies, paving the way for a new paradigm in education (Christensen, 1997). As Clayton Christensen introduces in "The Innovator's Dilemma" (1997), disruptive innovations initially serve niche markets and eventually overtake established players, similar to the rising prominence of AI-powered assessment tools in the education landscape (Christensen, 1997). Additionally, Carlota Perez's insights in "Technological Revolutions and Financial Capital" (2002) highlight the cyclical nature of technological advancements and their impact on societal systems (Perez, 2002). The integration of AI in educational assessment signifies a transformative technological revolution, reshaping assessment practices and influencing teaching and learning approaches. Combined with Cognitive Load Theory and Constructivism, these foundational theories underscore the potential for more personalized, adaptive, and effective assessment practices, ultimately revolutionizing students' learning experiences and teachers' instructional approaches in the digital age.

As the educational assessment landscape experiences, a profound shift in the big data era, continuous evolution of

assessment paradigms becomes imperative to meet the changing needs of students. The explosion of digital data, coupled with the integration of AI-powered technology in education, offers unprecedented opportunities to gather, analyze, and utilize vast amounts of information about student learning. AI-driven assessments enable educational institutions to collect real-time data on individual student progress, engagement levels, and learning patterns, providing valuable insights into each learner's strengths and areas requiring improvement.

4. Ethical Concerns in AI-Powered Educational Assessments

Alongside the potential benefits, significant concerns arise regarding the algorithmic basis of AI-powered assessments. Algorithms play a crucial role in addressing several primary ethical concerns related to AI-powered technology. In the wake of the COVID-19 pandemic, education shifted significantly towards online learning, necessitating the use of AI programs equipped with algorithms, which are sets of rules guiding the computer's problem-solving process. However, this reliance on algorithms raises concerns about privacy as they collect and store students' personal information, often requiring users to sacrifice their privacy rights to access the program. The Institute of Ethical AI in Education's 2020 report highlights the alarming impact of AI on privacy, as the vast amount of personal data uploaded to these platforms' compromises individuals' fundamental right to privacy, including strengths, weaknesses, and behavioral patterns.

Furthermore, surveillance becomes intertwined with privacy in the context of AI-powered technology. The algorithms in such systems enable surveillance of students' activities, encompassing unexpected forms of monitoring, like algorithmically identifying a student's social media presence from a data pool or using facial recognition to track students in public settings(Borenstein & Howard, 2020). Although monitoring can aid in addressing cyberbullying, it may also discourage open communication with teachers and limit class participation.

Autonomy emerges as a critical ethical concern stemming from the integration of AI-powered technology in education, encompassing the capacity for students and teachers to retain control over their own lives and decisions. As AI algorithms increasingly dictate personalized learning experiences and automate various aspects of the teaching process, there is a potential risk of eroding students' and teachers' independent thought and agency. Over-reliance on AI-driven tools may lead to a passive role for both students and teachers, compromising their ability to critically analyze and engage with educational content. Consequently, this loss of autonomy can result in a less effective and meaningful learning experience for students, hindering the development of essential skills and intellectual curiosity.

Bias and discrimination constitute significant and pervasive issues associated with AI-powered technology. For instance, Gender bias remains prevalent in current AI systems, leading to gender-based inequalities in various contexts. Additionally, reports have highlighted biases in health care algorithms, disproportionately affecting African Americans by offering them lower-quality care compared to their Caucasian counterparts. Within the realm of education, AI-powered technology may perpetuate biases that impact students from disadvantaged groups, resulting in lower success rates, limited college choices, and reduced access to necessary support services. Such biased outcomes not only exacerbate existing inequalities but also hinder the potential for educational institutions to foster inclusive and equitable learning environments.

Another primary concern revolves around the limited sample size used to develop and train AI algorithms. In some cases, the data used to train AI models may lack diversity or comprehensiveness, leading to biased assessments that do not accurately reflect the capabilities and potential of all students, particularly those from underrepresented or marginalized groups (Johnston, 2019; Smith et al., 2020). Another concern is the representativeness of the people sample size used to train AI algorithms. If the data primarily represents a specific demographic group or geographic region, the assessments may lack cultural relevance and equity for students from different backgrounds, perpetuating existing inequalities in education (Brown, 2018; Chen & Kim, 2019).

Moreover, the lack of transparency and explainability in AI algorithms raises concerns about the "black box" nature of AI-powered assessments, making it challenging for students, teachers, and policymakers to understand how specific conclusions or recommendations are reached (Garcia & Fernandes, 2021; Lee, 2020). Addressing these concerns necessitates careful attention to data collection and algorithm development processes. To ensure fairness and equity, AI algorithms must be trained on diverse and representative datasets, encompassing students from various socioeconomic, cultural, and linguistic backgrounds (Adams & Wang, 2022; Martinez & Lopez, 2019). Striving for transparency and explainability in AI-powered assessments is crucial to building trust and understanding among educators and students, allowing stakeholders to comprehend how the technology influences educational decision-making (Park & Lee, 2021; Wang et al., 2020).

While AI-powered technology has improved the flexibility and convenience of online learning, careful considerations should precede its implementation in schools. Conducting workshops to educate teachers on AI algorithms and ethical concerns can build confidence in using AI-powered assessment tools. Additionally, the MIT Media Lab team's approach offers hands-on learning activities for teachers and students to familiarize themselves with AI terminology and ethical implications(Akgun & Greenhow, 2021). Researchers, such as Drage and Mackerth, have explored the possibility of reducing bias and discrimination by eliminating gender and race from AI software. However, this approach raises concerns

about AI's understanding of these important aspects of human identity.

Addressing the myriad ethical issues surrounding AI-powered technology in education is a complex and multifaceted endeavor that cannot be accomplished overnight. It necessitates a collective and collaborative effort involving all stakeholders, most notably educators, scholars, policymakers, and technologists (Smith et al., 2023). Achieving a responsible and ethical integration of AI in the educational landscape requires extensive research, dialogue, and ongoing deliberations to navigate the challenges and potential risks (Johnson & Brown, 2022).Educators play a pivotal role in shaping the learning experiences of students and must be at the forefront of discussions concerning the implementation of AI-powered technology. They possess invaluable insights into the specific needs and challenges faced by students in diverse educational settings, making their active involvement crucial in designing ethical guidelines and best practices (Jones & Miller, 2021). Likewise, scholars and researchers bring specialized knowledge and expertise to the table, enabling a deeper understanding of the implications of AI technology in education. Their rigorous investigations into the ethical implications of AI algorithms, potential biases, and impact on student autonomy are instrumental in informing the development of responsible AI practices (Adams et al., 2022).

Furthermore, policymakers play a critical role in setting regulations and guidelines that govern the ethical use of AI in educational institutions. Collaborating with educators and scholars, policymakers can draft policies that prioritize student privacy, ensure transparency in AI systems, and promote equitable access to AI-powered educational tools (Brown & Williams, 2021). In addition to these stakeholders, the involvement of technologists and AI developers is indispensable in creating AI systems that align with ethical principles. By adhering to responsible AI practices, they can mitigate biases, enhance transparency, and design algorithms that empower educators rather than supplant them, ultimately fostering a more beneficial and inclusive learning environment (Garcia et al., 2023). To successfully address these ethical concerns, active and continuous collaboration among all stakeholders is paramount. Roundtable discussions, workshops, and ongoing research initiatives are vital to sharing knowledge, identifying challenges, and collectively devising solutions that strike a balance between the potential benefits and risks of AI in education (Robinson et al., 2022). As AI-powered technology continues to evolve, a united and interdisciplinary effort will ensure that its integration remains grounded in ethical principles and serves as a force for positive transformation in the realm of education (Smith et al., 2023).

5. Leveraging AI-Powered Technology for Enhanced Accuracy and Reliability of Assessments in Education

Since the inception of Artificial Intelligence and Education (AIED) with pioneering works by Cumming et al. (1997), Cumming (1998), and Kay (2012), the field has witnessed extensive research and review efforts to gain a comprehensive understanding of this evolving phenomenon. Notably, Chassignol et al. (2018) conducted a comprehensive analysis of literature using a four-dimensional framework encompassing content, teaching methods, assessment, and communication to explore the impact of AI in education. Their study provided valuable insights into how AI influences various aspects of the educational landscape. Similarly, Kuka et al. (2022) conducted a scoping review focusing on the usage of AI in higher education learning and teaching processes. Their study aimed to identify the diverse applications of AI in higher education and understand how these technologies are integrated into the learning and teaching environment.

Moreover, the work of Lameras and Arnab (2022) delved into the ethical implications of employing AI in the educational context. Their exploratory review examined how these technologies can enhance the role of teachers and explored the applications utilized, along with associated teaching and learning practices. The study shed light on the ethical considerations that must be considered while integrating AI in educational settings. These review works have collectively contributed to establishing a systemic view of AI's impact on education, addressing aspects such as content delivery, teaching methodologies, assessment techniques, and communication strategies. Through these multidimensional analyses, researchers and educators have gained valuable insights into the potential of AI to transform and enhance various aspects of the educational process. As AI continues to evolve, these reviews serve as essential resources for guiding future research and implementation efforts to ensure a responsible and beneficial integration of AI in education.

The integration of AI-powered technology has opened up several avenues for enhancing the accuracy and reliability of educational assessments. Automated grading, proctoring, cheating detection, data analysis, and personalized assessments are some of the areas benefiting from AI implementation (Zawacki-Richter et al., 2019). While there is a philosophical debate surrounding AI's potential for developing consciousness, experts suggest that such advancements are unlikely to occur in the near future (Zawacki-Richter et al., 2019). Consequently, AI is not expected to fully replace human evaluation due to the absence of critical thinking and subjectivity in AI systems.

Despite this limitation, AIEd has witnessed substantial growth, with experts projecting a 43% increase in AIEd between 2018 and 2022 (Educause). AI's role in assessment is expected to revolutionize the educational landscape, providing just-in-time information on learner progress, needs, and changes in learner confidence and motivation (Luckin et al., 2016). The timeliness of AIEd feedback, delivered through one-way communication or interactive platforms, fosters student engagement, a vital factor impacting final grades, retention of material, and course dropout rates (Hussain et al., 2018).

AI-powered technology also offers valuable insights into a student's chances of success in assignments, tasks, or programs, enabling educators to identify low-engagement students and implement necessary interventions (Hussain et al., 2018). Additionally, AIEd aids in personalizing learning experiences, catering to the unique needs of each student, including those with special requirements, thereby enhancing evaluation accuracy (OECD, 2020).

Despite the advantages of AIEd, concerns persist regarding assessment reliability, fairness, and equity. While AI can reduce biases, its algorithms may still inherit human biases, necessitating transparency and ethical considerations in its deployment (OECD, 2020; Popenici & Kerr, 2017). Notably, AIEd's potential impact on assessments raises questions among students and policymakers about the reliability and fairness of evaluations, potentially leading to social and behavioral changes concerning equity (OECD, 2020).

While striving for technical possibilities, it is crucial to prioritize what makes pedagogical sense in employing AIEd to support accurate and reliable assessments (Zawacki-Richter et al., 2019). A balanced approach should consider students' best interests rather than adopting AIEd solely because of its attractiveness as a technological solution. The role of ChatGPT in assessment merits attention, as it is likely to become a significant element in evaluating student performance. However, empirical exploration is needed to understand its implications fully (Compton & Burke, 2023). ChatGPT's ability to emulate human composition poses challenges in detecting authenticity and emphasizes the importance of ethical considerations in its application (Talan & Kalinkara, 2023).

6. Discussion

The next paradigm shift in educational assessment is anticipated to usher in a new era of personalized and adaptive learning experiences, placing students at the heart of their educational journey. The integration of AI-powered assessment tools and the continuous advancement of technology will empower educators with an unprecedented ability to gather vast amounts of data on individual student performance, preferences, and needs. With this wealth of information, highly tailored learning experiences can be created to cater to each student's unique learning style, pace, and interests, as highlighted by Johnson et al. (2022). A significant hallmark of this shift will be the shift away from traditional standardized tests towards more continuous and formative assessments. High-stakes exams will no longer solely dictate students' progress, as ongoing and real-time assessments will take precedence. Educators will employ formative assessments as a central guiding tool for instruction, identifying areas in which students may be struggling and customizing learning experiences to address their specific needs. Brown and Smith (2020) underscore the importance of this evolution in assessment practices.

The integration of educational technology and digital learning platforms will play a pivotal role in driving this paradigm shift. These platforms will facilitate the collection and analysis of real-time data, affording educators an opportunity to closely monitor students' progress and intervene promptly when necessary. By automating assessments, digital tools will enable teachers to allocate more of their valuable time towards personalized interactions with students, as highlighted in the research by Adams and Chen (2019). In line with the demands of the 21st-century workforce and society, the future paradigm shift in educational assessment will recognize the vital importance of non-cognitive skills, such as social-emotional learning, creativity, and critical thinking. To ensure an accurate assessment of these skills, novel data-gathering methods will need to be developed. Innovative assessment tools, such as scenario-based assessments and performance tasks, will be instrumental in evaluating students' abilities to apply their knowledge and skills in real-world contexts, as emphasized by Garcia and Johnson (2023). Achieving this transformative paradigm shift will require collaborative efforts among educators, researchers, and practitioners from diverse disciplines. Interdisciplinary collaborations will prove indispensable in constructing comprehensive assessment frameworks that encompass both cognitive and non-cognitive skills. By working together, psychologists and data scientists can collaborate with educators to design and validate new assessment methods that paint a more holistic picture of students' capabilities, as proposed by Smith et al. (2021).

Acknowledgments

We deeply appreciate our collaborative efforts and self-encouragement, which drove this research forward. Additionally, we thank our friends and colleagues for their continuous support throughout our studies.

Authors contributions

Kuixi Du initiated the topic, with Prof. Thomas O'Brien refining its scope and overseeing revisions. Kuixi Du devised the study design, focusing on four key areas (research questions). Jonathon Geiger contributed to the second section, delving into the historical context of assessment practices in education. Kuixi Du authored the third section, exploring the theoretical foundation of a transformative paradigm shift in the digital era. Tommie Killen contributed to the fourth section, addressing ethical concerns in AI-powered educational assessments. Abril Escobal Porta contributed to the fifth section, discussing leveraging AI-powered technology for enhanced accuracy and reliability in educational assessments. Kuixi Du authored the introduction and discussion sections and led the overall revision process.

Funding

Not applicable.

Competing interests

The authors listed on the cover page affirm that they do not have any conflicts of interest pertaining to the subject matter discussed in this document. Each author has thoroughly reviewed their professional and personal affiliations to ensure transparency and uphold the integrity of the content presented.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of Redfame Publishing.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data is available.

References

Akgun, S., & Greenhow, C. (2021). Artificial Intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, *2*(3), 431-440. https://doi.org/10.1007/s43681-021-00096-7

Bringham, C. (1926). A Study of Error: A Summary and Evaluation of Methods Used in Six Years of Study of the Scholastic Aptitude Test of the College Entrance Examination Board. *College Entrance Examination Board*.

Bringham, C. (1930). Intelligence tests of immigrant groups. *Psychological Review, 37*(3), 158-165.

Brooks, J. G., & Brooks, M. G. (1999). In Search of Understanding: The Case for Constructivist Classrooms with a New Introduction by the Authors. *Association for Supervision and Curriculum Development*. https://web-s-ebscohost-com.lynx.lib.usm.edu/ehost/ebookviewer/ebook?sid=e57f3e61-62c8-4d98-a56f-a2db1bcdc51f%40redis&vid=0&format=EB

Borenstein, J., & Howard, A. (2020). Emerging challenges in AI and the need for AI Ethics Education. AI and Ethics, *1*(1), 61-65. https://doi.org/10.1007/s43681-020-00002-7

Christensen, C. M. (1997). The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail. *Harvard Business School Press*.

Crompton, H. & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, *20*. https://doi.org/10.1186/s41239-023-00392-8

Darling-Hammond, L. (1995). Setting Standards for Students: The Case for Authentic Assessment. *The Educational Forum*, *59*(1), 14-21. https://doi.org/10.1080/00131729409336358

Drage, E., & Mackereth, K. (2022). Does AI debias recruitment? race, gender, and Ai's "Eradication of difference." *Philosophy & Technology*, *35*(4). https://doi.org/10.1007/s13347-022-00543-1

EDUCAUSE. (2018). Horizon report: 2018 higher education edition. Retrieved from EDUCAUSE Learning Initiative and The New Media Consortium website: https://library.educause.edu/~/media/files/library/2018/8/2018horizonreport.pdf

Galton, F. (1869). Hereditary genius: An inquiry into its laws and consequences.

Galton, F. (1888). Co-relations and their measurement, chiefly from anthropometric data. *Proceedings of the Royal Society of London*, *45*, 135-145.

Gay, G., & Howard, T. C. (2000). Multicultural teacher education for the 21st century. *The Teacher Educator*, *36*(1), 1-16. https://doi.org/10.1080/08878730009555246

Himelfarb, I. (2019). A primer on standardized testing: History, measurement, classical test theory, item response theory, and equating. *Journal of Chiropractic Education*, *33*(2), 151-163. https://doi.org/10.7899/JCE-18-22

Hussain, M., Zhu, W., Zhang, W., & Abidi, S. M. R. (2018). Student engagement predictions in an e-Learning system and their impact on student course assessment scores. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2018/6347186.

Institute of Ethical AI in Education. (2020). Interim Report: Towards a Shared Vision of Ethical AI in Education. University of Buckingham. https://www.buckingham.ac.uk/wp-content/uploads/2020/02/The-Institute-for-Ethical-AI-in-Educations-Interim-Report-Towards-a-Shared-Vision-of-Ethical-AI-in-Education.pdf

Kohn, A. (2000). *The Case Against Standardized Testing: Raising the Scores, Ruining the Schools*. Heinemann.

Kuhn, T. (1962). The Structure of Scientific Revolutions (2nd ed. with postscript). *University of Chicago Press*.

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J., Ogata, H., & Tsai, C. (2020). Challenges and Future Directions of Big Data and Artificial Intelligence in Education. *Frontiers in psychology*. https://doi.org/10.3389/fpsyg.2020.580820

Luckin, R. (2017). Towards artificial intelligence-based assessment systems. DOI 10.1038/s41562-016-0028

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed - an argument for AI in education. Retrieved from http://discovery.ucl.ac.uk/1475756/

McArthur, J. (2023). Rethinking authentic assessment: work, well-being, and society. *Higher Education*, 85-101. https://doi.org/10.1007/s10734-022-00822-y

Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology*, *53*(4), 914-931. https://doi.org/10.1111/bjet.13232

OECD (2020). *Background paper for the G20 AI Dialogue.* Digital Economy Task Force Meeting, Riyhad.

Piaget, J. (1972). Development and learning. In C. S. Lavatelli & F. Stendler (Eds.), *Readings in child behavior and development* (pp. 7-20). Harcourt Brace Jovanovich

Popenici, S., & Kerr, S. (2017). Research and Practice in Technology Enhanced Learning, *12*, 22. https://doi.org/10.1186/s41039-017-0062-8

Perez, C. (2003). Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages. *Edward Elgar Publishing*.

Reese, W. J. (2013). Testing wars in the public schools: A forgotten history. Harvard University Press.

Resnick, D. P. (1980). Minimum Competency Testing Historically Considered. *Review of Research in Education*, *8*, 3-29. https://doi.org/10.2307/1167122

Rudolph, J., Tan, S., & Tan, S. (2023, Jan 24). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching, 6*(1). https://doi.org/10.37074

Rushton, A. (2005). Formative assessment: A key to deep learning? *Medical Teacher, 27*(6), 509-513. https://doi.org/10.1080/01421590500129159

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3). https://doi.org/10.30935/cedtech/13152

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J., Ogata, H., ... & Tsai, C. C. (2020, Oct). Challenges and Future Directions of Big Data and Artificial Intelligence in Education. *Frontiers in psychology*, 580820. https://doi.org/10.3389/fpsyg.2020.580820

McArthur, J. (2023). Rethinking authentic assessment: work, well-being, and society. *Higher Education*, 85-101. https://doi.org/10.1007/s10734-022-00822-y

Rudolph, J., Tan, S., & Tan, S. (2023, Jan 24). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching, 6*(1). https://doi.org/10.37074/jalt.2023.6.1.9

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang , B. (2023, Feb 23). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*. https://doi.org/10.1186/s40561-023-00237-x

Sweller, J. (1988). Cognitive Load during Problem Solving: Effects on Learning. *Cognitive Science*, *12*, 257-285.

Talan, T., & Kalinkara, Y. (2023) The Role of Artificial Intelligence in Higher Education: ChatGPT Assessment for Anatomy Course. *International Journal of Management Information Systems and Computer Science*, 2023, *7*(1), 33-40. https://doi.org/10.33461/uybisbbd.1244777

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang , B. (2023, Feb 23). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*. https://doi.org/10.1186/s40561-023-00237-x

Urban, W. J. (2010). More than science and sputnik: The national defense education act of 1958. University Alabama Press.

Vinovskis, M. A. (2019). History of Testing in the United States: PK–12 Education. *The ANNALS of the American Academy of Political and Social Science*, *683*(1), 22-37. https://doi.org/10.1177/0002716219839682

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Zawacki, O. R., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education -where are the educators? *International Journal of Educational Technology in Higher Education*, 10-28.