JOURNAL OF BALTIC SCIENCE EDUCATION
ISSN 1648-3898 /Print/
ISSN 2538-7138 /Online/

SCIENCE TEACHERS' APPROACH TO CONTEMPORARY ASSESSMENT WITH A READING LITERACY EMPHASIS

Maja Kerneža,
Dejan Zemljak

Abstract. *In a sample of 1215 teachers, this study examined the readiness of science educators for assessment in the rapidly evolving landscape of artificial intelligence in education. Participants responded to an online questionnaire during the emergency remote teaching phase, offering insights into the frequency and nature of assessment methods utilized. The research draws a connection between assessment techniques during remote teaching and the emergence of AI in education. The results show that the selected assessment methods vary across teachers, with some specific differences observed in the assessment practices of science teachers. The study underscores the critical role of reading literacy in enhancing student engagement in contemporary learning environments. Moreover, the findings suggest that continuous professional development significantly improves the readiness of (science) teachers for AI-enhanced assessment. Drawing from these insights, recommendations for subsequent research are delineated.*

Keywords: *artificial intelligence, assessment, reading literacy, science teachers, teacher training*

Maja Kerneža, Dejan Zemljak
University of Maribor, Slovenia

Introduction

The OECD (2022) defines AI systems as machine-based entities capable of making predictions, recommendations, or decisions influencing real or virtual environments, aligned with human-defined objectives. Many educators and professionals remain unaware that national AI education policies were established even before ChatGPT's emergence, a milestone in AI's educational integration. The OECD began addressing this as early as 2021 (Galingo et al., 2021), summarizing various national strategies. Notably, AI educational guidelines, such as "Is education losing the race with technology" (OECD, 2023) and "OECD digital education outlook" (OECD, 2021), were established before ChatGPT's advent. The European Commission has also contributed with publications like "Ethical Guidelines on the use of artificial intelligence (AI) and data in teaching and learning for Educators" (European Commission, 2022) and "White paper on artificial intelligence – a European approach to excellence and trust" (European Commission, 2020). However, until Open AI popularized ChatGPT in November 2022, these documents seemed futuristic. Now, they hold immediate relevance. Natural language processing tools, exemplified by ChatGPT, have revolutionized education. This AI-driven model assists in diverse tasks, from coding to essay writing. As a result, the educational sector is undergoing significant transformation. The predictions of Seldon and Abidoye (2018) in "The Fourth Education Revolution" have rapidly materialized. The sudden emergence of such tools found many educators unprepared, leading to polarized views on AI's role in education. Regardless of perspective, AI's ubiquity necessitates strategic adaptation. Some nations have even implemented regulations or outright bans on tools like ChatGPT, emphasizing the need for ethical engagement with AI in education (Bhati, 2023; McCallum, 2023; Yang, 2023). Cotton et al. (2023) highlight the dual nature of AI in education, presenting both challenges and opportunities. Educational assessment serves multifaceted purposes, with its nature diverging based on the chosen method. Criteria such as construct validity, reliability, desired impact, and resource optimization

https://doi.org/10.33225/jbse/23.22.851

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

are essential for evaluating assessment techniques. Teacher assessment remains central to these criteria (Harlen, 2007). Conventional assessment methods, primarily paper-based, are familiar to educators (Montgomery, 2002), however, these methods are characterized by rigid question formats, offer a limited snapshot of an individual's capabilities often sidelining multifaced learning perspectives (Frey & Schmitt, 2010), and do not cater to students' unique knowledge, skills, and experiences and evaluate competencies typically beyond computerization (Swiecki et al., 2022). While teachers have incorporated technology to support conventional assessment practices (Grion et al., 2018) and have adapted their pedagogical and evaluative strategies during exigencies (Legvart et al., 2022; Turan-Güntepe et al., 2023), AI's advent in grading introduces novel dimension. AI's inherent nuances necessitate educators' cognizance during evaluation, transcending mere performance measurement. Educators must contemplate assessment's overarching role and objective, as it significantly influences both teachers' and students' value perceptions. Such reflection enables judicious student evaluation, catering to the varied expectations of stakeholders through the assessment trajectory and its outcomes (Dunn et al., 2004, p. 15). Yet, educators must remain vigilant of A's constraints, exemplified by ChatGPT, which is characterized by post-September 2021 knowledge gaps, misinformation, comprehension lapses, and hallucinations, potentially generating spurious data. Such scenarios, exacerbated by a failure to align with children's development nuances, risk jeopardizing their well-being. This challenge becomes pronounced with AI not explicitly designed for educational purposes, especially concerning younger students. This predicament starkly contrasts with the child-centric AI principles delineated by UNESCO (Miao et al., 2020, p. 4), which has underscored child protection from AI's adverse impacts, child agency in AI system shaping, and the judicious harnessing of AI's potential benefits.

While teachers may not have direct experience with assessment in the emerging landscape of artificial intelligence in education, parallels can be drawn with their experiences during remote teaching. The transition to remote teaching, necessitated by global circumstances, required educators to swiftly adapt and integrate technology into their assessment practices. This rapid shift, much like the introduction of AI in education, presented both challenges and opportunities for educators. Experiences from the period of emergency remote teaching have indicated that while traditional assessment methods possess inherent advantages, they may not fully capture the breadth and depth of learning and teaching in technologically advanced environments. The integration of AI into assessment necessitates a re-evaluation of existing practices. The parallel, link, and interplay between distance learning and artificial intelligence are explored in numerous studies (e.g., Aljarrah et al., 2021; Tang et al., 2023). These developments call for modifications in teaching methodologies, assessment practices, and the incorporation of digital tools, emphasizing the critical role of digital literacy competencies. The shift to remote teaching provides insight into teachers' adaptability and potential readiness for the integration of AI-driven assessment methods. Recent research further supports the perspective. Kim and Kim (2022) highlighted the significance of teachers' perceptions of AI-enhanced tools in STEM education, suggesting that successful AI integration is influenced by teachers' attitudes and prior experiences. Kerneža and Zemljak (2023) posited that teachers have preconceptions about future technologies, such as humanoid robots and AI, that influence their perceptions of these technologies. Salas-Pilco, Xiao, and Hu (2022) examined the relationship between AI and learning analytics in teacher education, emphasizing the importance of understanding teachers' digital competence and their views on AI's role in teaching. Assessment in remote settings frequently involves the utilization of digital tools, many of which are intrinsically linked to artificial intelligence. Understanding how teachers employ and adapt to remote assessment techniques can serve as an indicator of their readiness and capability to integrate AI into evaluation processes. Furthermore, their receptiveness to adopting new technologies within the context of remote assessment may reflect a broader preparedness for innovations, such as the incorporation of artificial intelligence in educational practices. Given these insights, it is justified to generalize teachers' feedback on distance education to the broader context of AI-driven assessment. The shared challenges and opportunities in both areas, informed by recent research findings, offer a comprehensive understanding of teachers' preparedness for the evolving landscape of AI in education.

Research Problem

The educational landscape is dynamically evolving in real time. The OECD (2018) highlighted the unparalleled advancements in science and technology, especially in biotechnology and AI, as pressing challenges for the future of education. Among the multifaced dimensions of AI in education, the rise of ChatGPT, alongside other current and forthcoming generative pre-trained transformers and large-scale language models, draws particular attention

https://doi.org/10.33225/jbse/23.22.851

to the assessment and unique features that advanced technologies bring. Following initial restrictions on ChatGPT in specific countries, there is a growing consensus that chatbots should not be banned, but rather that students should be taught how to use them responsibly (Crawford et al., 2023; Gimpel et al., 2023). The discourse underscores the pillars of integrity, ethics, and personal responsibility, emphasizing that individuals are responsible for the quality of their work (Rudolph et al., 2023). There's also a call for discerning utilization and understanding the inherent boundaries of non-human text generators, with a spotlight on the intrinsic worth of human composition (Anson & Straume, 2022; D'Agostino, 2022; Fyfe, 2022), which remains a cornerstone of intellectual growth (Mills, 2023). Currently, there's a push for assessment methodologies that prioritize oral presentations, self-reflection, performance-based assessment, and peer assessment, all underpinned by collaborative work (Gimpel et al., 2023; Rudolph et al., 2023). The incorporation of mentorship and coaching, which segment learning into smaller pieces and provide more feedback, is perceived as beneficial (Gimpel et al., 2023). Indeed, this pivotal moment may mark a paradigm shift from conventional pedagogical approaches towards not just innovative, but deeply transformative instructional strategies.

Research Focus

In the realm of pedagogical advancements, the deployment of intricate, non-traditional assessment tools becomes imperative. Specifically, within the curriculum framework, science teachers ought to plan assignments that encourage students to think critically (Rüütmann, 2019). Furthermore, teachers should review assignments and assessments in their courses (Teaching in the AI era, 2023). The initial stride toward this objective was catalyzed by emergency remote teaching (Khan et al., 2021). The next phase should include the integration of new AI technologies. The subsequent phase should encompass the seamless integration of emergent AI technologies. Contrary to the procedural aspects of AI-driven assessment, such as automated grading as described by Gardner et al., (2021), the primary research emphasis pivots towards the evaluation of knowledge, both facilitated by and rooted in AI. The deployment of AI in this context can manifest in concealed, explicit, or intentional forms, contingent upon the educator's strategic choices. The present study draws from teacher feedback collated at the end of the emergency remote teaching phase, wherein they elucidated the assessment modalities employed during remote instruction. Distance remote teaching has indeed opened the door to unconventional forms of learning and assessment, such as scalability, research innovation, flexible learning, diversity, adaptation of assessment methods, and potential for innovation (Gurajena et al., 2021). The primary research focus was to examine teachers' readiness to embrace novel assessment methods resulting from the application of AI. It seeks to understand the challenges teachers may encounter when evaluating novel environments, including those related to reliability, fairness, and objectivity of assessment using AI.

Research Aim and Research Questions

The research aim was to explore the preparedness of science teachers for assessment in the age of AI, considering the challenges and opportunities presented by the widespread accessibility of AI technologies in education. One of the major goals of this study was to understand the specific factors contributing to teachers' readiness or their lack thereof. Furthermore, the study had an objective to discern potential differences or similarities in the preparedness of science teachers compared to teachers in other subject areas. Through these detailed objectives, the overarching aim was to offer insights into teachers' readiness for AI-based assessments, thereby guiding educational practices and policies in adapting to the evolving landscape of technology in education. The following research questions were formulated:

RQ1: Are teachers ready for assessment in the era of widespread accessibility of AI in education?
RQ2: Are there differences between science teachers and teachers of other subjects in terms of assessment in the era of widespread accessibility of AI in education?
RQ3: Is assessment in the era of broad accessibility of AI in education based on novel models that require higher levels of reading literacy?

https://doi.org/10.33225/jbse/23.22.851

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Research Methodology

General Background

In the contemporary educational landscape, the integration of artificial intelligence (AI) is reshaping learning and teaching paradigms. The broader context of this transformation underscores the importance of adapting assessment strategies to these technologically advanced environments. A review of existing literature reveals a consensus on the need for novel approaches to learning, teaching, and assessment that align with 21st-century skills. However, a notable gap emerges when juxtaposing this consensus with the current state of teacher preparedness. Many educators, despite the wealth of research advocating for change, remain inadequately equipped to navigate the intricacies of assessment in an AI-augmented setting. This disparity between the recognized educational imperatives and the actual capabilities of teachers is further accentuated with the rapid proliferation of AI in educational settings. The urgency of this matter becomes even more pronounced when considering the responses from educators during the emergency remote teaching phase. During this period, the transformative potential of AI, rather than being an integral part of the educational process, was often perceived as a distant, futuristic concept. The study draws on data collected during the period of emergency remote teaching, presenting a unique context for understanding how teachers adapt and respond to technological shifts. This specific timeframe allows us to comprehend assessment practices under exceptional circumstances, which can serve as a foundation for understanding teacher readiness and needs in an era of rapid technological advancement, such as the integration of AI. The intention is to offer insights that can shape practices, policies, and further research, ensuring that the educational community can harness the full benefits of AI while addressing its challenges.

Sample

The survey is based on a questionnaire completed by 1215 teachers from primary and secondary schools in Slovenia. The sample size of 1215 was determined based on a power analysis to ensure adequate statistical power for detecting meaningful differences and associations in the data. A power analysis was conducted to determine the appropriate sample size. The power analysis was performed considering an anticipated effect size of d = .5 (Cohen, 1988). The standard deviation of responses, obtained from a pilot study, was also considered and found to be 1.2. The calculation was carried out using the SPSS software, an aim was set to achieve a test power of .85 to reliably detect statistically significant effects, should they exist. The Type I error rate was set at .05. Based on these parameters, the power analysis indicated that a sample of at least 1050 teachers was required to reliably detect the anticipated effects. Based on this analysis, a decision was made to obtain a sample of 1215 teachers, which exceeds the minimum required sample size determined by the power analysis. Teachers were selected from different regions, including urban and rural areas, to ensure a diverse and representative sample. This stratified sampling approach aimed to capture the varied experiences and perspectives of teachers from different geographical and demographic backgrounds, allowing more robust generalization of the results. Although a non-probabilistic method was employed, teachers were randomly selected based on their accessibility and willingness to participate. This approach was chosen to maximize participation while ensuring a diverse range of respondents. The breakdown of participating teachers is as follows: 182 primary school teachers (teaching all school subjects), 268 social science teachers, 227 science teachers, and 246 vocational teachers. In terms of teaching experience, most of the participating teachers had more than 20 years of experience (50.04 %), followed by 15.97 % of teachers with 15-20 years of experience, 13.00 % of teachers with 5 years of experience or less, 12.02 % of teachers who have been teaching for 10-15 years, and 8.97 % of teachers who reported 5-10 years of experience. The varied experience levels further enhance the representativeness of the sample, capturing insights from both seasoned educators and those newer to the profession.

Instrument and Procedures

The research was conducted at the end of the 2021/2022 school year, coinciding with the conclusion of emergency remote teaching. The primary objective was to gather insights into the pedagogical strategies employed by teachers during the remote teaching phase, to develop pedagogical recommendations and guidelines in case the need for remote teaching arises again. The initial questionnaire consisted of 18 items representing dependent

https://doi.org/10.33225/jbse/23.22.851

variables and 5 items representing independent variables. To ensure the instrument's robustness, a validation study was initiated to scrutinize the questionnaire's structure and affirm its construct validity. The study aimed to validate the measures used to assess teachers' assessment in the time of emergency remote teaching. The questions were based on teachers' experiences reported in various courses they attended during the pandemic. Participants included 2 teachers from the first triennium of primary school, 2 teachers from the second triennium of primary school, 2 teachers from the third triennium of primary school, 2 teachers from a vocational high school, and 2 teachers from a grammar school. Based on the results, the questionnaire was found acceptable.

The final questionnaire for teachers was designed and published on the online platform 1ka.si. It was distributed through various forums and websites to teachers who matched the predetermined sample. The questionnaire was also distributed among teachers with a request to share the link with their colleagues. For the original purpose of drawing conclusions about remote teaching, half of the results were used, except for the question about assessment during remote teaching. However, with the emergence of AI and its broader impact on education, the question of assessment in new learning environments and under new instructional and teaching conditions has become relevant again. To understand teachers' readiness for unconventional assessment methods, answers to questions posed in these new learning environments can provide valuable insights. Due to teachers' limited awareness of AI and the practical benefits it can have for them, as shown by a survey conducted in Estonia, which ranks first among the 27 European countries in the Index of Readiness for Digital Lifelong Learning (IRDLL) (Chounta et al., 2022), a new questionnaire was deliberately not designed, as the original one also covers new teaching methods, even in the age of AI (Zimmerman, 2018). This decision was made because a similar question related to the environments with which teachers are already familiar after the COVID crisis provides sufficiently clear results and answers regarding teachers' opinions about unconventional assessment methods in the classroom. These responses are later interpreted in the context of the emerging construct of AI. The responses to the questions asked in these new learning environments may help to understand teachers' readiness for unconventional assessment methods. The surveyed question consisted of five items that provided data on assessment approaches during emergency remote teaching. Teachers answered these questions on a 4-point Likert scale (1 – never, 2 – rarely, 3 – often, 4 – mostly). They rated the frequency of assessment through videoconferencing, quizzes and tasks in online classrooms, written assessments, evaluation of seminar papers, and assessment of authentic assignments, video products, and projects during emergency remote teaching. Finally, teachers also responded to a demographic question, from which the information about their primary educational domain was obtained. The options for selection were primary school teacher, social sciences teacher, natural sciences teacher, and vocational subject teacher.

The Cronbach's alpha coefficient, with a value of .771, showed reliable internal consistency (Nunnally, 1978). The commonality data indicated the extent to which each variable contributed to the extracted factors, and all variables had appropriate values for further interpretation (oral assessment using video conferencing = .568; quizzes and tasks in online classrooms = .523; written assessment = .567; evaluation of seminar papers = .553; assessment of authentic assignments, video products, and projects = .797).

Ethical Procedures

All measurements and interventions were conducted in accordance with applicable ethical guidelines and were voluntary for teachers who chose to participate in the study. Prior to their involvement, teachers provided informed consent, granting permission for the utilization of their data for analytical purposes and subsequent publication of findings. All information collected was processed and stored in accordance with applicable data protection regulations. A special emphasis was placed on safeguarding the privacy and ensuring the anonymity of the participating educators. To this end, all collected materials maintained teacher anonymity through the deployment of encrypted identification codes, eliminating any possibility of tracing back to individual identities. Throughout the entirety of the research process, the highest ethical standards were rigorously upheld. The rights, dignity, and overall well-being of all participants were consistently prioritized and respected.

Data Analysis

Reliability analysis was conducted to examine the questionnaire structure. To measure the internal consistency of the questionnaire, Cronbach's alpha coefficient was employed. The Kruskal-Wallis analysis for independent samples was performed to test for statistically significant differences among groups of teachers based on their

https://doi.org/10.33225/jbse/23.22.851

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

primary educational domain. In the context of the Kruskal-Wallis test, the magnitude of the effect of the observed differences was calculated. Furthermore, a post hoc analysis using Dunn's test was conducted to explore specific differences between groups of teachers. Descriptive statistics were employed to obtain basic information on mean scores, standard deviations, and ranges within each assessment category. The data were analyzed using IBM SPSS software.

Research Results

In accordance with the research conducted, the general results regarding the forms of assessment during emergency remote teaching are presented, based on how frequently they were chosen by the teachers.

Table 1
Frequency of Using Different Assessment Methods during Emergency Remote Teaching

Assessment Method	*N*	*M*	*SD*
Oral assessment using video conferencing	1215	2.39	1.176
Quizzes, tasks in online classrooms	1215	2.07	1.038
Written assessment	1215	1.54	.869
Assessment of seminar papers	1215	1.95	1.012
Assessment of authentic tasks, video productions, projects	1215	2.29	1.121

The results presented (Table 1) show that during emergency remote teaching, teachers predominantly assessed students' knowledge through oral assessment via video conferencing ($M = 2.39$). The average of all responses given is also found for the frequency of assessment through authentic tasks, video productions, and projects ($M = 2.29$). Less frequently, teachers chose quizzes and tasks in online classes ($M = 2.07$) and assessment of seminar papers ($M = 1.95$) less frequently, while they least frequently chose to assess student knowledge through written assessment ($M = 1.54$). The variable of written assessment showed the least variability in teachers' responses ($SD = .869$), indicating greater consensus among teachers regarding the choice of written assessment in distance learning. The greatest dispersion of results was found for oral assessment using video conferencing ($SD = 1.176$) and assessment of authentic tasks, video productions, and projects ($SD = 1.121$), suggesting that teachers vary in how well prepared they are to assess in this way. However, the variance is still relatively small, suggesting some degree of consistency in assessment using the above methods. Slightly, but not significantly, different scores were observed for quizzes and tasks in online classrooms ($SD = 1.038$) and assessment of seminar papers ($SD = 1.012$), where the teachers' scores were somewhat more similar.

Table 2
Assessment Methods during Emergency Remote Teaching according to Their Primary Field of Teaching

	Primary school teacher		Social sciences teacher		Natural sciences teacher		Vocational subject teacher	
	N	*M*	*N*	*M*	*N*	*M*	*N*	*M*
Oral	182	412.94	368	521.75	418	462.07	247	445.43
Quizzes	182	425.81	368	438.42	418	510.22	247	507.21
Written	182	364.60	368	465.78	418	487.32	247	515.57
Seminar	182	295.25	368	472.69	418	465.79	247	591.93
Authentic	182	406.52	368	484.33	418	441.34	247	507.60

Note. Oral – Oral assessment using video conferencing. *Quizzes* – Quizzes, tasks in online classrooms. *Written* – Written assessment. *Seminar* – Assessment of seminar papers. *Authentic* – Assessment of authentic tasks, video productions, projects.

https://doi.org/10.33225/jbse/23.22.851

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

The results presented in Table 2 show statistically significant differences in the choice of assessment methods among teachers based on their teaching area ($H(3) = 23.249$, $p = .001$, $\eta^2 = .02$). Video conferencing assessment was most frequently used by social science teachers ($M = 521.75$), which is consistent with their preference for interactive teaching methods and discussions with students. Science teachers ($M = 462.07$) and vocational subject teachers ($M = 445.43$), on the other hand, used it less frequently, possibly due to the practical nature of their subjects, which is more difficult to assess through videoconferencing. Primary school teachers ($M = 412.94$) were the least likely to use it, likely due to the challenges of providing appropriate technology and addressing the developmental needs of younger students. The small effect size ($\eta^2 = 0.02$) does not detract from the statistical significance of the differences found ($p = .001$). The teaching domain is only one of potentially many influencing factors, which means that further research is needed.

Based on the results of the Kruskal-Wallis test ($H(3)$: 19.642, $p = .001$, $\eta^2 = .02$), significant differences were found between at least two groups of teachers in their use of quizzes and assessment tasks. Teachers of science subjects ($M = 510.22$) and teachers of vocational subjects ($M = 507.21$) most frequently assessed their students using quizzes and tasks, which could be attributed to the interactive and practical nature of their subjects that lend themselves these assessment methods. Social science teachers ($M = 438.42$) and primary school teachers ($M = 425.81$) used quizzes and tasks less frequently, which may be due to the emphasis on conceptual understanding in social science subjects and the use of alternative assessment methods for primary school students to gauge comprehensive understanding. Despite the small effect size ($\eta^2 = 0.02$), the statistical significance of the observed differences ($p = .001$) remains unchanged. The pedagogical domain is only one of many possible elements influencing these differences, highlighting the need for further research.

Statistically significant differences in the frequency of written assessment ($H(3)$: 45.005, $p = .001$, $\eta^2 = .04$) were found between the different groups of teachers. Teachers of vocational subjects ($M = 515.57$) chose written assessment most frequently, probably because it is suitable for assessing skills and subject knowledge. Natural science teachers ($M = 487.32$) and social science teachers ($M = 465.75$) used written assessment less frequently because their subjects focus primarily on theoretical concepts that may be better assessed using other methods. Primary school teachers ($M = 364.60$) were least likely to use written assessment, possibly due to the challenges of developing writing and reading skills in younger students, leading them to choose more appropriate assessment approaches for their developmental stage. The observed effect size ($\eta^2 = .04$) means that although the differences are statistically significant, the variance in the frequency of writing assessments between the different teacher groups is only marginally explained by the instructional domain. This suggests that other factors may also play a significant role in teachers' choice of assessment method.

The analysis also revealed statistically significant differences in assessment method choices based on teaching domains ($H(3) = 124.784$, $p = .001$, $\eta^2 = .10$). Vocational subject teachers ($M = 591.93$) most frequently used seminar paper assessment, reflecting the specificity of vocational education, in which seminars are a common way of assessing practical skills and knowledge related to the chosen profession. Social science teachers ($M = 472.69$) and natural science teachers ($M = 465.79$) used this method less frequently, possibly reflecting the nature of their subjects, which require other appropriate assessment methods. Primary school teachers ($M = 295.25$) used seminar assessments least frequently, which may be related to the adaptation of assessment methods to the age group and developmental characteristics of younger students at this educational level. The relatively large effect size ($\eta^2 = 0.10$) in this analysis suggests that 10% of the variance in the frequency of seminar paper assessments among the different groups of teachers can be attributed to their teaching domain. This suggests that the teaching domain has a considerable influence on the choice of this assessment method. Despite the observed differences, factors other than the teaching area could also affect teachers' assessment preferences.

Similarly, significant differences were observed in the use of assessment of authentic tasks, video productions, and projects were observed ($H(3)$: 17.503, $p = .001$, $\eta^2 = .02$). Vocational subject teachers ($M = 507.60$) frequently emphasized practical skills, best assessed through authentic tasks, video productions, and projects. Social science teachers ($M = 484.33$) also frequently used this method, possibly to encourage critical thinking and idea expression. Natural science teachers ($M = 441.34$) used it less frequently, possibly due to a greater emphasis on concrete scientific concepts. Primary school teachers ($M = 406.52$) used this assessment method least often, likely due to the constraints of conducting such assessment formats with younger students. The effect size suggests that the teaching domain accounts for 2 % of the variance in the use of authentic tasks, video productions, and projects for assessment, suggesting that other factors contribute significantly to this choice and thus further research is needed.

A Dunn's post-hoc test aimed to compare the assessment practices of natural science with those from other

https://doi.org/10.33225/jbse/23.22.851

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

fields in new educational environments. The results suggest that natural science teachers' assessment methods in these contexts do not significantly differ from other teachers. However, some notable differences emerged in specific assessment types. Regarding oral assessment, a statistically significant difference was found between natural science teachers and social science teachers ($p = .008$, $M_{natural\ science} = 2.35$, $M_{social\ science} = 2.62$). Natural science teachers tended to use oral assessment less frequently compared to their counterparts in social science. In the case of quizzes, significant differences were observed between natural science teachers and primary school teachers ($p = .009$, $M_{natural\ science} = 2.26$, $M_{primary\ school} = 1.91$) as well as social science teachers ($p = .007$, $M_{natural\ science} = 2.26$, $M_{social\ science} = 1.96$). Regarding written assessment, statistically significant differences were seen in comparison with primary school teachers ($p = .001$, $M_{natural\ science} = 1.63$, $M_{primary\ school} = 1.17$). Natural science teacher teachers tended to use written assessments more frequently than primary school teachers. For the assessment of seminar papers, differences were found between natural science teachers and primary school teachers ($p = .001$, $M_{natural\ science} = 1.94$, $M_{primary\ school} = 1.29$), as well as vocational subject teachers ($p = .001$, $M_{natural\ science} = 1.94$, $M_{primary\ school} = 1.24$). Natural science teachers used assessment of seminar papers more frequently compared to primary school and vocational subject teachers. However, no statistically significant differences were observed in the use of authentic tasks among natural science teachers and teachers from other fields. In summary, while natural science teachers did not significantly differ from teachers in other fields in terms of general assessment practices in new environments, they showed variations in specific assessment methods like oral assessments, quizzes, written assessments, and assessment of seminar papers when compared to their colleagues in primary school and social science, as well as vocational subject teachers.

Discussion

The purpose of this study is to examine science teachers' preparedness for assessment in the age of AI, considering the challenges and opportunities presented by the ubiquitous presence of AI technologies in education (Cotton et al., 2023). Although the results were analyzed from the perspective of self-assessment during periods of emergency remote teaching, the findings suggest that factors beyond the teaching domain also play significant roles in determining teachers' choice of assessment methods. These conclusions are further discussed through the lens of assessment in the age of AI.

Based on the theoretical framework of the research conducted, it is hypothesized that teachers, in general, are not adequately prepared for assessment in the era of widespread accessibility of AI. In line with the theoretical underpinnings of the study, which highlights the need for new forms of knowledge, new forms of learning, new forms of teaching, and consequently new forms of assessment, it was expected that teachers would predominantly report choosing conventional assessment methods during emergency remote teaching. This period opened the door to computer-based learning and new forms of assessment in schools (Khan et al., 2023). Results show that teachers most often chose to use oral assessment via video conferencing during distance assessment. This method was instrumental in curbing potential plagiarism and other forms of cheating that students might employ in assessing their curriculum goals. Oral assessment is considered one of the fundamental forms of conventional assessment, alongside with written assessment, which was used least frequently during remote teaching due to concerns of plagiarism and cheating. Oral assessments can be successful and necessary in contemporary learning environments when they promote self-reflection, performance-based assessment, and peer assessment (Crawford et al., 2023; Gimpel et al., 2023). To a slightly lesser extent, but still infrequently, teachers chose to assess authentic tasks, video productions, and projects, which are forms of assessment anticipated to be foundational in future assessments. Teachers were somewhat less inclined to assess seminar assignments, which became problematic in the age of AI because students could complete the entire assignment using AI. Teachers' responses indicate that they are not fully prepared for assessment in the age of widespread AI in education, underscoring the need for enhanced teacher preparation for AI use in assessment. Adequate training and support for teachers are essential to help them navigate the challenges and harness the opportunities that AI technology presents in education, as it becomes evident that they are not yet ready for it (Kerneža et al., 2023). Adapting educational practices and policies to the evolving technological landscape of education triggered by distance learning (Gurajena et al., 2021) is crucial for achieving teacher readiness for assessment in the age of AI.

The results indicate that differences exist in the assessment approaches adopted by teachers during remote education, contingent on their primary area of instruction. Oral assessment via videoconferencing was most frequently chosen by social science teachers, less so by natural science and vocational teachers, and least by primary school teachers. The predilection for more interaction and discussion in social science subjects might explain this

https://doi.org/10.33225/jbse/23.22.851

trend, while natural science subjects entail more hands-on demonstrations or experiments, challenging to execute in a virtual milieu. Quizzes and assignments were predominantly utilized by natural science and vocational teachers to assess student knowledge, and least by social science and primary school teachers. Natural science and vocational teachers appear to emphasize the practical application of knowledge, assessable through assignments and quizzes, whereas social science and primary school teachers seem to prioritize conceptual understanding, assessable through alternative assessment forms. Written assessments were predominantly chosen by vocational teachers, less so by natural science and social science teachers, and least by primary school teachers. The variance in the preference for written assessment among different subject teachers could stem from the intrinsic nature of the subject and the imperative for written articulation of knowledge. Vocational subjects might necessitate extensive writing and formulation, aligning with conventional assessment methods (Montgomery, 2022) pertinent to skills typically not computer based (Swiecky et al., 2022). In contrast, social science subjects might emphasize argumentation and essay writing. Science problems are more easily assessed using alternative methods, potentially explaining why natural science teachers infrequently opt for written assessment. Assessment of seminar papers was most prevalent in vocational subjects, less so in social studies and natural science subjects, and least in primary school subjects. Vocational students might develop seminar papers emphasizing practical examples and real-world knowledge applications, while social science subjects might focus on data presentation, and natural science subjects on experimental work. Conversely, primary school students, especially in the early years, might not possess the requisite skills for independent seminar paper production. Authentic tasks, video productions, and projects were predominantly chosen by vocational and social science teachers, less frequently so by natural science teachers, and least by primary school teachers. This aligns with the practical and analytical skills students demonstrate in such assignments. Natural science subjects might necessitate experimental work or the creation/presentation of a scientific research project. Science teaching involves intricate concepts that students encounter and explore in school, with optimal teaching and learning defined by higher taxonomic levels (CAST, 2018). Primary school teachers, who also teach science, were observed to seldom use authentic tasks, video productions, and projects. Given the younger age of these students, it is not anticipated that these tasks and related skills are as developed as in older students. However, it becomes imperative to ascertain whether teachers are fostering basic digital literacy skills in students, preparing them for advanced assessment at the secondary level. Primary education is a pivotal juncture in establishing the foundation for teaching contemporary 21st-century learning environments (Kerneža & Kordigel Aberšek, 2023; Kordigel Aberšek & Kerneža, 2023).

Considering the research findings that address potential disparities between science teachers and educators from other disciplines concerning assessment practices in the age of pervasive AI accessibility and the unparalleled innovations in science and technology, as highlighted by OECD (2018), some differences were noted in the selection of assessment methods based on the educators' primary instructional domain. Social science teachers predominantly opted for oral assessments via videoconferencing, suggesting an amplified necessity for interaction and discourse within social science disciplines. In contrast, science and vocational subject teachers frequently employed quizzes and assignments as assessment tools, potentially reflecting the accentuation of knowledge's practical application within this domain. Vocational teachers exhibited a marked preference for written assessments, whereas their counterparts in the social sciences appeared to prioritize argumentative and essay-based evaluation. Science teachers demonstrated a reduced inclination towards written assessment, possibly due to the efficacy of alternative methods in evaluating scientific queries. The assessment of seminar papers was predominantly observed within the vocational subjects, facilitating students in showcasing practical knowledge applications in real-world context. Primary school teachers exhibited a reduced propensity for this assessment form, attributing this to the perceived underdevelopment of independent seminar paper production skills among primary students. Authentic tasks, video productions, and projects were predominantly favored by vocational and social science educators, mirroring the practical and analytical proficiencies students manifest in these tasks. Science teachers displayed a diminished preference for these assessment methods, potentially due to the inherent requirements of experimental undertakings or the formulation of scientific research projects, which might be deemed unsuitable for primary school students lacking the requisite competencies for autonomous task execution.

The findings from the research highlight the varying degrees of teacher readiness and their inclinations towards certain assessment methods during remote teaching. These can be interpreted through the lens of potential receptiveness to AI-driven innovations. Such receptiveness is not merely a reflection of adaptability to new technologies but also an indication of pedagogical flexibility and willingness to evolve in response to the changing educational landscape. Insights from the research phase suggest that the experiences and feedback of

https://doi.org/10.33225/jbse/23.22.851

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

teachers during distance education can provide valuable context when considering the broader implications of AI-driven assessment. For instance, the challenges faced, the solutions devised, and the overall sentiment towards technology-mediated teaching can serve as indicators of how educators might approach and integrate AI tools in their teaching and assessment practices. Moreover, the nuances of experiences during remote teaching, such as preferences for certain assessment methods or reservations about others, can offer clues about potential areas of comfort or concern when it comes to AI-driven assessment. This perspective is further supported by recent research, including the works of Salas-Pilco, Xiao, and Hu (2022). Their findings emphasize the multifaceted nature of teacher readiness, suggesting that it's not just about technological proficiency but also about pedagogical understanding, attitude towards innovation, and the ability to foresee the potential challenges and benefits of integrating AI. Collectively, these insights paint a comprehensive picture of the evolving landscape of AI in education, emphasizing the intricate interplay between technology, pedagogy, and the requisite preparedness of educators.

The data highlight the imperative for literacy skills that are essential for participation in various assessment activities, encompassing oral evaluations, reading assignments, essay writing, or comprehension of instructions. Teachers must recognize the significance of fostering students' literacy skills and catering to their diverse reading and writing requirements within the assessment context. Such an approach not only renders learning meaningful but also proffers students with significant and rigorous learning prospects, as underscored by CAST (2018). This is especially pronounced in primary schools, where the same teacher assesses both science and literacy, given the low reading scores shown in the new PIRLS data analysis (Mullis et al., 2023). This ensures equitable and quality knowledge assessment in the AI epoch, wherein the widespread availability of AI technologies in education represents merely one facet of the myriad opportunities and challenges confronting teachers. The connection between literacy and science instruction is crucial, as evidenced by other studies (e.g., Kim et al., 2021; Pearson et al., 2010). Grasping content and critical interpreting of scientific texts are vital for students' academic success in science, given their exposure to a plethora of scientific texts that often contain complex concepts, scientific language, and specific terminology. Proficiently navigating scientific literature is paramount for information assimilation, key idea identification, detail discernment, and holistic subject comprehension. Within contemporary assessment frameworks, students must be able to evaluate the credibility of sources, recognize scientific bias, analyze, and evaluate arguments, and identify possible errors or gaps in research. This requires developed critical thinking supported by advanced reading and writing skills. Following the process of data collection, students should articulate their perspectives coherently and systematically, defending their viewpoints and discoveries. Written expression skills include appropriate use of scientific language, text organization, logical connections, and use of appropriate scientific and technical terms. In summary, the teaching and assessment of science in the age of AI encompasses the full spectrum of both literacy and scientific content. In a modern science classroom supported by AI, the development and support of literacy skills are crucial. Students must have the opportunity to develop comprehension, critical reading, summarizing, and analytical thinking skills in the specific context of science content. This prepares them with the skills essential for active participation in assessment activities demanding the interpretation and application of scientific data. Students should instruct students to critically evaluate sources and information they obtain from the Internet, which should also be given more attention by teachers (Zemljak & Kerneža, 2023). A potential strategy is delineated by Leu et al., (2008) with the three-phase model of online reading instruction called Internet Reciprocal Teaching (IRT), resonating with the tenets of problem-based learning (Zemljak et al., 2023). This approach facilitates student collaboration, fostering critical thinking and problem-solving skills, with an emphasis on the learning trajectory rather than mere outcomes. Students need empowerment within the learning paradigm, given that the learning process itself constitutes a pivotal contribution to pragmatic life knowledge- Thus, education can be perceived as a life experience, encompassing attitudes, knowledge, and skills, tailored *for life-encompassing* behavior, cognition, and consideration (Broks, 2023).

Children today live in a very different world than their parents (Siraj, 2017). Not only do today's children not know the time before smart devices; they are also the first generation whose lives are in one way or another defined by AI-enabled applications and devices, while also being exposed to AI-related risks (UNICEF, 2020). The research provides insights into the need for training and support for teachers in the use of AI in assessment, but further study is needed to provide a more concrete and reliable picture of teacher readiness in this area. If Turetzky et al., (2019) were already calling for AI researchers to become AI educators in 2019 and create resources to help teachers and students understand AI intelligence, this is even more important in 2023. Kerneža (2023) noted that pre-service teachers frequently overestimated the competencies essential for interpreting chatbot-generated content, a perception that often diverged from evaluators' assessments. Such disparities might have led to potential

https://doi.org/10.33225/jbse/23.22.851

misjudgments of their capabilities or suboptimal evaluations, potentially curtailing their progression opportunities. This reinforces the findings outlined by Farell (2007) that teachers need to continually revise their knowledge of teaching and learning to teach effectively and competently in the rapidly changing field of education. With additional research and analysis focused directly on AI's role in assessment, more definitive conclusions can be drawn regarding teacher readiness in this sphere, facilitating the proposition of suitable strategies and interventions to recalibrate educational practices. Our responsibility as individuals is to ensure the ethical development and application of AI (Dignum, 2019). Therefore, as UNESCO (Miao et al., 2020) emphasizes it is particularly important, that AI supports and promotes children's growth. More than ever, education must become more systemic through a balance between technological and humanitarian education. Today, we need to find better ways to combine our physical and spiritual pursuits, which is the fulfillment of a truly sustainable (balanced, long-term) development of our way of life and our future education (Broks, 2016). Furthermore, it is important to emphasize that, as UNESCO (Miao et al., 2020, p. 39) point out, the teacher must support children's development and well-being; ensure inclusion of and for children; prioritize fairness and nondiscrimination for children; protect children's data and privacy; ensure children's safety; provide transparency, explainability, and accountability for children; prepare children for current and future developments in AI; equip governments and businesses with knowledge about AI and children's rights; and create an enabling environment for child-centred AI. This can only be achieved through an AI-competent teacher.

The uniqueness of this study lies in its emphasis on the readiness of science teachers for assessment in the age of AI. While there have been studies on the use of AI in education and its impact on student learning, this research specifically addresses teachers' readiness to use AI-based assessment methods. By examining teachers' perspectives and practices, it sheds light on the current state of their readiness and highlights areas that need further attention and support. This focus on teacher readiness in the context of AI assessment sets this study apart from others. The implications of the study also extend to the broader educational community, highlighting the need for professional development programs, subject-specific adaptations of AI-based assessment, and the integration of literacy into assessment practices. By considering these implications, the findings can contribute to the effective and ethical implementation of AI in education, ultimately benefiting student learning outcomes.

The limitations of the study are also considered in the analysis of the results and discussion. A non-probability sample was used for the study, which may lead to selective participation and potential bias, as only those teachers who are more interested in the topic or have experience may have chosen to participate. Therefore, the generalizability of the results should be taken with caution, as they may be more representative of this group of teachers. With respect to the administration of the questionnaire, it is also important to consider the possibility that teachers may have provided subjective responses that are consistent with socially desirable responses or idealized perceptions of their work. There are also concerns about the accuracy of self-reflection. In addition, it is important to note that the study was conducted during a time when teaching occurred remotely, which may have influenced teachers' perceptions, instructional decisions, and willingness to adopt new assessment methods. Therefore, the application of the study's findings under "normal" circumstances should be carefully considered.

Conclusions and Implications

This research highlights the need for better preparation and support of teachers in the use of AI in assessment. Despite the challenges and opportunities AI presents in education, teachers predominantly chose conventional assessment methods during emergency remote teaching. This approach helped address concerns about plagiarism and cheating. However, the assessment methods teachers chose varied depending on their primary subject. Certain assessment methods were preferred across subjects, such as quizzes and assignments in natural science and vocational subjects and written assessments in vocational subjects. Primary school teachers were less likely to use certain forms of assessment because of the age and readiness of their students. The study also highlights the importance of literacy skills and their connection to science education.

The findings suggest that teachers are not fully prepared for assessment in the era of widespread availability of AI technologies. Understanding teacher readiness for assessment using AI is critical to the successful adaptation of this technology in education. Only with adequately trained teachers can we ensure that AI serves as a tool to enhance the learning process and not as a substitute for teacher expertise. Professional development programs, subject-specific adaptations of AI-based assessment, and the integration of literacy into assessment practices are essential to ensure the effective implementation of AI in education. To further this goal, it is recommended that

https://doi.org/10.33225/jbse/23.22.851

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

future efforts be directed towards the development of comprehensive educational programs and materials. For the organization of the educational process, it is recommended that ongoing professional development programs focused on the use of artificial intelligence in assessment are introduced. Additionally, it is suggested that additional resources and tools are provided for teachers to assist them in adapting to new technologies and assessment methods. In the context for planning further research, it is proposed that studies be conducted to directly assess specific teacher skills reflecting their readiness to work in the reality of artificial intelligence. Such studies would allow for a more precise understanding of where the exact shortcomings lie and how best to address them. This will ensure that teachers are not only introduced to AI but are also equipped to utilize it effectively in their teaching practices.

As we stand at the brink of an educational shift powered by artificial intelligence, the readiness of our teachers is the key to its success. The interplay between traditional teaching methods and emerging AI technologies is intricate, and it's up to us to ensure this collaboration is seamless. The future of education, shaped by AI, will be as strong as the foundation we lay today. Thus, investing in our teachers' preparedness is not just a necessity but a commitment to a brighter, more informed future.

Acknowledgements

The authors want to thank the research program P5-0433; Digital Restructuring of Deficit Occupations for Society 5.0 (Industry 4.0). The research program is financed by the Slovenian Research Agency (ARRS).

Declaration of Interest

The authors declare no competing interest.

References

Aljarrah, A., Ababneh, M., Karagozlu, D., & Özdamlı, F. (2021). Artificial intelligence techniques for distance education: A systematic literature review. *TEM Journal, 9*(4), 1621–1629. https://doi.org/10.18421/TEM104-18

Anson, C. M., & Straume, I. (2022). Amazement and trepidation: Implications of AI-based natural language production for the teaching of writing. *Journal of Academic Writing, 12*(1), 1–9. https://doi.org/10.18552/joaw.v12i1.820

Bhati, D. (2023, April 5). Viral AI chatbot ChatGPT is banned in many countries, but why? Full list of countries. *India Today*. https://www.indiatoday.in/technology/news/story/viral-ai-chatbot-chatgpt-is-banned-in-many-countries-but-why-full-list-of-countries-2355938-2023-04-05

Broks, A. (2020). General remarks on basic actualities within our life and education during the first 20 years of 21st century. *Journal of Baltic Science Education, 19*(5), 692–695. https://doi.org/10.33225/jbse/20.19.692

Broks, A. (2016). Thinking about thinking: Start-up for modern education for tomorrow. *Gamtamokslinis Ugdymas / Natural Science Education, 13*(3), 94–95. https://www.doi.org/10.48127/gu-nse/16.13.94

CAST (2018). *Universal Design for Learning Guidelines version 2.2.* http://udlguidelines.cast.org

Chounta, I. A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education, 32*(3), 725–755. https://doi.org/10.1007/s40593-021-00243-5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.

Cotton, D. R. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International* 1–12. https://doi.org/10.1080/14703297.2023.2190148

Crawford, J., Cowling, M., & Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice, 20*(3). https://doi.org/10.53761/1.20.3.02

D'Agostino, S. (2022, October 26). Machines can craft essays. How should writing be taught now? *Inside Higher Education.* https://www.insidehighered.com/news/2022/10/26/machines-can-craft-essays-how-should-writing-be-taught-now

Dignum, V. (2019). What Is artificial intelligence? In *Responsible Artificial Intelligence: How to develop and use AI in a responsible way.* Springer International Publishing. https://doi.org/10.1007/978-3-030-30371-6

Dunn, L., Morgan, C., O'Reilly, M., & Parry, S. (2004). *The student assessment handbook. New directions in traditional and online assessment*. RoutledgeFalmer.

European Commission. (2020). *White paper on artificial intelligence: a European approach to excellence and trust*. https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

European Commission, Directorate-General for Education, Youth, Sport and Culture. (2022). *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators.* Publications Office of the European Union. https://data.europa.eu/doi/10.2766/153756

Farrell, T. S. C. (2007). *Reflective language teaching: From research to practice*. Continuum.

Frey, B. B., & Schmitt, V. L. (2010). Teachers' classroom assessment practices. *Middle Grades Research Journal, 5*(3), 107–117.

https://doi.org/10.33225/jbse/23.22.851

Fyfe, P. (2022). How to cheat on your final paper: Assigning AI for student writing. *AI & Society*, 2022, 1–11. https://doi.org/10.1007/s00146-022-01397-z

Galindo, L., Perset, K., & Sheeka, F. (2021). An overview of national AI strategies and policies. *Going Digital Toolkit Note*, No. 14. https://goingdigital.oecd.org/data/notes/No14_ToolkitNote_AIStrategies.pdf

Gardner, J., O'Leary, M., & Yuan, L. (2021). Artificial intelligence in educational assessment: 'Breakthrough? Or buncombe and ballyhoo? Journal of Computer Assisted Learning, 37(5), 1207–1216. https://doi.org/10.1111/jcal.12577

Gimpel, H., Hall, K., Decker, S., Eymann, T., Lämmermann, L., Mädche, A., Röglinger, R., Ruiner, C., Schoch, M., Schoop, M., Urbach, N., & Vandirk, S. (2023). *Unlocking the power of generative AI models and systems such as GPT-4 and ChatGPT for higher education: A guide for students and lecturers*. University of Hohenheim.

Grion, V., Serbati, A., & Nicol, D. (2018). Technology as a support to traditional assessment practices. *Italian Journal of Educational Technology, 26*(3), 3–5. https://doi.org/10.17471/2499-4324/1082

Gurajena, V., Mbunge, E., & Fashoto, S. (2021). Teaching and learning in the new normal: Opportunities and challenges of distance learning amid COVID-19 Pandemic. *International Journal of Education and Teaching, 1*(2), 9–15. https://doi.org/10.51483/IJEDT.1.2.2021.9-15

Harlen, W. (2007). Criteria for evaluation systems for student assessment. *Studies in Educational Evaluation, 33*(1), 15–28. https://doi.org/10.1016/j.stueduc.2007.01.003

Kerneža, M. (2023). Fundamental and basic cognitive skills required for teachers to effectively use Chatbots in education. In V. Lamanauskas (Ed.), *Science and technology education: New developments and innovations. Proceedings of the 5th International Baltic Symposium on Science and Technology Education (BalticSTE2023)* (pp. 99–110). Scientia Socialis Press. https://doi.org/10.33225/BalticSTE/2023.99

Kerneža, M., & Kordigel Aberšek, M. (2023). Specifics of digital literacy development during the reading process in digital learning environments in primary school. In L. Gómez Chova, C. González Martínez, & J. Lees (Eds.), *INTED 2023: conference proceedings: 17th annual International Technology, Education and Development Conference: 6-8 March 2023, Valencia (Spain)* (pp. 4877–4883). IATED Academy. https://doi.org/10.21125/inted.2023.1271

Kerneža, M., & Zemljak, D. (2023). Teachers' perspectives on the use of humanoid robots at primary and secondary education level, with a focus on native language teachers. In L. Gómez Chova, C. González Martínez, & J. Lees (Eds.), *EDULEARN 2023: conference proceedings: 15th International Conference on Education and New Learning Technologies: 3-5 July, 2023, Palma (Spain)* (pp. 8504–8511). IATED Academy. https://doi.org/10.21125/edulearn.2023.2235

Khan, Z. R., Sivasubramaniam, S., Anand, P., & Hysaj, A. (2021). E-thinking teaching and assessment to uphold academic integrity: lessons learned from emergency distance teaching. *International Journal for Educational Integrity,* Article 17. https://doi.org/10.1007/s40979-021-00079-5

Kim, J. S., Burkhauser, M. A., Mesite, L. M., Asher, C. A., Relyea, J. E., Fitzgerald, J., & Elmore, J. (2021). Improving reading comprehension, science domain knowledge, and reading engagement through a first-grade content literacy intervention. *Journal of Educational Psychology, 113*(1), 3–26. https://doi.org/10.1037/edu0000465

Kim, N. J., & Kim, M. K. (2022). Teachers' perceptions of Using an Artificial Intelligence-Based Educational Tool for Scientific Writing. *Frontiers in Education, 7*, Article 755914. https://doi.org/10.3389/feduc.2022.971287

Kordigel Aberšek, M., & Kerneža, M. (2023). Age conditioned online research and comprehension skills in primary school students. In L. Gómez Chova, C. González Martínez, & J. Lees (Eds.), *INTED 2023: conference proceedings: 17th annual International Technology, Education and Development Conference: 6-8 March 2023, Valencia (Spain)* (pp. 5496–5504). IATED Academy. https://doi.org/10.21125/inted.2023.1271

Legvart, P., Kordigel Aberšek, M., & Kerneža, M. (2022). Developing communication competence in digital learning environments for primary science students. *Journal of Baltic Science Education, 21*(5), 836–848. https://doi.org/10.33225/jbse/22.21.836

Leu, D. J., Coiro, J., Castek, J., Hartman, D. K., Henry, L. S., & Reinking, D. (2008). Research on instruction and assessment of the new literacies of online reading comprehension. In C. C. Block, S. Parris, & P. Afflerbach (Eds.), *Comprehension instruction: Research-based best practices* (pp. 321–346). Guilford Press.

McCallum, S. (2023, April 1). ChatGPT banned in Italy over privacy concerns. *BBC News*. https://www.bbc.com/news/technology-65139406

McMurtrie, B. (2022, December 13). AI and the future of undergraduate writing. *The Chronicle of Higher Education*. https://www.chronicle.com/article/ai-and-the-future-of-undergraduate-writing

McMurtrie, B. (2023, January 5). Teaching: Will ChatGPT change the way you teach? *The Chronicle of Higher Education*. https://www.chronicle.com/newsletter/teaching/2023-01-05

Miao, F., Homes, W., Ronghuai, H., & Hui, Z. (2021). *AI and education: guidance for policy-makers*. UNESCO. https://www.unesco.org/en/digital-education/artificial-intelligence

Mills, A. (2023). ChatGPT just got better. What does that mean for our writing assignments? *Chronicle of Higher Education.* https://www.chronicle.com/article/chatgpt-just-got-better-what-does-that-mean-for-our-writing-assignments

Montgomery, K. (2002). Authentic tasks and rubrics: going beyond traditional assessments in college teaching. *College Teaching, 50*(1), 34–40. https://doi.org/10.1080/87567550209595870

Nunnally, J. C. (1978). *Psychometric theory*. McGraw Hill.

OECD. (2018). *The future of education and skills. Education 2030.* OECD Publishing. https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf

OECD. (2021). *OECD Digital education outlook 2021. Publishing the frontiers with AI, blockchain, and robots.* OECD Publishing. https://doi.org/10.1787/589b283f-en

https://doi.org/10.33225/jbse/23.22.851

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

OECD. (2022). *Recommendation of the Council on Artificial Intelligence*. OECD Publishing. https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449

OECD. (2023). *Is education losing the race with technology? AI's progress in maths reading*. OECD Publishing. https://doi.org/10.1787/73105f99-en

Mullis, I. V. S., von Davier, M., Foy, P. ,Fishbein, B., Reynolds, K. A., & Wry, E. (2023). *PIRLS 2021. International results in reading.* Boston College, TIMSS & PIRLS International Study Center. https://doi.org/10.6017/lse.tpisc.tr2103.kb5342

Pearson, P. D., Moje, E., & Greenleaf, C. (2010). Literacy and science: Each in the service of the other. *Science, 382*(5977), 459–463. https://doi.org/10.1126/science.1182595

Ramachandran, A., & Scassellati, B. (2014). *Adapting Difficulty Levels in Personalized Robot-Child Tutoring Interactions*. https://scazlab.yale.edu/sites/default/files/files/Ramachandran_AAAI14Workshop.pdf

Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching, 6*(1), 409–413. https://doi.org/10.37074/jalt.2023.6.1.2

Rüütmann, T. (2019). Development of Critical Thinking and Reflection. In: M. Auer, T. Tsiatsos (Eds.), *The Challenges of the Digital Transformation in Education. ICL 2018. Advances in Intelligent Systems and Computing, vol. 917*. Springer International Publishing. https://doi.org/10.1007/978-3-030-11935-5_85

Salas-Pilco, S. Z., Xiao, K., & Hu, X. (2022). Artificial intelligence and learning analytics in teacher education: A systematic review. *Education Sciences, 12*(8), 569. https://doi.org/10.3390/educsci12080569

Seldon, A., & Abidoye, O. (2018). *The fourth education revolution: will artificial intelligence liberate or infantilize humanity*? The University of Buckingham Press.

Siraj, I. (2017, November 1–2). Teaching kids 21st century skills early will prepare them for their future. *The Conversation*. https://ro.uow.edu.au/cgi/viewcontent.cgi?article=4276&context=sspapers

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence, 3*. https://doi.org/10.1016/j.caeai.2022.100075

Tang, K. Y., Chang, C. Y., & Hwang, G. J. (2023). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments, 31*(4), 2134–2152. https://doi.org/10.1080/10494820.2021.1875001

Teaching in the AI era. (2023, March 6). *Stanford graduate school of business*. Teaching and learning hub. https://tlhub.stanford.edu/wp-content/uploads/2023/04/TLH_Teaching-in-the-AI-Era-Handout.pdf

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K-12: What should every child know about AI? In P. van Hentenryck, & Z. H. Zhou (Eds.), Proceedings *of the AAAI-19-Thirty-Third AAAI Conference on Artificial Intelligence*, *33*(1), 9795-9799. https://doi.org/10.1609/aaai.v33i01.33019795

Turan-Güntepe, E., Durmuş, T., & Dönmez-Usta, N. (2023). Assessment of distance learning practices during the COVID-19 pandemic in grades K-12. *Athens Journal of Education, 10*(2), 249–270. https://doi.org/10.30958/aje.10-2-4

UNICEF. (2020). *Policy guidance on AI for children.* https://www.unicef.org/globalinsight/media/2356/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.p

Yang, M. (2023, January 6). New York City schools AI chatbot that writes essays and answers prompts. *The Guardian*. https://www.theguardian.com/us-news/2023/jan/06/new-york-city-schools-ban-ai-chatbot-chatgpt

Zemljak, D., & Kerneža, M. (2023). Povezava med recipročnim učenjem s spleta (RUS) in problemskim učenjem [The connection between Internet reciprocal teaching (IRT) and problem-based learning. *Dianoia*, *7*(1), 9–18. https://www.fnm.um.si/wp-content/uploads/2023/04/Dianoia_2023_1.pdf

Zimmerman, M. (2018). *Teaching AI: Exploring new frontiers for learning*. International Society for Technology in Education.

Received: *August 01, 2023* Revised: *August 25, 2023* Accepted: *October 03, 2023*

Cite as: Kerneža, M., & Zemljak, D. (2023). Science teachers' approach to contemporary assessment with a reading literacy emphasis. *Journal of Baltic Science Education*, *22*(5), 851-864. https://doi.org/10.33225/jbse/23.22.851

Maja Kerneža
(Corresponding author)
PhD, Teaching Assistant, Faculty of Education, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia.
E-mail: maja.kerneza1@um.si
ORCID: https://orcid.org/0000-0002-0813-8675

Dejan Zemljak
Teaching Assistant, Faculty of Mathematics and Natural Sciences, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia.
E-mail: dejan.zemljak1@um.si
ORCID: https://orcid.org/0000-0002-7757-5457

https://doi.org/10.33225/jbse/23.22.851