Cumhuriyet International Journal of Education, 15(1): 313-337, 2026
DOI: https://dx.doi.org/10.30703/cije.1758405

Cumhuriyet International Journal of Education

| cije.cumhuriyet.edu.tr | Founded: 2011 | ISSN: 2147-1606 | Publisher: Sivas Cumhuriyet University

Pedagogical, Ethical, and Psychometric Orientations of Digital Assessment Applications within a Conceptual Framework

İsmail Demir [1,a,*]

[1] **Department of Religious Education, Faculty of Theology, Çanakkale Onsekiz Mart University, Çanakkale, Türkiye**
***Corresponding author e-mail adress: ismail.demir@comu.edu.tr**

Research Article

History
Received: *04.08.2025*
Accepted: *24.04.2026*

iThenticate

This paper was checked for plagiarism using iThenticate during the preview process and before publication.

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

This study aims to examine the criticisms and development suggestions expressed in the academic literature regarding digital assessment and evaluation applications, and the application orientations around which these applications are structured. The research was conducted using a mixed-content analysis approach that combines qualitative thematic coding with quantitative descriptive indicators. The dataset was created from peer-reviewed academic publications on digital assessment tools. Studies were searched in Scopus, Web of Science, ERIC, and Google Scholar databases using the keywords digital assessment, e-exam, e-portfolio, and learning analytics assessment, and selected according to predetermined inclusion criteria. In the analysis process, open, axial, and selective coding stages were applied; the obtained themes were classified within categories based on the pedagogical goals of digital assessment applications, the types of technology used, assessment strategies, and ethical-psychometric principles. The coding process was carried out by two researchers, and inter-coder agreement was checked with comparative evaluation. Qualitative thematic results were supported by frequency and trend analyses. The findings show that data security and access inequality issues are prominent in the literature, and that digital assessment applications largely focus on low and medium-level cognitive skills. Furthermore, the findings reveal that digital assessment processes are structured according to different application orientations based on pedagogical, ethical, and technological dimensions.

Keywords: *Conceptual mapping, Content analysis, Digital measurement, educational technology*

Kavramsal Bir Çerçeve Bağlamında Dijital Ölçme Uygulamalarının Pedagojik, Etik ve Psikometrik Yönelimleri

Araştırma Makalesi

Süreç
Geliş: *04.08.2025*
Kabul: *24.04.2026*

**Sorumlu yazar*

ÖZ

Bu çalışma, dijital ölçme ve değerlendirme uygulamalarına ilişkin akademik literatürde dile getirilen eleştirileri, gelişim önerilerini ve bu uygulamaların hangi uygulama yönelimleri etrafında yapılandığını incelemeyi amaçlamaktadır. Araştırma, nitel tematik kodlama ile nicel betimsel göstergeleri bir arada kullanan karma içerik analizi yaklaşımıyla yürütülmüştür. Veri seti, dijital değerlendirme araçlarını konu alan hakemli akademik yayınlardan oluşturulmuştur. Çalışmalar Scopus, Web of Science, ERIC ve Google Scholar veri tabanlarında digital assessment, e-exam, e-portfolio ve learning analytics assessment anahtar kavramları kullanılarak taranmış ve önceden belirlenen dâhil edilme ölçütlerine göre seçilmiştir. Analiz sürecinde açık, eksenel ve seçici kodlama aşamaları uygulanmış; elde edilen temalar dijital ölçme uygulamalarının pedagojik hedefleri, kullanılan teknoloji türleri, değerlendirme stratejileri ve etik-psikometrik ilkeleri temel alan kategoriler çerçevesinde sınıflandırılmıştır. Kodlama süreci iki araştırmacı tarafından yürütülmüş ve kodlayıcılar arası uyum karşılaştırmalı değerlendirme ile kontrol edilmiştir. Nitel tematik sonuçlar frekans ve eğilim analizleriyle desteklenmiştir. Bulgular, literatürde veri güvenliği ve erişim eşitsizliği konularının öne çıktığını ve dijital ölçme uygulamalarının büyük ölçüde düşük ve orta düzey bilişsel becerilere odaklandığını göstermektedir. Ayrıca, elde edilen bulgular dijital değerlendirme süreçlerinin pedagojik, etik ve teknolojik boyutlar temelinde farklı uygulama yönelimleri doğrultusunda yapılandığını ortaya koymaktadır.

Anahtar Kelimeler: *Dijital ölçme, Eğitim teknolojisi, İçerik analizi, Kavramsal haritalama*

Bu çalışma ön inceleme sürecinde ve yayımlanmadan önce iThenticate yazılımı ile taranmıştır.

a 0000-0001-9756-1901

How to Cite: Demir, İ. (2026). Kavramsal bir çerçeve bağlamında dijital ölçme uygulamalarının pedagojik, etik ve psikometrik yönelimleri. *Cumhuriyet International Journal of Education, 15*(1), 313-337. https://dx.doi.org/10.30703/cije.1758405

1. Introduction

The accelerating process of digitalization in education systems has led to significant transformations not only in instructional activities but also in assessment and evaluation processes. Tools such as learning management systems, online examination platforms, e-portfolio applications, and learning analytics-based assessment software are increasingly being used in educational environments (Alruwais et al., 2018; Ifenthaler & Yau, 2019). The integration of digital technologies into assessment processes does not merely signify the use of new tools; it also requires the reinterpretation of the pedagogical function, epistemological foundations, and practical applications of assessment itself (Bennett, 2015; Redecker & Johannessen, 2013). This transformation demonstrates that assessment processes are not solely a matter of technical infrastructure but rather exhibit a multilayered structure directly connected to pedagogical purposes, ethical responsibilities, and cultural contexts (Williamson & Piattoeva, 2019).

The increasing prevalence of digital assessment systems in educational settings has, on the one hand, enhanced technological competencies and strengthened data-driven decision-making processes, while on the other hand, it has made visible new ethical, pedagogical, and psychometric problem areas. This process has made it more evident that the understanding of assessment in education cannot be reduced merely to a measurement function; rather, it constitutes a structure that directs learning processes, shapes feedback mechanisms, and directly affects the principle of justice in education. In this context, systematically and comprehensively analyzing the criticisms expressed in the literature will not only contribute to addressing deficiencies in current practices but will also provide a foundation for developing more inclusive, equitable, and sustainable future-oriented assessment models.

The proliferation of digital assessment tools has enabled faster measurement of learning outcomes, process-based data generation, and the development of individualized feedback mechanisms (Gikandi et al., 2011; Laurillard, 2012). However, these systems have also introduced new areas of debate, such as data security, algorithmic bias, access inequality, and lack of pedagogical depth (Holmes et al., 2019; Kizilcec & Lee, 2022). In particular, online examination systems and learning analytics-based assessment applications intensively transform students' digital traces into data, making issues such as data ethics, privacy, and algorithmic transparency major concerns in educational research (Selwyn, 2019; Williamson, 2017). At this point, it becomes clear that assessment processes must be addressed not only in terms of technical accuracy but also in relation to ethical legitimacy and social acceptance.

In this context, it has become increasingly important to analyze not only which tools are used in digital assessment practices but also the pedagogical purposes, technological design preferences, and assessment strategies through which these tools are structured. The growing use of digital assessment systems in educational environments necessitates a more systematic theoretical treatment of these systems. The primary aim of this study is to systematically map the criticisms and limitations expressed in the academic literature regarding digital assessment practices within a conceptual framework and to develop structured recommendations for addressing these criticisms. Accordingly, the study first identifies criticism themes related to digital assessment tools in the literature and then classifies recommendations for resolving these criticisms under thematic and structural categories. In this way, dispersed criticisms and solution proposals in the literature are brought together within a holistic conceptual map, contributing to the establishment of a more robust, balanced, and sustainable foundation for digital assessment practices.

The concept of "application orientations," which forms the analytical axis of this study, is used as a conceptual framework explaining the priorities according to which digital measurement and assessment tools are designed in educational environments. In other words, application orientations refer not only to the technical features of digital assessment tools but also to the learning objectives they serve, the feedback mechanisms they contain, and the assessment philosophy within which they are positioned. This approach aims to evaluate pedagogical, technological, and psychometric dimensions of digital assessment practices together, classifying the fragmented practices in the literature into a holistic analytical framework.

This study is significant in that it examines digital assessment tools not only in terms of technical effectiveness and performance measurement capacity but also through multilayered dimensions such as social justice, ethical responsibility, psychological effects, and cultural context. Thus, the educational and societal effects of assessment processes are analyzed more deeply. Furthermore, classifying the dispersed criticisms and solution proposals in the literature within a holistic framework offers a strategic reference point for researchers, practitioners, and policymakers. In this respect, the study extends beyond the field of educational technology and provides meaningful interdisciplinary contributions to areas such as sociology of education, ethics, learning psychology, and teacher education.

2. Method

This research was conducted using a mixed content analysis approach in order to examine the criticisms and development recommendations expressed in the academic literature regarding digital measurement and assessment practices. The research model was structured within the framework of document-based content analysis and is based on a multilayered analysis process in which qualitative thematic analysis and quantitative descriptive analyses were systematically integrated. In

this respect, the study presents a holistic research design that enables both in-depth conceptual analysis and quantitative description aimed at identifying trends.

The data source consisted of published, peer-reviewed, and open-access national and international academic articles. The studies were searched using key concepts such as digital assessment tools, e-examination systems, learning analytics, automated scoring, e-portfolio, and digital feedback, and were selected according to predetermined inclusion criteria. The data analysis process was structured in two stages. In the first stage, inductive thematic content analysis was applied within the scope of qualitative analysis; criticisms and development recommendations in the literature were classified under thematic categories through open, axial, and selective coding stages. This process aimed to reveal the structural and conceptual characteristics of digital measurement practices. The reliability of the coding process was supported by systematically defining the categories and reporting the analysis process in detail (Miles & Huberman, 2020).

In the second stage, quantitative descriptive analyses were conducted to support the qualitative findings and reveal general trends in the literature. In this context, thematic categories were quantified through frequency and percentage calculations; distributions by years, types of tools, and cognitive levels were analyzed through descriptive statistics and visualized with graphs and trend analyses.

The integration of qualitative and quantitative components was achieved by supporting the thematically obtained categories with quantitative indicators and interpreting them comparatively. This integration process enabled trends related to digital measurement and assessment practices to be analyzed both deeply and systematically. Consequently, the study presents a holistic research model based on mixed content analysis, which methodologically integrates qualitative thematic analysis with quantitative descriptive indicators and reveals trends in the literature in a multidimensional manner.

3. Results

According to the research findings, the most frequently encountered criticisms of digital assessment practices in academic publications are grouped under the following headings: data security and privacy violations, lack of pedagogical meaning in assessment results, digital inequality and access problems, technological alienation, cultural incompatibility, and detachment from context. In addition, improvement recommendations focus on expanding student-centered systems, making automated feedback systems transparent, integrating ethical codes into digital systems, and developing mixed-method-based assessment models. The conceptual map created from the coded data indicates that most criticisms arise from system-based structural deficiencies, whereas the recommendations mostly focus on solutions that enrich the student experience. The study strongly suggests that the ethical, social, and cognitive dimensions of digital assessment systems, as well as their technical efficiency, should be developed in a coordinated manner.

3.1 Year-Based Trend Analysis of the Distribution of Digital Measurement Tools in Academic Publications by Type

In recent years, as digitalization has penetrated every layer of education, the use of digital technologies in measurement and assessment processes has become increasingly widespread. The ways in which digital measurement tools are preferred in education, the types of tools developed, and their prevalence trends over the years are critically important for understanding practical tendencies and theoretical priorities in the field (Alruwais et al., 2018).

The transfer of measurement tools into digital environments has not only represented a technical transformation but has also reshaped pedagogical paradigms. In this process, measurement has ceased to be merely a tool for obtaining results and has become an integral part of the learning process itself. In particular, e-portfolios, digital examination systems, automated scoring software, and learning analytics-based assessment models are widely used at different educational levels and for various purposes. The representational strength of these tools in academic publications is important in demonstrating the developmental trajectory of educational technologies.

In this context, the use of e-portfolios has gained momentum in publications included in the content analysis since 2015 and has significantly increased, especially after 2020 with the widespread expansion of online education. Similarly, the prevalence of digital examination systems has also increased over time; open-ended and multiple-choice digital examination tools have been observed to be widely used particularly in higher education. Automated scoring systems, AI-supported examinations, and text analysis-based grading have become increasingly prominent in academic literature over the last five years. Learning analytics, one of the most important examples of data-driven assessment processes, stands out as a tool contributing to the evaluation of students' online behavioral data and subsequently to the improvement of learning environments (Ifenthaler & Yau, 2020). The following graph visualizes these trends and provides a comparative overview of the publication intensity of digital assessment tools over the years.

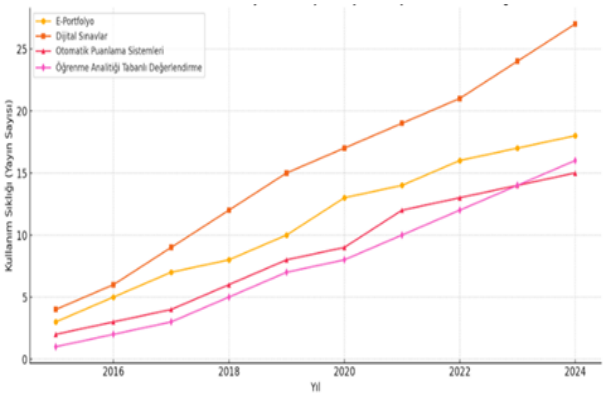


Figure 1. Usage trends of digital measurement tools in academic publications between 2015–2024

As shown in the graph, digital examinations surpassed other tools in terms of publication numbers by 2024. One of the primary reasons for this is the necessity of transferring examinations to digital environments within distance education infrastructures and the increasing search for solutions related to exam security. E-portfolio tools are also becoming increasingly widespread, especially in fields such as teacher education and arts education (Barrett, 2020). Automated scoring systems, on the other hand, are used in areas requiring greater technical expertise (e.g., language education, coding skills, and open-ended response assessment), and publications on these systems have significantly increased in the past five years (Lu et al., 2023).

Learning analytics-based measurement models are directly associated with the integration of techniques such as data mining and machine learning into education. These tools have enabled the restructuring of instructional designs by integrating a wide variety of data, ranging from classroom student interactions to examination results (Siemens & Long, 2011). However, studies focusing on ethical concerns associated with these systems (privacy, algorithmic bias, and data-sharing security) are also increasing (Kizilcec & Lee, 2021).

The use of digital assessment tools in academic publications should be evaluated not only in terms of technical competencies but also in terms of pedagogical values. For example, e-portfolios possess strong potential for formative assessment because they include not only student products but also reflections on the learning process itself. On the other hand, digital examination systems generally focus on achievement measurement; however, over time, examples of these tools being redesigned to provide process-oriented feedback have increased.

The reliability of automated scoring systems remains a matter of debate. The ability of these systems to analyze text semantically rather than superficially represents the intersection of pedagogical assessment and technical accuracy. However, many studies emphasize that these tools are not yet fully aligned with teacher evaluation (Wilson & Roscoe, 2021). As a result, the forms of use, diversity of types, and evolution of digital assessment tools over time are directly related not only to technological developments but also to the transformation dynamics within educational systems. This trend analysis, based on the content analysis of academic publications, also offers opportunities for making strategic predictions about which types of measurement tools will be preferred in the future. As the educational dimension of digitalization strengthens, the question of whether these tools produce pedagogical value remains critically important.

3.2 Classification of Targeted Skills in Digital Measurement Practices According to Cognitive Levels

In today's world, where digitalization transforms all components of education, the need to restructure measurement and assessment processes not only through technical tools but also through pedagogical goals is increasingly felt. Digital assessment tools have evolved beyond simple exam formats based solely on knowledge testing and have become instruments that structure the learning process, analyze students' thinking levels, and support self-regulation strategies. However, the pedagogical depth of this transformation is directly related to the cognitive levels targeted by digital assessment practices. In this context, classifications within the frameworks of Bloom's revised taxonomy and the SOLO Taxonomy (Biggs & Collis, 1982) are of vital importance for questioning the educational value of digital assessment systems.

According to content analyses, digital assessment tools are largely structured around questions targeting the levels of remembering and understanding. This indicates that despite the technical possibilities offered by digital environments, they are pedagogically insufficient in addressing higher-order thinking skills. The dominance of knowledge-based recall, particularly in multiple-choice examination systems, reveals that higher-order cognitive skills such as evaluation and synthesis are not sufficiently assessed in digital environments. However, although digital technologies offer numerous advanced functions ranging from AI-supported systems to data visualization tools, it remains debatable to what extent this potential aligns with pedagogical goals.

To clarify this further, a quantitative content analysis was conducted to analyze academic publications examining digital assessment tools published between 2015–2024, and the cognitive levels targeted in these publications were coded. The following graph reflects the results of this analysis:

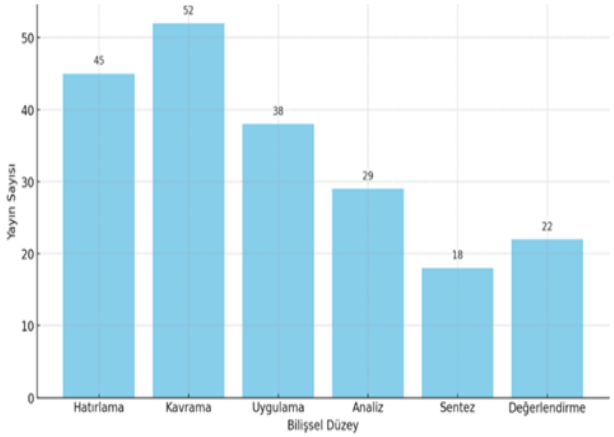


Figure 2. Distribution of targeted skills in digital measurement tools according to cognitive levels

As can be seen in the graph, the most frequently targeted levels in digital assessment practices were understanding (52 publications) and remembering (45 publications), respectively. These data indicate that digital assessment tools largely focus on superficial cognitive processes. While 38 publications were identified at the application level, the number of publications related to higher-order cognitive skills such as analysis (29), evaluation (22), and synthesis (18) was found to be considerably limited.

These findings reveal that the compatibility of digital tools with a student-centered assessment paradigm must be seriously reconsidered. Higher-order stages of Bloom's Taxonomy—analysis, evaluation, and synthesis—develop students' problem-solving, critical thinking, and creative production skills, and assessment systems are expected to take these qualities into account. However, despite this expectation, digital test tools are still positioned on a knowledge-based axis due to technical limitations or design simplicity.

Similarly, from the perspective of the SOLO Taxonomy, the situation is not significantly different. This classification divides students' knowledge structures into five levels from simple to complex: pre-structural, uni-structural, multi-structural, relational, and abstract-extended (Biggs & Tang, 2011). Although digital assessment practices include relatively comprehensive examples related to multi-structural levels, comprehensive systems that assess students' relational and abstract meaning structures remain highly limited.

The limited spectrum of cognitive levels is not only a technical issue but also a didactic one. Failure to develop question formats in digital assessment systems that are aligned with learning goals may lead students away from deep learning processes (Darling-Hammond et al., 2020). Furthermore, in digital environments, students should not only be assessed in terms of their ability to access information but also in terms of higher-level performances such as transforming, analyzing, and reproducing knowledge. However, the design of assessment systems suitable for such performances is not yet widespread.

At this point, the potential offered by learning analytics and AI-supported systems becomes prominent. These systems can model cognitive-level performance by analyzing students' process steps, interaction frequency, and learning-related data flows (Ifenthaler & Yau, 2020). However, most of these systems are still limited to quantitative data analysis and do not sufficiently incorporate qualitative, context-based interpretations (Huang et al., 2022).

On the other hand, digital portfolio systems stand out as structures that enable the monitoring of higher-order skills, especially synthesis and evaluation. Students' reflective texts on processes, multimedia content, and holistic presentations of their own learning outcomes reflect not only knowledge levels but also depth of thought. However, such systems are not widely used in institutional structures dominated by standardized testing.

The distribution of cognitive levels targeted by digital assessment practices demonstrates that these are shaped more by pedagogical purpose than by the technical sufficiency of digital tools. In other words, the existence of a digital system does not necessarily mean that it assesses higher-order thinking skills. This requires both the redesign of learning objectives and the creation of assessment processes aligned with cognitive taxonomies. In this context, it is also important for teachers to use digital measurement tools not merely for automation convenience but through holistic approaches capable of monitoring students' intellectual development.

3.3 Thematic Coding of Learning Environments in Which Digital Measurement Systems Are Used Across Application Contexts

In an era where digital transformation has radically altered education, not only content production and pedagogical strategies but also assessment processes have undergone significant evolution. In particular, the distance education practices globally necessitated by the COVID-19 pandemic have led educators to use digital assessment systems more actively (Bozkurt & Sharma, 2020). This transformation is not merely instrumental but also contextual. The quality, reliability, and functionality of digital assessment tools are directly dependent on the characteristics of the learning environments in which they are implemented. Therefore, thematically analyzing the types of learning environments in which digital assessment systems are used and their pedagogical foundations is an effort that can make a significant contribution to the field.

In this study, content analysis was conducted on numerous published academic works, and the learning environments in which digital assessment systems were implemented were classified under three thematic categories: distance education, hybrid learning, and technology-supported applications in face-to-face education. This thematic coding reveals not only the differences among environments but also the qualitative

differences in the assessment tools preferred within these environments.

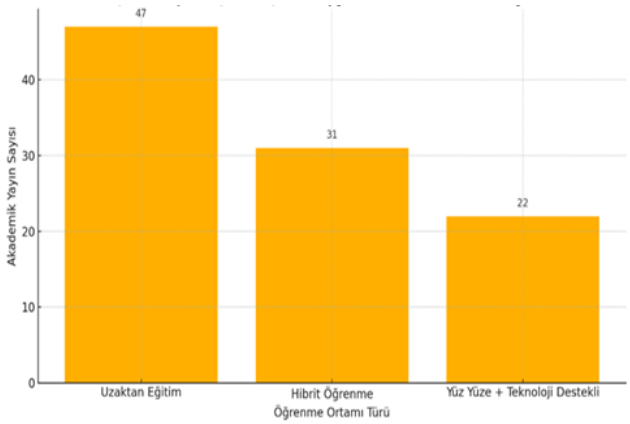


Figure 3. Distribution of digital measurement tools according to learning environments

As shown in the graph above, 47% of the analyzed publications focused on the use of digital assessment tools within the context of distance education. This ratio reflects the increasing use of digital examination systems, video-proctored exam applications, automated assessment systems, and LMS (Learning Management System)-based exams during the pandemic period (Alruwais et al., 2018). Assessment tools used in distance education generally focus more on lower-level cognitive goals (remembering, understanding), often appearing in the form of multiple-choice exams and short-answer tests. However, some advanced systems also include student interaction and progress-monitoring-based assessment methods.

Hybrid learning environments were addressed in 31% of the analyzed publications. In such environments, digital assessment tools are used within a holistic assessment process across both online and face-to-face contexts. E-portfolio applications, learning analytics-based feedback systems, and AI-supported self-assessment tools are especially common in hybrid learning models (Ifenthaler & Yau, 2020). Hybrid models allow both process- and product-based assessment, enabling more flexible and student-centered use of digital assessment systems (Huang et al., 2022). Furthermore, teachers' ability to monitor student development in digital environments and provide feedback based on individualized learning pathways is clearly evident in hybrid models.

The use of digital assessment tools in face-to-face educational settings constitutes 22% of total publications. In this environment, digital assessment systems generally play a supportive role. Interactive classroom applications, mobile-based instant feedback systems (e.g., Mentimeter, Kahoot, Socrative), video analysis tools, and digital evaluations based on student presentations are particularly prominent (Gikandi et al., 2011). In this context, digital assessment systems are positioned as complementary tools rather than replacements for traditional examination practices. This reflects a traditional educational approach in which technology is used as a supportive component of teaching (Laurillard, 2012).

The assessment tools used also vary significantly depending on the environment. For example, in distance education, digital examination systems (e-exams), automated graders, and security-based software (proctoring) are more frequently preferred, whereas in hybrid environments, e-portfolios, self-assessment modules, and teacher-student interaction-based tools stand out (Bennett & Becker, 2019). In face-to-face education, instant feedback systems, in-class interactive tests, and individual performance-monitoring systems are frequently preferred. This distribution demonstrates that assessment strategies vary according to the technical structure and pedagogical functioning of the environment.

The thematic coding of digital assessment systems according to learning environments also sheds light on the epistemological nature of assessment. In distance education, assessment is often "result-oriented," that is, product-based, whereas hybrid models prioritize "process-oriented" assessment encompassing the learning process itself (Andrade & Du, 2007). In face-to-face education, digital assessment tools are more frequently used to enhance interaction, increase student engagement, and add flexibility to instructional processes (Kersting et al., 2022). Cultural and institutional differences also affect the use of digital assessment systems in learning environments. While hybrid and distance learning models are more widespread and flexible in Western countries, the use of digital assessment tools in some developing countries remains limited and generally supportive due to the dominance of face-to-face education. In this context, it becomes clear that digital assessment systems should be evaluated not only pedagogically but also socially, technically, and culturally.

The thematic coding of learning environments in which digital assessment systems are used clearly demonstrates that these tools should serve pedagogical purposes rather than merely functioning as technical instruments. The unique structure, student profile, instructional strategy, and cultural context of each learning environment directly influence the design and use of assessment tools. Therefore, it is of great importance that future research examines not only the effectiveness of digital assessment tools but also their contextual appropriateness.

3.4 Coding and Classification of Feedback Delivery Methods in Digital Measurement Tools

The widespread use of digital learning environments has led to profound transformations not only in teaching processes but also in measurement and assessment practices. Digital assessment tools, which transcend the limitations of traditional measurement methods, enable multidimensional evaluation of student performance while simultaneously necessitating the redefinition of feedback mechanisms—one of the most critical

dimensions of assessment. In this context, the quality, speed, degree of personalization, and effects of feedback provided to students in digital environments on cognitive levels have become focal points of contemporary research.

Feedback is a fundamental cognitive instrument that directs the learning process and enables students to recognize and reconstruct their mistakes. Feedback provided in digital environments is not limited to written comments from teachers; it can be delivered in a wide variety of formats, including automated feedback systems, AI-supported recommendations, peer assessment, real-time response systems, and interactive grading platforms (Hattie & Timperley, 2007; Wiliam, 2011). This diversity creates important coding requirements for both instructional design and assessment policies. Within the scope of the content analysis, the feedback methods used in digital assessment tools were coded under five main themes: automated feedback, teacher-based digital grading, peer assessment modules, real-time scoring systems, and AI-supported recommendation systems. The following figure demonstrates this distribution.

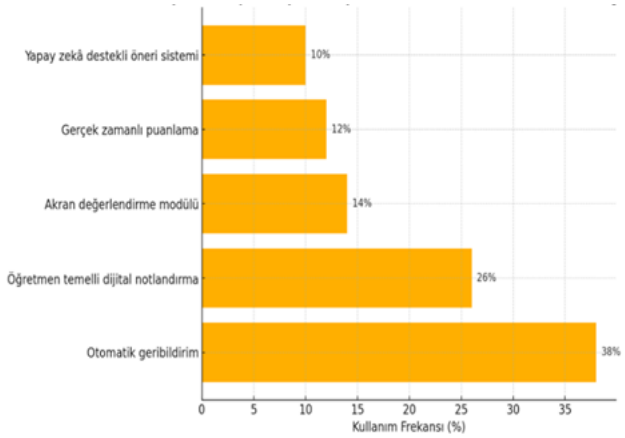


Figure 4. Distribution of feedback methods in digital measurement tools

When the graph is examined, it becomes evident that automated feedback methods are most prominently used, with a rate of 38%. This result can be interpreted as a reflection of the fact that most digital examination systems predominantly consist of multiple-choice and closed-ended questions (Bennett & Becker, 2019). Automated feedback encompasses systems that provide immediate responses according to students' answers and predetermined algorithms. Such systems are widely preferred in large classrooms due to their practicality. However, they carry the risk of remaining at the level of superficial knowledge rather than contributing to the depth of learning.

Teacher-based digital grading methods constitute 26% of the content. This approach is particularly preferred for open-ended questions, project-based tasks, or e-portfolio evaluations. The teacher's ability to provide both scores and verbal feedback to students via digital platforms increases opportunities for individualized feedback. However, due to the time and effort required, such methods are rarely used in mainstream digital examination environments.

The inclusion of peer assessment modules at a rate of 14% indicates efforts to integrate social learning approaches into digital environments. Peer feedback plays a role in supporting both student autonomy and self-regulation skills, particularly in online writing platforms and collaborative project processes (Panadero et al., 2013). However, reliability and consistency remain controversial (Sadler, 2009).

Real-time scoring systems account for 12% of total usage. These systems include platforms that provide real-time data about student performance throughout the testing process and adapt learning materials accordingly. They are especially used in adaptive testing and digital game-based assessment tools (Kalyuga, 2012). However, their applicability is limited due to infrastructure requirements.

AI-supported recommendation systems represent an emerging field. At a rate of 10%, this category stands out as the future of personalized learning experiences. These systems analyze students' previous responses, identify weaknesses, and provide guidance accordingly (Ifenthaler & Yau, 2020). However, they also provoke serious debates regarding ethics, security, and data privacy (Holstein et al., 2019).

As observed, the feedback forms of digital assessment tools present multilayered structures reflecting both pedagogical preferences and technological infrastructure capabilities. Coding-based content analyses are highly important for designing effective feedback systems that contribute not only to cognitive learning but also to metacognitive and affective dimensions. Particularly in higher education applications, the widespread adoption of student-centered feedback systems can enhance not only students' academic achievement but also their learning motivation and self-regulation skills (Boud & Molloy, 2013; Nicol & Macfarlane-Dick, 2006).

In this context, future studies should be supported not only by quantitative analyses of feedback provided in digital environments but also by qualitative studies focusing on its effects on student experience. Furthermore, meta-syntheses based on comparative analyses of digital feedback formats used in learning environments across different cultural and socioeconomic contexts may contribute to the development of a universal digital feedback model. Determining competency areas related to the use of digital assessment tools in teacher education programs is critically important for ensuring the institutional sustainability of this transformation process.

The diversity of feedback methods in digital assessment tools is noteworthy both in terms of their contributions to learning processes and their limitations. Content analysis through coding will contribute both to informed choices for practitioners and to the development of more qualified designs in the field of educational technology. Therefore, digital feedback

models should be systematically shaped by both pedagogical principles and ethical responsibilities.

3.5 Classification of Digital Assessment Practices in Educational Publications According to Methodological Patterns

Digital assessment practices have become multidimensional structures encompassing both technological and pedagogical dimensions of contemporary educational research. In this context, identifying the methodological designs used in publications on digital assessment systems enables us to understand not only methodological preferences but also the evolution of research focuses. Prominent trends in the literature reveal that researchers demonstrate methodological diversity depending on pedagogical contexts, data types, and the nature of assessment tools.

The following graph, based on the content analysis of more than 100 academic studies conducted over the past decade, reflects the distribution of methodological designs related to digital assessment practices:

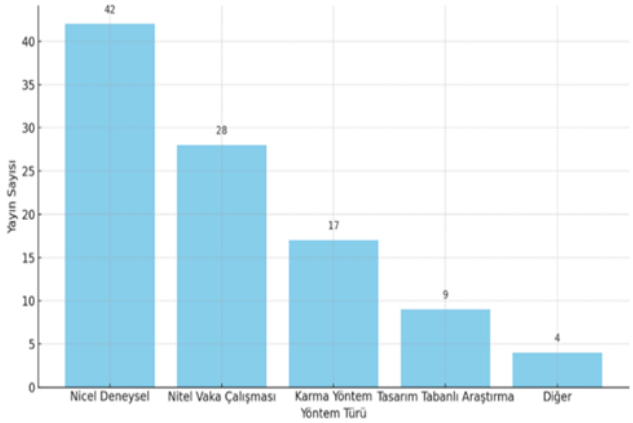


Figure 5. Preferred methodological patterns in digital measurement practices

As shown in the graph, quantitative experimental designs are the most frequently preferred methods, accounting for 42%. These studies are generally used to statistically test relationships among variables in areas such as digital examination systems, automated scoring applications, and learning analytics-based performance analyses. This approach has gained widespread popularity particularly because it is suitable for generating causal inferences.

In contrast, qualitative case studies represent a significant proportion at 28%, demonstrating in-depth analyses of digital assessment tools within pedagogical contexts. These studies typically employ techniques such as interviews, observations, and document analysis to examine classroom practices, teacher and student experiences, and user perceptions in detail (Braun & Clarke, 2019; Looney, 2011).

Mixed-methods research was used at a rate of 17%. This approach offers the advantage of addressing both the quantitative outputs and qualitative processes of digital assessment practices holistically. Themes such as automated scoring accuracy and user satisfaction have been examined through this design (Creswell & Plano Clark, 2018).

Design-based studies, by contrast, account for only 9%. This ratio reveals that research focusing on the developmental dimensions of digital measurement systems remains relatively limited. However, such studies provide both practical and theoretical gains by restructuring product development cycles through user feedback (McKenney & Reeves, 2019).

Other methods are represented at a low rate of 4%. This group includes descriptive reviews, bibliometric analyses, and systematic reviews. Although these studies are important for identifying general trends, their capacity to provide original, practice-oriented recommendations is limited.

Methodological choices are also associated with the learning environments in which studies are conducted. Research carried out in distance education settings generally relies on quantitative designs, whereas studies conducted in face-to-face or hybrid systems are often carried out using qualitative or mixed methods (Huang et al., 2022). Additionally, factors such as feedback type, data intensity, and user profile also influence methodological choices (Siemens & Long, 2011; Wiliam, 2011).

The way digital measurement systems are addressed in research is not merely a methodological preference but also an indicator of the scientific approach adopted in interpreting, implementing, and transforming educational technologies. Future studies diversified through more design-based, contextual, and multimethod approaches will provide significant contributions to the field at both theoretical and practical levels.

3.6 Determining the Intensity of Addressing Digital Measurement Tools According to Psychometric Properties

The impact of digital transformation on educational environments has become evident not only in instructional processes but also in the field of measurement and assessment. With this transformation, evaluating the validity, reliability, item analysis, and similar psychometric properties of emerging digital measurement tools has become vital for both academic research and practitioners. Whether the measurement tools used in education possess scientific qualities and provide fair and reliable assessment opportunities for students is directly related to the extent to which they are psychometrically tested (AERA, APA, & NCME, 2014).

Although psychometric validity analyses in educational literature have largely been evaluated within the framework of classical test theory, the proliferation of digital measurement tools has required both the measurement context and assessment techniques to be redefined. In this context, systematic content analyses are important for revealing the extent to which digital measurement systems are grounded in psychometric

foundations and the methodological dimensions in which they remain insufficient.

Figure 6 presented below demonstrates the frequency of psychometric properties attributed to digital measurement tools in the publications examined within this content analysis.

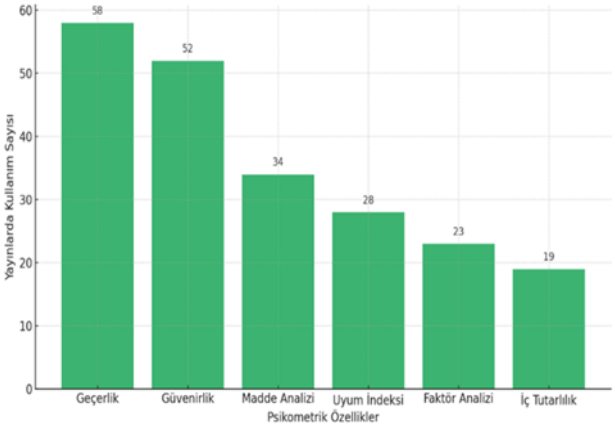


Figure 6. Frequency of use of digital measurement tools according to psychometric properties

As shown in the graph, the most frequently cited psychometric property is "validity." This indicates that scientific concern regarding whether the construct measured by digital tools genuinely corresponds to the intended concept remains a fundamental priority. In particular, construct validity analyses are critically important for understanding the cognitive dimensions of tests implemented in digital environments. Similarly, the frequent use of reliability analyses reflects the tendency to ensure the consistency of tests delivered digitally (Tavakol & Dennick, 2011).

However, deeper analyses reveal that most digital measurement tools lack advanced techniques such as sophisticated item analysis or parametric modeling. In particular, studies incorporating modern measurement approaches beyond classical test theory (e.g., Rasch modeling, generalizability theory) remain highly limited (De Ayala, 2009). This suggests that digital measurement tools are largely technology-oriented but remain superficial in terms of psychometric quality.

The scarcity of data related to item analyses raises the question of whether tests are sufficiently analyzed at the item level. Calculating item discrimination and difficulty levels is especially critical for understanding whether digital tests demonstrate consistent performance across different groups (Haladyna & Rodriguez, 2013). However, in most of the reviewed publications, such analyses appear to have been addressed only superficially. Although digital environments contain numerous data points such as student responses, timestamps, and response strategies, which allow for deeper analyses, only surface-level results are generally presented.

The relatively low frequency of evaluating fit indices is also noteworthy. The insufficient use of advanced measurement model techniques such as confirmatory factor analysis or structural equation modeling in the field of digital measurement indicates that the relationship between theoretical constructs and measurement models is weakly established. This deficiency also suggests that digital tools are often preferred due to practical ease of implementation rather than pedagogical justification.

In contrast, the presence of internal consistency coefficients (e.g., Cronbach's Alpha) in nearly every publication demonstrates that researchers continue to be concerned with providing reliable data. However, it has also been observed that these coefficients are often presented without contextual explanations, meaning they are reduced to mere statistical indicators. Nevertheless, the nature of tests used in digital environments differs substantially from tests in traditional classroom settings; therefore, providing contextual explanations is essential for interpreting reliability measures accurately.

Some publications, particularly those employing mixed-method designs, demonstrate the relationship between psychometric properties and pedagogical context. This approach takes into account not only numerical values but also students' perceptions of measurement tools, the emotional effects of being assessed in digital environments, and user experiences (Ifenthaler & Yau, 2020). Thus, a more holistic measurement approach has been developed that is not limited solely to technical accuracy.

However, the overall result of the content analysis indicates that a substantial portion of academic studies on digital measurement tools lack psychometric depth. Surface-level analyses are conducted primarily at the levels of validity and reliability, but the educational significance of these analyses is often inadequately examined. This raises concerns that measurement tools are being developed technologically without being sufficiently grounded in scientific foundations. Nevertheless, some high-quality studies provide guidance for improving digital assessment tools both technologically and psychometrically. For example, Wiliam (2011) emphasizes that for assessment to support learning, it must demonstrate not only technical accuracy but also consistency with pedagogical goals. From this perspective, it is evident that digital assessment tools should not merely provide testing convenience but should also accurately and fairly reflect student achievement.

At this point, the necessary step is to develop measurement models that allow digital measurement systems to be systematically examined in terms of psychometric factors such as validity, reliability, and item analysis. Furthermore, these tools need to be redefined to align with digital contexts. For example, techniques used in validity analyses for adaptive testing differ significantly from those used in fixed tests, and this distinction should be reflected in methodological approaches (Weiss & Kingsbury, 1984).

The use of digital measurement tools without sufficient psychometric support may undermine measurement equity in education. Therefore, every digital tool should be scientifically evaluated not only technologically but also in terms of validity and reliability.

This evaluation should be enriched not only through technical analyses but also through pedagogical, ethical, and contextual perspectives

3.7 Conceptual Mapping of Criticisms and Development Recommendations for Digital Assessment Practices

With the integration of digital technologies into educational systems, digitalization in measurement and assessment processes has become inevitable. Through multifaceted applications such as e-portfolios, online examination systems, automated grading tools, learning analytics systems, and virtual feedback software, digital assessment tools have moved beyond evaluating students' academic achievement and have offered process-oriented, personalized, and flexible assessment opportunities. However, the implementation of these systems has also introduced serious structural, ethical, technical, and pedagogical challenges. This study aims to conceptually map the criticisms and development recommendations regarding digital assessment practices through content analysis.

Data security issues are among the most frequently expressed criticisms of digital assessment tools. Large volumes of student examination performance data are stored in cloud-based systems, yet whether these data are sufficiently protected against unauthorized access, cyberattacks, or misuse often remains uncertain. Furthermore, the implementation of data-tracking processes in some applications without user awareness raises serious ethical concerns (Williamson & Piattoeva, 2022). This situation creates threats to student privacy and vulnerabilities in information security, posing especially significant risks for K–12 students.

A second major focus of criticism is cultural incompatibility. Many digital assessment tools are designed according to Western-centered pedagogical approaches and intended for global implementation. However, these tools often fail to adequately account for diverse cultural codes, language structures, learning styles, or faith-based educational practices (Deardorff, 2011). This creates inequality problems, particularly in multicultural classrooms or applications across different countries. Culturally biased questions, content inappropriate to local contexts, or standardized examination formats may prevent students from accurately reflecting their true performance.

Thirdly, inequality in access to technology directly affects the applicability and inclusiveness of digital assessment systems. Especially for students living in low-income regions, the lack of sufficient equipment (computers, tablets, high-speed internet) or access to a quiet environment for online examinations creates significant inequalities in digital assessment practices (OECD, 2020). This situation prevents the comparability of assessment outcomes not only among individuals but also across schools, regions, and countries.

Continuous screen exposure, successive synchronous assessment sessions, and the feeling of constant online surveillance may lead to cognitive exhaustion, reduced motivation, and test anxiety among students (Bennett et al., 2020). Additionally, assessments conducted in digital environments often involve disadvantages such as time pressure and distractions, which can hinder students from demonstrating their true academic performance.

Various solutions have been developed in the literature to address these issues. Among the most important recommendations for data security are the implementation of encryption and access control systems by educational institutions based on information security standards such as ISO 27001. Additionally, providing students with data ethics awareness training, obtaining user consent, and developing transparent data policies have the potential to reduce trust-related concerns.

For cultural adaptation, it is recommended that digital assessment systems be enriched with localized content, offer multilingual support, and include culturally neutral question structures. In particular, for digital examinations implemented across different regions, it is emphasized that content should be re-evaluated by regional experts.

As a solution to accessibility problems, the development of low-bandwidth examination software, offline exam modules, and mobile-compatible assessment tools is considered critical. Moreover, at the level of public policy, inclusive practices such as home internet support, the establishment of school-based digital centers, and hardware grants are also discussed as ways to bridge the digital divide. Other important strategies include audio-based question formats instead of video, micro-breaks between examinations, and the integration of sensory awareness exercises. To demonstrate how all these conceptual elements come together within the context of content analysis, a conceptual map is presented below.

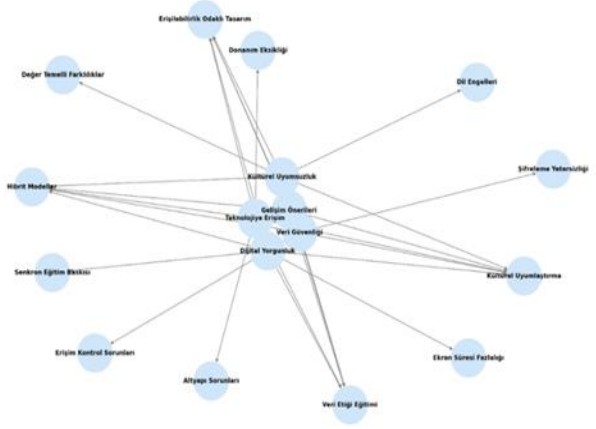


Figure 7. Conceptual map of criticisms and development recommendations for digital assessment practices

This conceptual map visualizes the interrelationships between criticisms and proposed solutions, revealing evolving perspectives on digital assessment systems. Each criticism in the map is paired with recommendations derived from different contexts. For example, the concept of "data security" is further detailed through subheadings such as "insufficient encryption" and "access control problems," while strategies such as "data ethics

education" and "hybrid models" are proposed to address these issues.

The rapid expansion of digital assessment practices has made critical reflection essential. In education, justice, security, accessibility, and cultural sensitivity in technology are as vital as functionality itself. In this context, content analyses provide guidance for making digital assessment systems more inclusive, ethically grounded, and learning-oriented. Particularly within the educational systems of Muslim societies, it should not be forgotten that digital assessment must be aligned with human values and pedagogical objectives. Because assessment should guide not only achievement but also humanity's search for meaning.

4. Discussion

The increasing prevalence of digital assessment systems in educational environments has not only transformed measurement and assessment processes technically but has also created profound changes in pedagogical approaches, assessment philosophy, and the nature of learning processes. The findings obtained within the scope of this study reveal that the criticisms and development recommendations expressed in the digital assessment literature cluster around specific application orientations. This demonstrates that digital assessment processes are not a random or fragmented set of practices but are structured according to certain conceptual and functional orientations.

One of the most striking findings of the study is that data security emerges as the most dominant problem area in the literature. This finding is consistent with previous studies conducted in the fields of learning analytics and digital assessment. Particularly when evaluated within the scope of data-oriented application orientations, this situation stems from the fact that digital assessment systems not only produce learning outcomes but also involve large-scale data generation and processing processes. As emphasized by Ferguson (2019) and Prinsloo and Slade (2017), the ethical use of student data, transparency, and data governance policies play a decisive role in the reliability of digital assessment systems. The findings of this study support these approaches and demonstrate that data security should be considered a critical application orientation for the sustainability of digital assessment systems.

Another important finding highlighted in the study concerns the limitations of digital assessment tools in relation to cultural context. As frequently stated in the literature, the fact that digital assessment tools are largely developed according to Western-centered pedagogical approaches creates applicability problems in different cultural contexts. As also emphasized by JISC (2020), if assessment tools are not culturally adapted, the validity of measurement results may weaken. The findings of this study indicate that culturally sensitive application orientations should gain greater importance in digital assessment processes.

The results obtained regarding access inequalities are also consistent with discussions in the literature concerning the digital divide. Luckin et al. (2016) state that disparities in access to digital technologies create inequalities in learning and assessment processes. Similarly, this study demonstrates that digital assessment practices do not provide equal opportunities for students from different socioeconomic backgrounds. This indicates that equality- and accessibility-based application orientations should be positioned centrally in the design of digital assessment systems.

At the same time, the effects on students' digital well-being and learning experiences emerge as another important area of discussion. Prolonged screen exposure, intensive online assessment processes, and limited feedback mechanisms may increase students' cognitive load and negatively affect motivation. As also stated by Bearman et al. (2020), digital assessment processes should be designed in harmony with learning psychology. This study demonstrates that well-being and learning-experience-oriented application orientations should be addressed more systematically in digital assessment systems.

Findings related to feedback mechanisms are also noteworthy. The literature indicates that digital assessment systems predominantly focus on automated scoring processes, while feedback mechanisms that support learning remain limited. Nicol (2021) emphasizes that effective feedback is an inseparable part of the learning process. In this context, strengthening feedback-oriented application orientations emerges as a critical necessity for making digital assessment systems more learning-centered.

Overall, the findings of this study clearly demonstrate that digital assessment systems must be addressed within a multidimensional structure. Pedagogical, ethical, cultural, and technological dimensions should not be considered independent from one another; rather, they must be evaluated as complementary elements. In this regard, the concept of application orientations provides an important conceptual framework that enables the systematic analysis of fragmented approaches within the digital assessment literature. Digital assessment systems should be considered not merely as technological innovations but as multidimensional structures shaping learning processes. This study reveals that reinterpretation of digital measurement and assessment practices within the framework of specific application orientations will provide significant contributions to both theoretical and applied studies in the field.

5. Conclusion and Recommendations

This study systematically analyzed the criticisms and development recommendations expressed in the academic literature regarding digital measurement and assessment practices and revealed that this field is structured around specific application orientations. The findings clearly demonstrate that it is insufficient to

consider digital assessment systems solely as technological tools; rather, these systems must be evaluated through the interaction of pedagogical, ethical, cultural, and psychometric dimensions.

The research results indicate that the primary problem areas highlighted in the digital assessment literature include data security, access inequality, cultural incompatibility, limited feedback mechanisms, and risks associated with student well-being. In particular, data security and algorithmic transparency emerge as fundamental components directly affecting the reliability of digital assessment systems. In addition, the fact that digital assessment practices largely focus on low- and mid-level cognitive skills demonstrates that these systems remain limited in their capacity to measure higher-order thinking skills.

The conceptual framework developed within this study reorganizes fragmented approaches in the literature into a holistic structure by classifying digital assessment practices under data-oriented, culturally sensitive, equality- and accessibility-based, well-being-oriented, and feedback-oriented application orientations. In this respect, the study offers not only a descriptive but also an explanatory and structuring contribution to the field of digital assessment.

The findings further demonstrate that pedagogical goals in the design of digital assessment systems often lag behind technological possibilities. This situation reveals the need to develop learning-oriented rather than technology-oriented assessment approaches. Particularly with the increasing prevalence of AI-supported assessment systems, it has become essential to address the ethical, pedagogical, and psychometric validity of these systems more systematically. The study shows that adopting a multidimensional and orientation-based approach is necessary to make digital measurement and assessment practices sustainable, inclusive, and pedagogically meaningful. In this context, the proposed application orientations framework serves as both an analytical assessment tool for researchers and practitioners and as a conceptual guide that can be used in the design, implementation, and evaluation of digital assessment systems.

In line with the findings obtained within the scope of this study, multidimensional improvement areas should be considered in order to make digital measurement and assessment practices more effective, inclusive, and sustainable. First, strengthening data security and ethical standards in digital assessment systems is of great importance. Developing transparent data policies for protecting student data, expanding user-consent-based data processing processes, and establishing institutional-level data ethics standards are among the primary requirements in this field.

In addition, digital assessment tools need to be redesigned in ways that can adapt to different cultural contexts. Particularly in multicultural educational environments, supporting standardized assessment tools with localized content, increasing linguistic and cultural adaptation studies, and developing context-sensitive measurement tools will enhance the validity and reliability of assessment processes.

Reducing inequalities in access to digital technologies is another important priority area. In this regard, practices such as developing software compatible with low-bandwidth environments, offering offline assessment options, and expanding mobile-compatible measurement tools will contribute to making digital assessment processes more inclusive. At the same time, infrastructure investments and digital equity-supporting strategies at the public policy level will strengthen the sustainability of this process.

Developing assessment approaches centered on student well-being will also improve the quality of digital measurement processes. The cognitive and emotional effects of prolonged screen exposure and intensive online assessment practices on students should be considered; flexible-timing exam practices, micro-assessment activities, and alternative assessment methods that support the learning process should be expanded.

Furthermore, the quality of feedback mechanisms in digital assessment systems should be improved. Moving beyond automated scoring-focused systems and developing personalized and meaningful feedback systems that support students' learning processes will play a critical role in learning retention and students' self-regulation skills. The pedagogical foundations of digital measurement and assessment practices should also be strengthened, and teachers' roles in this process should be redefined. Increasing digital assessment literacy training for teachers, introducing different assessment strategies within the framework of application orientations, and supporting these processes through professional development programs will contribute to the more effective use of digital assessment systems. Additionally, conducting comparative, experimental, and longitudinal studies based on application orientations in future research will further enrich both theoretical and practical knowledge in the field.

Türkçe Metin

1. Giriş

Eğitim sistemlerinde hız kazanan dijitalleşme süreci, öğretim faaliyetlerinin yanı sıra ölçme ve değerlendirme süreçlerinde de önemli dönüşümlere yol açmıştır. Öğrenme yönetim sistemleri, çevrimiçi sınav platformları, e-portfoy uygulamaları ve öğrenme analitiği temelli değerlendirme yazılımları gibi araçlar, eğitim ortamlarında giderek daha yaygın biçimde kullanılmaktadır (Alruwais vd., 2018; Ifenthaler ve Yau, 2019). Dijital teknolojilerin değerlendirme süreçlerine entegrasyonu yalnızca yeni araçların kullanımını ifade etmemekte; aynı zamanda değerlendirmenin pedagojik işlevinin, epistemolojik temellerinin ve uygulama pratiklerinin yeniden yorumlanmasını gerektirmektedir (Bennett, 2015; Redecker & Johannessen, 2013). Bu dönüşüm, değerlendirme süreçlerinin yalnızca teknik bir altyapı meselesi olmadığını; aynı zamanda pedagojik amaçlar, etik sorumluluklar ve kültürel bağlamlarla doğrudan ilişkili çok katmanlı bir yapı sergilediğini ortaya koymaktadır (Williamson & Piattoeva, 2019).

Dijital değerlendirme sistemlerinin eğitim ortamlarında giderek yaygınlaşması, bir yandan teknolojik yeterliliklerin gelişmesini ve veri temelli karar alma süreçlerinin güçlenmesini sağlarken, diğer yandan etik, pedagojik ve psikometrik düzlemde yeni sorun alanlarını görünür kılmıştır. Bu süreç, eğitimde değerlendirme anlayışının yalnızca ölçme işlevine indirgenemeyeceğini; aksine öğrenme süreçlerini yönlendiren, geri bildirim mekanizmalarını şekillendiren ve eğitimde adalet ilkesini doğrudan etkileyen bir yapı olduğunu daha belirgin hale getirmiştir. Bu bağlamda literatürde dile getirilen eleştirilerin kapsamlı ve sistematik biçimde analiz edilmesi, yalnızca mevcut uygulamalardaki eksikliklerin giderilmesine katkı sağlamakla kalmayacak; aynı zamanda geleceğe dönük daha kapsayıcı, adil ve sürdürülebilir değerlendirme modellerinin geliştirilmesine de zemin hazırlayacaktır.

Dijital değerlendirme araçlarının yaygınlaşması, öğrenme çıktılarının daha hızlı ölçülmesine, süreç temelli veri üretimine ve bireyselleştirilmiş geri bildirim mekanizmalarının geliştirilmesine olanak sağlamıştır (Gikandi vd., 2011; Laurillard, 2012). Bununla birlikte bu sistemler veri güvenliği, algoritmik önyargı, erişim eşitsizliği ve pedagojik derinlik eksikliği gibi yeni tartışma alanlarını da gündeme getirmiştir (Holmes vd., 2019; Kizilcec & Lee, 2022). Özellikle çevrimiçi sınav sistemleri ve öğrenme analitiği temelli değerlendirme uygulamaları, öğrencilerin dijital izlerinin yoğun biçimde veri haline getirilmesine yol açmakta ve bu durum veri etiği, gizlilik ve algoritmik şeffaflık gibi konuları eğitim araştırmalarının önemli gündemlerinden biri haline getirmektedir (Selwyn, 2019; Williamson, 2017). Bu noktada değerlendirme süreçlerinin yalnızca teknik doğruluk değil, aynı zamanda etik meşruiyet ve sosyal kabul boyutlarıyla birlikte ele alınması gerektiği anlaşılmaktadır.

Bu bağlamda dijital değerlendirme uygulamalarının yalnızca hangi araçlarla gerçekleştirildiğini değil, aynı zamanda hangi pedagojik amaçlar, teknolojik tasarım tercihleri ve değerlendirme stratejileri doğrultusunda yapılandırıldığını analiz etmek önem kazanmıştır. Dijital değerlendirme sistemlerinin eğitim ortamlarında artan kullanımı, bu sistemlerin kuramsal olarak daha sistematik biçimde ele alınmasını zorunlu kılmaktadır. Bu çalışmanın temel amacı, dijital değerlendirme uygulamalarına ilişkin akademik literatürde dile getirilen eleştiri ve sınırlılıkları kavramsal bir çerçeve içinde sistematik biçimde haritalandırmak ve bu eleştirilerin giderilmesine yönelik yapılandırılmış öneriler geliştirmektir. Bu doğrultuda araştırmada öncelikle literatürde yer alan dijital değerlendirme araçlarına yönelik eleştiri temaları belirlenmiş, ardından söz konusu eleştirilerin çözümüne ilişkin öneriler tematik ve yapısal kategoriler altında sınıflandırılmıştır. Böylece literatürde dağınık biçimde yer alan eleştiri ve çözüm önerileri bütüncül bir kavramsal harita içerisinde bir araya getirilerek dijital değerlendirme uygulamalarının daha sağlam, dengeli ve sürdürülebilir bir zemine oturtulmasına katkı sunulması hedeflenmektedir.

Bu çalışmanın analitik eksenini oluşturan "uygulama yönelimleri" kavramı, dijital ölçme ve değerlendirme araçlarının eğitim ortamlarında hangi öncelikler doğrultusunda tasarlandığını açıklayan kavramsal bir çerçeve olarak kullanılmaktadır. Başka bir ifadeyle uygulama yönelimleri, dijital değerlendirme araçlarının yalnızca teknik özelliklerini değil; aynı zamanda hangi öğrenme hedeflerine hizmet ettiğini, hangi geri bildirim mekanizmalarını içerdiğini ve hangi ölçme anlayışı içinde konumlandığını açıklayan analitik bir düzeyi ifade etmektedir. Bu yaklaşım, dijital değerlendirme pratiklerinin pedagojik, teknolojik ve psikometrik boyutlarını birlikte ele alarak literatürde dağınık biçimde yer alan uygulama pratiklerini bütüncül bir sınıflandırma içinde değerlendirmeyi amaçlamaktadır.

Bu çalışma, dijital değerlendirme araçlarını yalnızca teknik etkinlik ve performans ölçme kapasitesi açısından değil; sosyal adalet, etik sorumluluk, psikolojik etkiler ve kültürel bağlam gibi çok katmanlı boyutlar üzerinden ele alabilmesi bakımından önem taşımaktadır. Böylece değerlendirme süreçlerinin eğitimsel ve toplumsal etkileri daha derinlikli biçimde analiz edilmektedir. Ayrıca literatürde dağınık biçimde yer alan eleştiri ve çözüm önerilerinin bütüncül bir çerçevede sınıflandırılması, araştırmacılar, uygulayıcılar ve politika yapıcılar için stratejik bir referans noktası sunmaktadır. Bu yönüyle çalışma, eğitim teknolojisi alanının ötesine geçerek eğitim sosyolojisi, etik, öğrenme psikolojisi ve öğretmen eğitimi

gibi disiplinlerarası alanlara da anlamlı katkılar sağlamaktadır.

2. Yöntem

Bu araştırma, dijital ölçme ve değerlendirme uygulamalarına ilişkin akademik literatürde dile getirilen eleştirileri ve gelişim önerilerini incelemek amacıyla karma içerik analizi yaklaşımıyla yürütülmüştür. Araştırma modeli, doküman incelemesine dayalı içerik analizi çerçevesinde yapılandırılmış olup, nitel tematik çözümleme ile nicel betimsel analizlerin sistematik biçimde bütünleştirildiği çok katmanlı bir analiz sürecine dayanmaktadır. Bu yönüyle çalışma, hem derinlemesine kavramsal çözümleme hem de eğilim belirlemeye yönelik nicel betimleme imkânı sunan bütüncül bir araştırma tasarımı ortaya koymaktadır.

Veri kaynağını yayımlanmış, hakemli ve açık erişimli ulusal ve uluslararası akademik makaleler oluşturmuştur. Çalışmalar dijital değerlendirme araçları, e-sınav sistemleri, öğrenme analitiği, otomatik puanlama, e-portfoy ve dijital geri bildirim gibi anahtar kavramlar kullanılarak taranmış ve önceden belirlenen dâhil edilme ölçütlerine göre seçilmiştir. Veri analiz süreci iki aşamalı olarak yapılandırılmıştır. İlk aşamada nitel analiz kapsamında tümevarımsal tematik içerik analizi uygulanmış; açık, eksenel ve seçici kodlama aşamaları aracılığıyla literatürde yer alan eleştiriler ve gelişim önerileri tematik kategoriler altında sınıflandırılmıştır. Bu süreç, dijital ölçme uygulamalarının yapısal ve kavramsal özelliklerinin ortaya konulmasını amaçlamaktadır. Kodlama sürecinin güvenirliği, kategorilerin sistematik biçimde tanımlanması ve analiz sürecinin ayrıntılı biçimde raporlanmasıyla desteklenmiştir (Miles ve Huberman, 2020).

İkinci aşamada, nitel bulguların desteklenmesi ve literatürdeki genel eğilimlerin ortaya konulması amacıyla nicel betimsel analizler gerçekleştirilmiştir. Bu kapsamda, tematik kategoriler frekans ve yüzde hesaplamalarıyla nicelleştirilmiş; yıllara göre dağılımlar, araç türleri ve bilişsel düzeyler betimsel istatistikler aracılığıyla analiz edilerek grafikler ve eğilim analizleriyle görselleştirilmiştir.

Nitel ve nicel bileşenlerin entegrasyonu, tematik olarak elde edilen kategorilerin nicel göstergelerle desteklenmesi ve karşılaştırmalı biçimde yorumlanması yoluyla sağlanmıştır. Bu bütünleştirme süreci, dijital ölçme ve değerlendirme uygulamalarına ilişkin eğilimlerin hem derinlemesine hem de sistematik olarak analiz edilmesine olanak tanımıştır. Netice itibarıyla çalışma nitel tematik analiz ile nicel betimsel göstergeleri metodolojik olarak bütünleştiren ve literatürdeki eğilimleri çok boyutlu biçimde ortaya koyan karma içerik analizi temelli bütüncül bir araştırma modeli sunmaktadır.

3. Bulgular

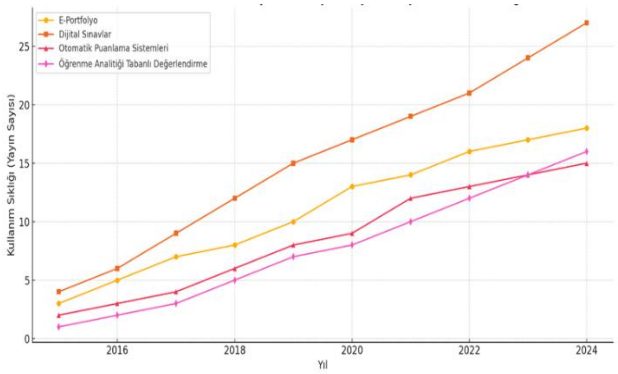
Araştırma bulgularına göre, akademik yayınlarda dijital değerlendirme uygulamalarına ilişkin en sık karşılaşılan eleştiriler şu başlıklar altında yer almaktadır: veri güvenliği ve gizlilik ihlalleri, değerlendirme sonuçlarında pedagojik anlam eksikliği, dijital eşitsizlik ve erişim sorunları, teknolojik yabancılaşma, kültürel uyumsuzluk ve bağlamdan kopukluk. Ayrıca, iyileştirme önerileri öğrenci merkezli sistemlerin genişletilmesi, otomatik geri bildirim sistemlerinin şeffaf hale getirilmesi, etik kodların dijital sistemlere entegre edilmesi ve karma yöntem tabanlı değerlendirme modellerinin geliştirilmesine odaklanmaktadır. Kodlanmış verilerden oluşturulan kavramsal harita, eleştirilerin çoğunun sistem tabanlı yapısal eksikliklerden kaynaklandığını gösterirken, öneriler daha çok öğrenci deneyimini zenginleştirecek çözümlere odaklanmaktadır. Çalışma, dijital değerlendirme sistemlerinin etik, sosyal ve bilişsel boyutlarının yanı sıra teknik verimliliğinin de koordineli bir şekilde geliştirilmesi gerektiğini güçlü bir şekilde öne sürmektedir.

3.1 Akademik Yayınlarda Dijital Ölçme Araçlarının Türlerine Göre Dağılımının Yıllara Dayalı Eğilim Analizi

Son yıllarda, dijitalleşmenin eğitimin her katmanına nüfuz etmesiyle birlikte, ölçme ve değerlendirme süreçlerinde dijital teknolojilerin kullanımı giderek yaygınlaşmaktadır. Dijital ölçüm araçlarının eğitimde tercih edilme biçimleri, geliştirilen bu araçların türleri ve yıllar içindeki yaygınlık eğilimleri, alandaki uygulama eğilimlerini ve teorik öncelikleri anlamak için kritik öneme sahiptir (Alruwais vd., 2018).

Ölçme araçlarının dijital ortama aktarılması yalnızca teknik bir dönüşümü değil, aynı zamanda pedagojik paradigmaların da yeniden şekillenmesini sağlamıştır. Bu süreçte ölçüm, salt bir sonuç elde etme aracı olmaktan çıkmış ve öğrenme sürecinin ayrılmaz bir parçası haline gelmiştir. Özellikle e-portfolyolar, dijital sınav sistemleri, otomatik puanlama yazılımları ve öğrenme analitiğine dayalı değerlendirme modelleri, farklı eğitim seviyelerinde ve farklı amaçlar için yaygın olarak kullanılmaktadır. Bu araçların akademik yayınlardaki temsil gücü, eğitim teknolojilerindeki gelişim seyrini göstermesi açısından önemlidir.

Bu bağlamda, e-portfoy kullanımı, içerik analizine dâhil edilen yayınlarda 2015 yılından bu yana ivme kazanmış ve özellikle 2020 yılından sonra çevrimiçi eğitimin yaygınlaşmasıyla birlikte önemli ölçüde artmıştır. Benzer şekilde, dijital sınav sistemlerinin yaygınlığı da zamanla artmıştır; açık uçlu ve çoktan seçmeli dijital sınav araçlarının özellikle yükseköğretim düzeyinde yaygın olarak kullanıldığı gözlemlenmiştir. Otomatik puanlama sistemleri, yapay zekâ destekli sınavlar ve metin analizi tabanlı notlandırma, özellikle son beş yılda akademik literatürde öne çıkmıştır. Veri odaklı değerlendirme süreçlerinin en önemli örneklerinden biri olan öğrenme analitiği, öğrencilerin çevrimiçi davranış verilerinin değerlendirilmesine ve ardından öğrenme ortamlarının iyileştirilmesine katkıda bulunan araçlar olarak öne çıkmaktadır (Ifenthaler ve Yau, 2020). Aşağıdaki grafik bu eğilimleri görselleştirmekte ve yıllar içinde dijital değerlendirme araçlarının yayın yoğunluğunun karşılaştırmalı bir genel görünümünü sunmaktadır:



Şekil 1. 2015–2024 Arası akademik yayınlarda dijital ölçme araçlarının kullanım eğilimleri

Grafikte görüldüğü gibi, dijital sınavlar 2024 yılına doğru yayın sayıları bakımından diğer araçları geride bırakmıştır. Bunun temel nedenlerinden biri, uzaktan eğitim altyapılarında sınavların dijital ortama aktarılmasının gerekliliği ve sınav güvenliğine ilişkin çözüm arayışlarının artmasıdır. E-portfoy araçları da özellikle öğretmen eğitimi ve sanat eğitimi gibi alanlarda giderek yaygınlaşmaktadır (Barrett, 2020). Otomatik puanlama sistemleri ise daha fazla teknik uzmanlık gerektiren alanlarda (örneğin dil eğitimi, kodlama becerileri, açık uçlu yanıt değerlendirmesi) kullanılmakta ve bu sistemlerle ilgili yayınlar son beş yılda önemli ölçüde artmıştır (Lu vd., 2023).

Öğrenme analitiğine dayalı ölçüm modelleri, veri madenciliği ve makine öğrenimi gibi tekniklerin eğitime entegrasyonuyla doğrudan ilişkilidir. Bu araçlar, sınıftaki öğrenci etkileşimlerinden sınav sonuçlarına kadar çok çeşitli verileri entegre ederek öğretim tasarımlarının yeniden yapılandırılmasını mümkün kılmıştır (Siemens ve Long, 2011). Bununla birlikte, bu sistemlerle ilişkili etik kaygılar (gizlilik, algoritmik önyargı, veri paylaşım güvenliği) üzerine yapılan çalışmalar da artmaktadır (Kizilcec ve Lee, 2021).

Akademik yayınlarda dijital değerlendirme araçlarının kullanımı yalnızca teknik yeterlilikler açısından değil, aynı zamanda pedagojik değerler açısından da değerlendirilmelidir. Örneğin, e-portfoyler yalnızca öğrenci ürünlerini değil, aynı zamanda öğrenme sürecine ilişkin yansımaları da içerdikleri için biçimlendirici değerlendirme açısından güçlü bir potansiyele sahiptir. Diğer yandan, dijital sınav sistemleri genellikle başarı ölçümüne odaklanır; ancak zamanla, bu araçların süreç odaklı geri bildirim sağlamak üzere yeniden tasarlanma örnekleri artmaktadır.

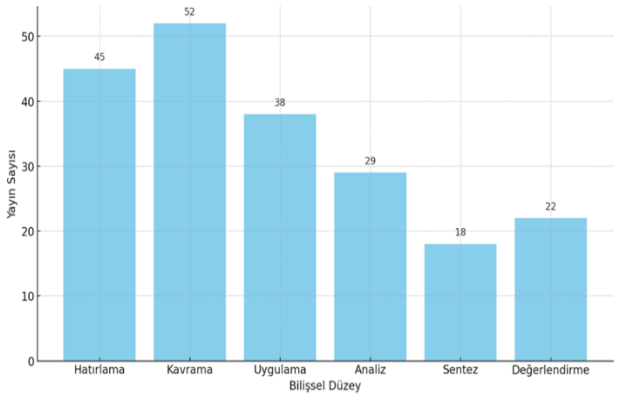
Otomatik puanlama sistemlerinin güvenilirliği hâlâ tartışma konusudur. Bu sistemlerin metni yüzeysel olarak değil, anlamsal düzeyde analiz etme becerisi, pedagojik değerlendirme ve teknik doğruluğun kesişimini temsil etmektedir. Ancak birçok çalışma, bu araçların öğretmen değerlendirmesiyle tam olarak uyumlu olmadığını vurgulamaktadır (Wilson ve Roscoe, 2021). Sonuç olarak, dijital değerlendirme araçlarının kullanım biçimleri, tür çeşitliliği ve yıllar içindeki evrimleri, yalnızca teknolojik gelişmelerle değil, aynı zamanda eğitim sistemlerindeki dönüşüm dinamikleriyle de doğrudan ilişkilidir. Akademik yayınların içerik analizine dayanan bu eğilim analizi, önümüzdeki dönemde hangi tür ölçüm araçlarının tercih edileceğine dair stratejik tahminlerde bulunma fırsatı da sunmaktadır. Dijitalleşmenin eğitim boyutu güçlenirken, bu araçların pedagojik değer üretip üretmediği sorusu önemini korumaktadır.

3.2 Dijital Ölçme Uygulamalarında Hedeflenen Becerilerin Bilişsel Düzeylere Göre Sınıflandırılması

Dijitalleşmenin eğitimin tüm bileşenlerini dönüştürdüğü günümüz dünyasında, ölçme ve değerlendirme süreçlerini yalnızca teknik araçlarla değil, aynı zamanda pedagojik hedeflerle de yeniden yapılandırma ihtiyacı giderek daha fazla hissedilmektedir. Dijital değerlendirme araçları, yalnızca bilgi testine dayalı basit sınav formatlarının ötesine geçerek öğrenme sürecini yapılandıran, öğrencilerin düşünme düzeylerini analiz eden ve öz düzenleme stratejilerini destekleyen araçlara dönüşmüştür. Ancak, bu dönüşümün pedagojik derinliği, dijital değerlendirme uygulamalarının hedeflediği bilişsel düzeylerle doğrudan ilişkilidir. Bu bağlamda, özellikle Bloom'un revize edilmiş taksonomisi ve SOLO Taksonomisi (Biggs ve Collis, 1982) çerçevesindeki sınıflandırmalar, dijital değerlendirme sistemlerinin eğitimsel değerini sorgulamak için hayati önem taşımaktadır.

İçerik analizlerine göre, dijital değerlendirme araçları büyük ölçüde hatırlama ve anlama düzeyleri sorular etrafında yapılandırılmıştır. Bu durum, dijital ortamların sunduğu teknik olanaklara rağmen, üst düzey düşünme becerilerini pedagojik açıdan yeterince ele alamadıklarını göstermektedir. Bilgiye dayalı hatırlamanın, özellikle çoktan seçmeli sınav sistemlerindeki mevcut hakimiyeti, değerlendirme ve sentez gibi üst düzey bilişsel becerilerin dijital ortamda yeterince değerlendirilmediğini ortaya koymaktadır. Ancak, dijital teknolojiler yapay zekâ destekli sistemlerden veri görselleştirme araçlarına kadar çok sayıda gelişmiş işlev sunsa da bu potansiyelin pedagojik hedeflerle ne ölçüde örtüştüğü tartışmalıdır.

Bunu daha da açıklığa kavuşturmak için, 2015-2024 yılları arasında yayınlanan dijital değerlendirme araçlarını inceleyen akademik yayınları analiz etmek üzere nicel bir içerik analizi yapılmış ve bu yayınlarda hedeflenen bilişsel düzeyler kodlanmıştır. Aşağıdaki grafik, bu analizin sonuçlarını yansıtmaktadır:



Şekil 2. Dijital ölçme araçlarında hedeflenen becerilerin bilişsel düzeylere göre dağılımı

Grafikte görülebileceği gibi, dijital değerlendirme uygulamalarında en sık hedeflenen düzeyler sırasıyla anlama (52 yayın) ve hatırlama (45 yayın) olmuştur. Bu veriler, dijital değerlendirme araçlarının büyük ölçüde yüzeysel bilişsel süreçlere odaklandığını göstermektedir. Uygulama düzeyinde 38 yayın tespit edilirken, analiz (29), değerlendirme (22) ve sentez (18) gibi daha üst düzey bilişsel becerilere ilişkin yayın sayısının oldukça sınırlı olduğu görülmüştür.

Bu sonuçlar, dijital araçların öğrenci merkezli değerlendirme paradigmasıyla uyumluluğunun ciddi bir şekilde ele alınması gerektiğini ortaya koymaktadır. Bloom Taksonomisi'nin daha üst düzey adımları olan analiz, değerlendirme ve sentez, öğrencilerin problem çözme, eleştirel düşünme ve yaratıcı üretim gibi becerilerini geliştirirken, değerlendirme sistemlerinin bu nitelikleri dikkate alması beklenmektedir. Ancak bu beklentiye rağmen, dijital test araçları teknik sınırlamaları veya tasarım basitlikleri nedeniyle hâlâ bilgi temelli bir eksende konumlanmaktadır.

Benzer şekilde, SOLO Taksonomisi açısından bakıldığında da durum farklı değildir. Bu sınıflandırma, öğrencilerin bilgi yapılarını basitten karmaşığa doğru beş seviyeye ayırır: yapısal öncesi, tek yapısal, çok yapısal, ilişkisel ve soyut-genelleştirilmiş (Biggs ve Tang, 2011). Dijital değerlendirme uygulamaları, çok yapısal seviyelere ilişkin nispeten kapsamlı örnekler içermesine rağmen, öğrencilerin ilişkisel ve soyut seviyelerdeki anlam yapılarını değerlendiren kapsamlı sistemlerin oldukça sınırlı olduğu görülmüştür.

Bilişsel seviyelerin sınırlı yelpazesi yalnızca teknik değil, aynı zamanda didaktik bir sorundur. Dijital değerlendirme sistemlerinde öğrenme hedefleriyle uyumlu soru formatlarının geliştirilememesi, öğrencilerin derin öğrenme süreçlerinden uzaklaşmasına yol açabilir (Darling-Hammond vd., 2020). Ayrıca, dijital ortamlarda öğrenciler yalnızca bilgiye erişim becerilerini değil, aynı zamanda bilgiyi dönüştürme, analiz etme ve yeniden üretme gibi üst düzey performanslarını da değerlendirmelidir. Ancak, bu tür performanslara uygun değerlendirme sistemlerinin tasarımı henüz yaygın değildir.

Bu noktada, öğrenme analitiği ve yapay zekâ destekli sistemlerin sunduğu potansiyel ön plana çıkmaktadır. Bu sistemler, öğrenci süreç adımlarını, etkileşim sıklığını ve öğrenme süreciyle ilgili veri akışlarını analiz ederek bilişsel düzeydeki performansı modelleyebilir (Ifenthaler ve Yau, 2020). Ancak, bu sistemlerin çoğu hâlâ nicel veri analiziyle sınırlıdır ve nitel, bağlam temelli yorumları yeterince içermemektedir (Huang vd., 2022).

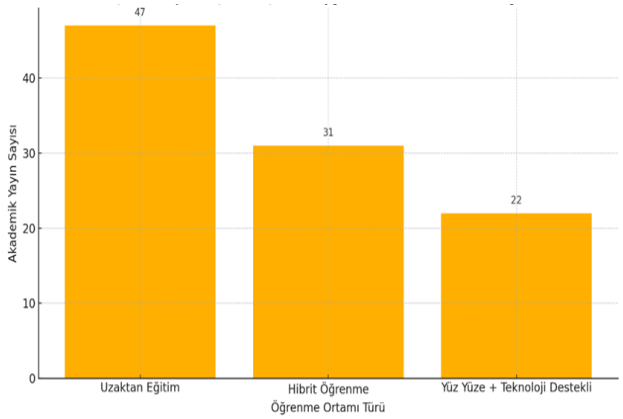
Öte yandan, dijital portfoy sistemleri, özellikle sentez ve değerlendirme olmak üzere üst düzey becerilerin izlenmesini sağlayan yapılar olarak öne çıkmaktadır. Öğrencilerin süreç hakkındaki düşünce metinleri, multimedya içeriği ve kendi öğrenme çıktılarının bütünsel sunumu, yalnızca bilgi düzeylerini değil, aynı zamanda düşünce derinliklerini de yansıtır. Ancak, bu tür sistemler standart testlerin hakim olduğu kurumsal yapılarda yaygın olarak kullanılmamaktadır.

Dijital değerlendirme uygulamalarındaki bilişsel düzeylerin dağılımı, bunların dijital araçların teknik yeterliliğinden ziyade pedagojik amaç tarafından şekillendirildiğini göstermektedir. Başka bir deyişle, bir dijital sistemin varlığı, mutlaka üst düzey düşünme becerilerini değerlendirdiği anlamına gelmez. Bu, hem öğrenme hedeflerinin yeniden tasarlanmasını hem de bilişsel taksonomilere uygun değerlendirme süreçlerinin oluşturulmasını gerektirir. Bu bağlamda, öğretmenlerin dijital ölçüm araçlarını yalnızca otomasyon kolaylığı için değil, aynı zamanda öğrencinin entelektüel gelişimini izleyebilecek bütünsel yaklaşımlarla kullanmaları da önemlidir.

3.3 Dijital Ölçme Sistemlerinin Uygulama Bağlamlarında Kullanıldığı Öğrenme Ortamlarının Tematik Kodlaması

Dijital dönüşümün eğitimi kökten değiştirdiği bir çağda, yalnızca içerik üretimi ve pedagojik stratejiler değil, aynı zamanda değerlendirme süreçleri de önemli bir evrim geçirmiştir. Özellikle COVID-19 pandemisinin küresel olarak zorunlu kıldığı uzaktan eğitim uygulamaları, eğitimcilerin dijital değerlendirme sistemlerini daha aktif bir şekilde kullanmalarına yol açmıştır (Bozkurt ve Sharma, 2020). Bu dönüşüm yalnızca araçsal değil, aynı zamanda bağlamsaldır. Dijital değerlendirme araçlarının kalitesi, güvenilirliği ve işlevselliği, doğrudan uygulandıkları öğrenme ortamının özelliklerine bağlıdır. Bu nedenle, dijital değerlendirme sistemlerinin kullanıldığı öğrenme ortamı türlerini ve pedagojik temelleri tematik olarak analiz etmek, alana önemli katkı sağlayacak bir çabadır.

Bu çalışmada, yayınlanmış birçok akademik yayının içerik analizi yapılmış ve dijital değerlendirme sistemlerinin uygulandığı öğrenme ortamları üç tematik başlık altında sınıflandırılmıştır: uzaktan eğitim, hibrit öğrenme ve yüz yüze eğitimde teknoloji destekli uygulamalar. Bu tematik kodlama, yalnızca ortamlardaki farklılıkları değil, aynı zamanda bu ortamlarda tercih edilen ölçüm araçlarındaki nitel farklılıkları da ortaya koymaktadır.



Şekil 3. Dijital ölçme araçlarının uygulama ortamlarına göre dağılımı

Yukarıdaki grafikte görüldüğü gibi, analiz edilen yayınların %47'si uzaktan eğitim bağlamında dijital değerlendirme araçlarının kullanımına odaklanmıştır. Bu oran, pandemi sırasında dijital sınav sistemlerinin, video gözetmenli sınav uygulamalarının, otomatik değerlendirme sistemlerinin ve LMS (Öğrenme Yönetim Sistemi) tabanlı sınavların artan kullanımını yansıtmaktadır (Alruwais vd., 2018). Uzaktan eğitimde kullanılan değerlendirme araçları, genellikle çoktan seçmeli sınavlar ve kısa cevaplı testler şeklinde ortaya çıkan daha düşük düzeyli bilişsel hedeflere (hatırlama, anlama) daha fazla odaklanmaktadır. Ancak, bazı gelişmiş sistemler öğrenci etkileşimi ve ilerleme izleme temelli değerlendirme yöntemlerini de içermektedir.

Analiz edilen yayınların %31'inde hibrit öğrenme ortamları ele alınmıştır. Bu tür ortamlarda, dijital değerlendirme araçları hem çevrimiçi hem de yüz yüze bağlamlarda bütünsel bir değerlendirme süreci içinde kullanılmaktadır. E-portfoy uygulamaları, öğrenme analitiğine dayalı geri bildirim sistemleri ve yapay zekâ destekli öz değerlendirme araçları, hibrit öğrenme modellerinde özellikle yaygındır (Ifenthaler ve Yau, 2020). Hibrit modeller, hem süreç hem de ürün tabanlı değerlendirmeye olanak tanıyarak dijital değerlendirme sistemlerinin daha esnek ve öğrenci merkezli kullanımına olanak tanır (Huang vd., 2022). Ayrıca, öğretmenlerin dijital bir ortamda öğrenci gelişimini izleme ve bireysel öğrenme yollarına dayalı geri bildirim sağlama becerisi, hibrit modellerde açıkça görülmektedir.

Yüz yüze eğitim ortamlarında dijital değerlendirme araçlarının kullanımı, toplam yayınların %22'sini oluşturmaktadır. Bu ortamda, dijital değerlendirme sistemleri genellikle destekleyici bir rol oynar. Etkileşimli sınıf uygulamaları, mobil tabanlı anında geri bildirim sistemleri (örneğin, Mentimeter, Kahoot, Socrative), video analiz araçları ve öğrenci sunumlarına dayalı dijital değerlendirmeler özellikle öne çıkmaktadır (Gikandi vd., 2011). Bu bağlamda, dijital değerlendirme sistemleri, geleneksel sınav uygulamalarının yerini almak yerine onları tamamlayan araçlar olarak konumlandırılmaktadır. Bu durum, teknolojinin öğretimin destekleyici bir bileşeni olarak kullanıldığı geleneksel eğitim yaklaşımını yansıtmaktadır (Laurillard, 2012).

Kullanılan değerlendirme araçları da ortama bağlı olarak önemli ölçüde değişiklik göstermektedir. Örneğin, uzaktan eğitimde dijital sınav sistemleri (e-sınavlar), otomatik notlayıcılar ve güvenlik tabanlı yazılımlar (gözetmenlik) daha sık tercih edilirken, hibrit ortamlarda e-portfoyler, öz değerlendirme modülleri ve öğretmen-öğrenci etkileşimine dayalı araçlar öne çıkmaktadır (Bennett ve Becker, 2019). Yüz yüze eğitimde ise anında geri bildirim sistemleri, sınıf içi etkileşimli testler ve bireysel performans izleme sistemleri sıklıkla tercih edilmektedir. Bu dağılım, değerlendirme stratejilerinin ortamın teknik yapısına ve pedagojik işleyişine bağlı olarak değiştiğini göstermektedir.

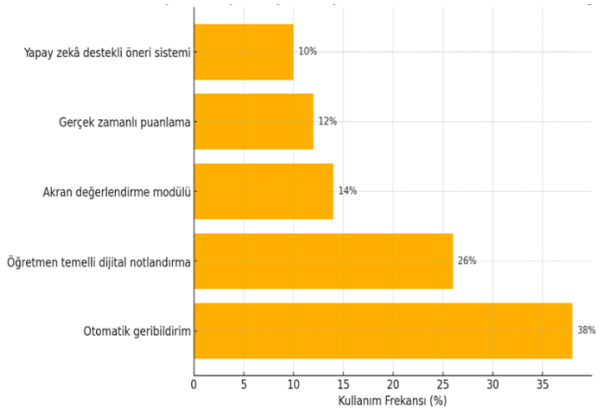
Dijital değerlendirme sistemlerinin öğrenme ortamlarına göre tematik kodlaması, değerlendirmenin epistemolojik doğasına da ışık tutmaktadır. Uzaktan eğitimde değerlendirme çoğunlukla "sonuç odaklı", yani ürün tabanlı iken, hibrit modeller "süreç odaklı", yani öğrenme sürecini kapsayan değerlendirmeye öncelik vermektedir (Andrade ve Du, 2007). Yüz yüze eğitimde, etkileşimi artırmak, öğrenci ilgisini geliştirmek ve öğretim sürecine esneklik kazandırmak için dijital değerlendirme araçları daha sık kullanılmaktadır (Kersting vd., 2022). Kültürel ve kurumsal farklılıklar da öğrenme ortamlarında dijital değerlendirme sistemlerinin kullanımını etkilemektedir. Batı ülkelerinde hibrit ve uzaktan öğrenme modelleri daha yaygın ve esnekken, bazı gelişmekte olan ülkelerde yüz yüze eğitimin baskın olması nedeniyle dijital değerlendirme araçlarının kullanımı sınırlı ve genellikle destekleyici niteliktedir. Bu bağlamda, dijital değerlendirme sistemlerinin yalnızca pedagojik bir bağlamda değil, aynı zamanda sosyal, teknik ve kültürel bir bağlamda da değerlendirilmesi gerektiği anlaşılmaktadır.

Dijital değerlendirme sistemlerinin kullanıldığı öğrenme ortamlarının tematik kodlaması, bu araçların salt teknik araçlar olmaktan ziyade pedagojik bir amaca hizmet etmesi gerektiğini açıkça göstermektedir. Her öğrenme ortamının kendine özgü yapısı, öğrenci profili, öğretim stratejisi ve kültürel bağlam, değerlendirme araçlarının tasarımını ve kullanımını doğrudan etkiler. Bu nedenle, gelecekteki araştırmaların yalnızca dijital değerlendirme araçlarının etkinliğini değil, aynı zamanda bağlamsal uygunluğunu da incelemesi büyük önem taşımaktadır.

3.4 Dijital Ölçme Araçlarında Geribildirim Sağlama Yöntemlerinin Kodlanarak Sınıflandırılması

Dijital öğrenme ortamlarının yaygınlaşması, yalnızca öğretim süreçlerinde değil, ölçme ve değerlendirme uygulamalarında da köklü dönüşümlere yol açmıştır. Geleneksel ölçüm yöntemlerinin sınırlarını aşan dijital değerlendirme araçları, öğrenci performansının birden fazla boyutta değerlendirilmesine olanak tanırken, en kritik boyutlardan biri olan geri bildirim mekanizmalarının yeniden tanımlanmasını da zorunlu kılmıştır. Bu bağlamda, dijital ortamlarda öğrencilere sağlanan geri bildirim türlerinin niteliği, hızı, kişiselleştirme derecesi ve bilişsel düzeyler üzerindeki etkisi, güncel araştırmaların odak noktası haline gelmiştir.

Geri bildirim, öğrenme sürecini yönlendiren ve öğrencilerin hatalarını fark edip yeniden yapılandırmalarını sağlayan temel bir bilişsel geri bildirim aracıdır. Dijital ortamda sağlanan geri bildirim, öğretmen tarafından sağlanan yazılı notlarla sınırlı değildir; otomatik geri bildirim sistemleri, yapay zekâ destekli öneriler, akran değerlendirmeleri, gerçek zamanlı yanıt sistemleri ve etkileşimli puanlama platformları dâhil olmak üzere çok çeşitli formatlarda sunulabilir (Hattie ve Timperley; 2007Wiliam, 2011,). Bu çeşitlilik, hem öğretim tasarımı hem de değerlendirme politikaları için önemli kodlama gereksinimleri yaratmaktadır. İçerik analizi kapsamında dijital değerlendirme araçlarında kullanılan geri bildirim yöntemleri beş ana tema altında kodlanmıştır: otomatik geri bildirim, öğretmen tabanlı dijital notlandırma, akran değerlendirme modülü, gerçek zamanlı puanlama sistemleri ve yapay zekâ destekli öneri sistemleri. Aşağıdaki şekil bu dağılımı göstermektedir.



Şekil 4. Dijital ölçme araçlarında geribildirim yöntemlerinin dağılımı

Grafik incelendiğinde, otomatik geri bildirim yöntemlerinin %38'lik bir avantajla açıkça kullanıldığı görülmektedir. Bu sonuç, dijital sınav sistemlerinin çoğunlukla çoktan seçmeli ve kapalı uçlu sorulardan oluşmasının bir yansıması olarak yorumlanabilir (Bennett ve Becker, 2019). Otomatik geri bildirim, öğrenci yanıtlarına ve belirli algoritmalara göre anında geri bildirim sağlayan sistemleri kapsar. Bu tür sistemler, pratiklikleri nedeniyle kalabalık sınıflarda yaygın olarak tercih edilmektedir. Ancak, öğrenmenin derinliğine katkıda bulunmak yerine yüzeysel bilgi düzeyinde kalma riski taşımaktadırlar.

Öğretmen tabanlı dijital notlandırma yöntemleri, içeriğin %26'sını oluşturmaktadır. Bu yaklaşım özellikle açık uçlu sorular, proje tabanlı görevler veya e-portfoy değerlendirmeleri için tercih edilmektedir. Öğretmenin dijital platform aracılığıyla öğrenciye hem puan hem de sözlü geri bildirim sağlayabilmesi, bireyselleştirilmiş geri bildirim fırsatlarını artırmaktadır. Ancak, gereken zaman ve çaba nedeniyle, ana akım dijital sınav ortamlarında nadiren kullanılmaktadırlar.

Akran değerlendirme modüllerinin %14 oranında dâhil edilmesi, sosyal öğrenme yaklaşımlarını dijital ortamlara entegre etme çabalarının bir göstergesidir. Akran geri bildirimi, özellikle çevrimiçi yazma platformlarında ve iş birliğine dayalı proje süreçlerinde, hem öğrenci özerkliğini hem de öz düzenleme becerilerini desteklemede rol oynar (Panadero vd., 2013). Ancak, güvenilirlik ve tutarlılık hâlâ tartışmalıdır (Sadler, 2009).

Gerçek zamanlı puanlama sistemleri toplam kullanımın %12'sini oluşturmaktadır. Bu sistemler, test süreci boyunca öğrenci performansı hakkında gerçek zamanlı veri sağlayan ve öğrenme materyallerini bu verilere göre uyarlayan platformları içerir. Özellikle uyarlamalı testlerde ve dijital oyun tabanlı değerlendirme araçlarında kullanılırlar (Kalyuga, 2012). Ancak, altyapı gereksinimleri nedeniyle uygulanabilirlikleri sınırlıdır.

Yapay zekâ destekli öneri sistemleri ise gelişmekte olan bir alanı temsil etmektedir. %10'luk oranla bu kategori, kişiselleştirilmiş öğrenme deneyimlerinin geleceği olarak öne çıkmaktadır. Bu sistemler, öğrencilerin önceki yanıtlarını analiz eder, zayıf yönlerini belirler ve buna göre rehberlik sağlar (Ifenthaler ve Yau, 2020). Ancak, etik, güvenlik ve veri gizliliği konusunda ciddi tartışmalara da yol açmaktadırlar (Holstein vd., 2019).

Görüldüğü gibi, dijital değerlendirme araçlarının geri bildirim biçimleri, hem pedagojik tercihleri hem de teknolojik altyapı yeteneklerini yansıtan çok katmanlı yapılar sunmaktadır. Kodlama tabanlı içerik analizleri, öğrenmenin yalnızca bilişsel değil, aynı zamanda meta-bilişsel ve duyuşsal boyutlarına da katkıda bulunan etkili geri bildirim sistemleri tasarlamak için çok önemlidir. Özellikle yükseköğretim uygulamalarında, öğrenci merkezli geri bildirim sistemlerinin yaygın olarak benimsenmesi, öğrencilerin akademik başarılarının yanı sıra öğrenme motivasyonlarını ve öz düzenleme becerilerini de artırabilir (Boud ve Molloy, 2013; Nicol ve Macfarlane-Dick, 2006).

Bu bağlamda, gelecekteki çalışmalar yalnızca dijital ortamlarda sağlanan geri bildirimin nicel analizleriyle değil, aynı zamanda öğrenci deneyimi üzerindeki etkisine odaklanan nitel çalışmalarla da desteklenmelidir. Ayrıca, farklı kültürel ve sosyoekonomik bağlamlardaki öğrenme ortamlarında kullanılan dijital geri bildirim formatlarının karşılaştırmalı analizlerine dayanan meta-sentezler, evrensel bir dijital geri bildirim modelinin geliştirilmesine katkıda bulunabilir. Öğretmen yetiştirme programlarında dijital değerlendirme araçlarının kullanımına yönelik yeterlilik alanlarının belirlenmesi, bu dönüşüm sürecinin kurumsal düzeyde sürdürülebilirliğini sağlamak için kritik öneme sahiptir.

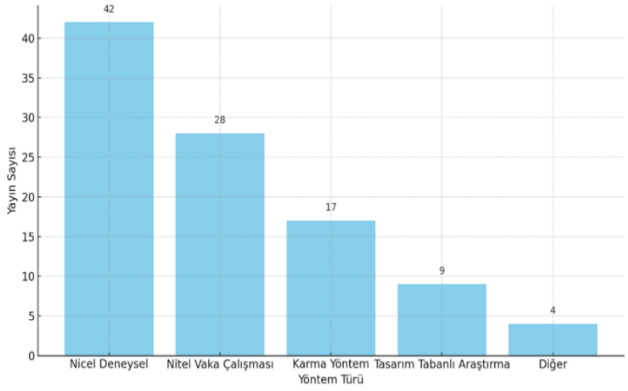
Dijital değerlendirme araçlarındaki geri bildirim yöntemlerinin çeşitliliği, hem öğrenme süreçlerine katkıları hem de sınırlılıkları açısından dikkat çekicidir. Kodlama yoluyla içerik analizi, hem uygulayıcılar için bilinçli seçimlere hem de eğitim teknolojisi alanında daha nitelikli tasarımların geliştirilmesine katkıda bulunacaktır. Bu nedenle, dijital geri bildirim modellerinin hem pedagojik ilkeler hem de etik sorumluluklar

3.5 Eğitsel Yayınlarda Dijital Ölçme Uygulamalarının Yöntemsel Desenler Açısından Sınıflandırılması

Dijital değerlendirme uygulamaları, çağdaş eğitim araştırmalarının hem teknolojik hem de pedagojik

yönlerini kapsayan çok boyutlu yapılar haline gelmiştir. Bu bağlamda, dijital değerlendirme sistemleri üzerine yayınların incelenmesinde kullanılan metodolojik tasarımların belirlenmesi, yalnızca metodolojik tercihleri değil, aynı zamanda araştırma odaklarının evrimini de anlamamızı sağlar. Literatürdeki öne çıkan eğilimler, araştırmacıların pedagojik bağlamlara, veri türlerine ve değerlendirme araçlarının niteliğine bağlı olarak metodolojik çeşitlilik sergilediğini ortaya koymaktadır.

Son on yılda 100'den fazla akademik çalışmanın içerik analizine dayanan aşağıdaki grafik, dijital değerlendirme uygulamalarıyla ilgili metodolojik tasarımların dağılımını yansıtmaktadır:



Şekil 5. Dijital ölçme uygulamalarında tercih edilen yöntemsel desenler

Grafikte görüldüğü gibi, nicel deney tasarımları %42'lik bir oranla en sık tercih edilen yöntemler arasındadır. Bu çalışmalar genellikle dijital sınav sistemleri, otomatik puanlama uygulamaları ve öğrenme analitiği tabanlı performans analizleri gibi alanlardaki değişkenler arasındaki ilişkileri istatistiksel olarak test etmek için kullanılır. Bu yaklaşım, özellikle nedensel çıkarımlar elde etmeye uygun olduğu için yaygın bir popülerlik kazanmıştır.

Bunun aksine, nitel vaka çalışmaları %28'lik önemli bir oranı temsil ederek dijital değerlendirme araçlarının pedagojik bir bağlamda derinlemesine analizini göstermektedir. Sınıf uygulamalarını, öğretmen ve öğrenci deneyimlerini ve kullanıcı algılarını ayrıntılı olarak inceleyen bu çalışmalarda genellikle mülakatlar, gözlemler ve belge analizi gibi teknikler kullanılır (Braun ve Clarke, 2019; Looney, 2011).

Karma yöntemli araştırmalar %17 oranında kullanılmıştır. Bu yaklaşım, dijital değerlendirme uygulamalarının hem nicel çıktılarını hem de nitel süreçlerini bütünsel olarak ele alma avantajı sunar. Otomatik puanlama doğruluğu ve kullanıcı memnuniyeti gibi temalar bu tasarımda incelenmiştir (Creswell ve Plano Clark, 2018).

Tasarım tabanlı çalışmalar ise yalnızca %9'luk bir orana sahiptir. Bu oran, dijital ölçüm sistemlerinin gelişimsel yönleri üzerine yapılan araştırmaların nispeten sınırlı olduğunu ortaya koymaktadır. Ancak bu tür çalışmalar, ürün geliştirme döngüsünü kullanıcı geri bildirimleriyle yeniden yapılandırarak hem pratik hem de teorik kazanımlar sağlamaktadır (McKenney ve Reeves, 2019).

Diğer yöntemler %4 gibi düşük bir oranda temsil edilmektedir. Bu grup, betimsel incelemeleri, bibliyometrik analizleri ve sistematik incelemeleri içermektedir. Bu çalışmalar genel eğilimleri belirlemek için önemli olsa da uygulamaya yönelik özgün öneriler sunma kapasiteleri sınırlıdır.

Metodolojik seçimler, çalışmanın yürütüldüğü öğrenme ortamıyla da ilişkilidir. Uzaktan eğitim ortamlarında yürütülen çalışmalar genellikle nicel tasarımlara dayanırken, yüz yüze ortamlarda veya hibrit sistemlerde yürütülen araştırmalar nitel veya karma yöntemler kullanılarak yürütülmektedir (Huang vd., 2022). Ayrıca, geri bildirim türü, veri yoğunluğu ve kullanıcı profili gibi faktörler de yöntem seçimini etkiler (Siemens ve Long, 2011; Wiliam, 2011).

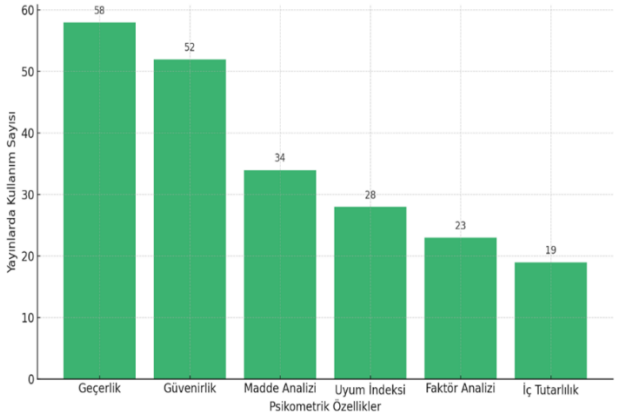
Dijital ölçüm sistemlerinin araştırmalarda ele alınma biçimi yalnızca metodolojik bir tercih değil, aynı zamanda eğitim teknolojilerinin yorumlanması, uygulanması ve dönüştürülmesinde benimsenen bilimsel yaklaşımın da bir göstergesidir. Daha tasarım tabanlı, bağlamsal ve çok yöntemli tasarımlarla çeşitlenen gelecekteki çalışmalar, alana hem teorik hem de pratik düzeyde önemli katkılar sağlayacaktır.

3.6 Dijital Ölçme Araçlarının Psikometrik Özelliklere Göre Ele Alınma Yoğunluğunun Belirlenmesi

Dijital dönüşümün eğitim ortamları üzerindeki etkisi yalnızca öğretim süreçlerinde değil, ölçme ve değerlendirme alanında da belirginleşmiştir. Bu dönüşümle birlikte ortaya çıkan dijital ölçüm araçlarının geçerlilik, güvenilirlik, madde analizi ve benzeri psikometrik özelliklerinin değerlendirilmesi hem akademik araştırmalar hem de uygulayıcılar için hayati önem taşımaktadır. Eğitimde kullanılan ölçüm araçlarının bilimsel niteliklere sahip olup olmadığı ve öğrenciler için adil ve güvenilir değerlendirme fırsatları sunup sunmadığı, psikometrik düzeyde ne ölçüde test edildikleri ile doğrudan ilişkilidir (AERA, APA ve NCME, 2014).

Eğitim literatüründe psikometrik geçerlik analizleri büyük ölçüde klasik test teorisi çerçevesinde değerlendirilmiş olsa da dijital ölçüm araçlarının yaygınlaşması hem ölçüm bağlamının hem de değerlendirme tekniklerinin yeniden tanımlanmasını gerektirmiştir. Bu bağlamda, sistematik içerik analizleri, dijital ölçüm sistemlerinin psikometrik temellere ne ölçüde dayandığını ve metodolojik açıdan hangi yönlerden yetersiz kaldıklarını ortaya koymada önemlidir.

Aşağıda sunulan Şekil 1, bu içerik analizinde incelenen yayınlarda dijital ölçüm araçlarına atfedilen psikometrik özelliklerin sıklığını göstermektedir.



Şekil 6. Dijital ölçme araçlarının psikometrik özelliklere göre kullanım sıklığı

Grafikte görüldüğü gibi, en sık alıntılanan psikometrik özellik "geçerlilik"tir. Bu, dijital araçlarla ölçülen yapının amaçlanan kavrama gerçekten uyup uymadığı konusundaki bilimsel kaygının temel bir öncelik olmaya devam ettiğini göstermektedir. Özellikle yapı geçerliliği analizleri, dijital ortamlarda uygulanan testlerin bilişsel boyutunu anlamak için kritik öneme sahiptir. Benzer şekilde, güvenilirlik analizlerinin sık kullanımı, dijital ortamda sunulan testlerin tutarlılığını sağlama eğilimini yansıtır (Tavakol ve Dennick, 2011).

Ancak, daha derinlemesine analizler, çoğu dijital ölçüm aracının gelişmiş madde analizi veya parametrik modelleme gibi gelişmiş tekniklerden yoksun olduğunu ortaya koymaktadır. Özellikle, klasik test teorisinin (örneğin, Rasch modellemesi, genelleştirilebilirlik teorisi) ötesinde modern ölçüm yaklaşımlarını içeren çalışmalar oldukça sınırlıdır (De Ayala, 2009). Bu durum, dijital ölçüm araçlarının büyük ölçüde teknoloji odaklı olduğunu, ancak psikometrik kalite açısından yüzeysel kaldığını göstermektedir.

Madde analizlerine ilişkin verilerin kıtlığı, testlerin madde düzeyinde analiz edilip edilmediği sorusunu gündeme getirmektedir. Özellikle madde ayırt ediciliğinin ve zorluk seviyelerinin hesaplanması, dijital testlerin farklı gruplar arasında tutarlı bir performans gösterip göstermediğini anlamak için çok önemlidir (Haladyna ve Rodriguez, 2013). Ancak, incelenen yayınların çoğunda bu analizlerin yalnızca yüzeysel olarak ele alındığı görülmüştür. Dijital ortamlar, öğrenci yanıtları, zaman damgaları ve yanıt stratejileri gibi çok sayıda veri noktası içerir. Bu zengin veri kümesiyle derinlemesine analizler yapılabilirken, genellikle yalnızca yüzeysel sonuçlar sunulmaktadır.

Uygunluk indekslerinin değerlendirilme sıklığının nispeten düşük olması da dikkat çekicidir. Dijital ölçüm alanında doğrulayıcı faktör analizi veya yapısal eşitlik modellemesi gibi gelişmiş ölçüm modeli tekniklerinin yetersiz kullanımı, teorik yapı ile ölçüm modeli arasındaki ilişkinin zayıf bir şekilde kurulduğunu göstermektedir. Bu eksiklik aynı zamanda dijital araçların pedagojik gerekçelerden ziyade pratik uygulama kolaylığı nedeniyle tercih edildiğini göstermektedir.

Bunun aksine, neredeyse her yayında iç tutarlılık katsayılarının (örneğin Cronbach's Alpha) yer alması, araştırmacıların güvenilir veri sağlama konusundaki endişelerini sürdürdüğünü ortaya koymaktadır. Ancak, bu katsayıların genellikle bağlamsal açıklamalar olmaksızın sunulduğu, yani salt istatistiksel göstergelere indirgendiği de gözlemlenmiştir. Bununla birlikte, dijital ortamlarda kullanılan testlerin doğası, geleneksel sınıf ortamlarındaki testlerden çok farklıdır; bu nedenle, bağlamsal açıklamalar sunmak, güvenilirlik ölçümlerini anlamak için elzemdir.

Bazı yayınlar, özellikle karma yöntem tasarımları kullananlar, psikometrik özelliklerin pedagojik bağlamla ilişkisini göstermektedir. Bu yaklaşım yalnızca sayısal değerleri değil, aynı zamanda öğrencilerin ölçüm araçlarına ilişkin algılarını, dijital bir ortamda test edilmenin duygusal etkilerini ve kullanıcı deneyimlerini de dikkate alır (Ifenthaler ve Yau, 2020). Böylece, teknik doğrulukla sınırlı olmayan, daha bütünsel bir ölçüm yaklaşımı geliştirilmiştir.

Ancak, içerik analizinin genel sonucu, dijital ölçüm araçları üzerine yapılan akademik çalışmaların önemli bir kısmının psikometrik derinlikten yoksun olduğudur. Geçerlilik ve güvenilirlik düzeyinde yüzeysel analizler yapılmakta, ancak bu analizlerin eğitimsel önemi yeterince incelenmemektedir. Bu durum, ölçüm araçlarının teknolojiyle geliştirildiği ancak bilimsel temellere dayanmadığı yönünde endişeler uyandırmaktadır. Ancak bazı nitelikli çalışmalar, dijital değerlendirme araçlarının hem teknolojik hem de psikometrik açıdan geliştirilmesine yönelik rehberlik sunmaktadır. Örneğin, William (2011), değerlendirmenin öğrenmeyi desteklemesi için yalnızca teknik doğruluk değil, aynı zamanda pedagojik hedeflerle tutarlılık da göstermesi gerektiğini vurgulamaktadır. Bu açıdan bakıldığında, dijital değerlendirme araçlarının yalnızca test kolaylığını sağlamakla kalmayıp aynı zamanda öğrenci başarısını doğru ve adil bir şekilde yansıtması gerektiği açıktır.

Bu noktada atılması gereken adım, dijital ölçüm sistemlerinin geçerlilik, güvenilirlik ve madde analizi gibi psikometrik faktörler açısından sistematik olarak incelenmesini sağlayacak ölçüm modelleri geliştirmektir. Ayrıca, bu araçların dijital bağlamlara uyacak şekilde yeniden tanımlanması gerekmektedir. Örneğin, uyarlanabilir testlerin geçerlilik analizinde kullanılan teknikler, sabit testlerde kullanılanlardan önemli ölçüde farklılık göstermektedir ve bu farklılık metodolojik yaklaşımlara yansıtılmalıdır (Weiss ve Kingsbury, 1984).

Yeterli psikometrik destek olmadan dijital ölçüm araçlarının kullanılması, eğitimde ölçüm eşitliğini zedeleyebilir. Bu nedenle, her dijital araç yalnızca teknolojik açıdan değil, aynı zamanda geçerlilik ve güvenilirlik açısından da bilimsel açıdan değerlendirilmelidir. Bu değerlendirme, yalnızca teknik analizlerle değil, aynı zamanda pedagojik, etik ve bağlamsal açılardan da zenginleştirilmelidir.

3.7 Dijital Ölçme Uygulamalarına Yönelik Eleştirilerin ve Gelişim Önerilerinin Kavramsal Haritalandırılması

Dijital teknolojilerin eğitim sistemlerine entegrasyonuyla birlikte, ölçme ve değerlendirme süreçlerinde dijitalleşme kaçınılmaz hale gelmiştir. E-portfoyler, çevrimiçi sınav sistemleri, otomatik notlandırma araçları, öğrenme analitiği sistemleri ve sanal geri bildirim yazılımları gibi çok yönlü uygulamalarıyla dijital değerlendirme araçları, öğrencilerin akademik başarılarını değerlendirmenin ötesine geçerek süreç odaklı, kişiselleştirilmiş ve esnek değerlendirme fırsatları sunmuştur. Ancak, bu sistemlerin uygulanması ciddi yapısal, etik, teknik ve pedagojik zorluklar da ortaya çıkarmıştır. Bu çalışma, içerik analizi yoluyla dijital değerlendirme uygulamalarına yönelik eleştirileri ve geliştirme önerilerini kavramsal olarak haritalamayı amaçlamaktadır.

Veri güvenliği sorunları, dijital değerlendirme araçlarına yönelik en sık dile getirilen eleştiriler arasındadır. Büyük miktarda öğrenci sınav performans verisi bulut tabanlı sistemlerde depolanmakta, ancak bu verilerin yetkisiz erişime, siber saldırılara veya kötüye kullanıma karşı yeterince korunup korunmadığı genellikle belirsizdir. Ayrıca, bazı uygulamalardaki veri izleme süreçlerinin kullanıcı farkındalığı olmadan yürütülmesi ciddi etik soruları gündeme getirmektedir (Williamson ve Piattoeva, 2022). Bu durum, öğrenci gizliliğine yönelik tehditler ve bilgi güvenliğindeki zafiyetler yaratarak özellikle K-12 düzeyindeki öğrenciler için risk oluşturmaktadır.

Eleştirilerin ikinci önemli odağı kültürel uyumsuzluktur. Birçok dijital değerlendirme aracı, Batı merkezli pedagojik yaklaşımlarla tasarlanmış ve küresel ölçekte uygulanması amaçlanmıştır. Ancak bu araçlar genellikle farklı kültürel kodları, dil yapılarını, öğrenme stillerini veya inanç temelli öğrenme uygulamalarını yeterince hesaba katmamaktadır (Deardorff, 2011). Bu durum, özellikle çok kültürlü sınıflarda veya farklı ülkelerdeki uygulamalarda eşitsizlik sorunlarına yol açmaktadır. Kültürel önyargı içeren sorular, yerel bağlama uygun olmayan içerik veya standart sınav formatları, öğrencilerin performanslarını doğru bir şekilde yansıtmalarını engelleyebilir.

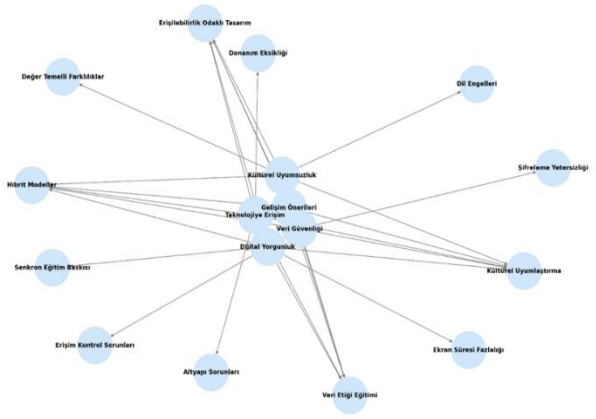
Üçüncüsü, teknolojiye erişimdeki eşitsizlik, dijital değerlendirme sistemlerinin uygulanabilirliğini ve kapsayıcılığını doğrudan etkilemektedir. Özellikle düşük gelirli bölgelerde yaşayan öğrenciler için yeterli ekipman (bilgisayar, tablet, yüksek hızlı internet) veya çevrimiçi sınavlar için uygun sessiz bir ortama erişim eksikliği, dijital değerlendirme uygulamalarında önemli eşitsizlikler yaratmaktadır (OECD, 2020). Bu durum, değerlendirme sonuçlarının yalnızca bireyler arasında değil, aynı zamanda okullar, bölgeler ve ülkeler arasında da karşılaştırılabilirliğini engellemektedir.

Ekranların önünde sürekli geçirilen zaman, ardışık eş zamanlı değerlendirme oturumları ve sürekli çevrimiçi gözetim altında olma hissi, öğrencilerde bilişsel tükenmişliğe, düşük motivasyona ve sınav kaygısına yol açabilir (Bennett vd., 2020). Ayrıca, dijital ortamda yapılan değerlendirmeler genellikle zaman baskısı ve dikkat dağıtıcı unsurlar gibi dezavantajlarla birlikte gelir ve öğrencilerin gerçek akademik performanslarını göstermelerini zorlaştırır.

Literatürde bu sorunları ele almak için çeşitli çözümler geliştirilmiştir. Veri güvenliği için en önemli öneriler arasında, eğitim kurumlarının ISO 27001 gibi bilgi güvenliği standartlarına dayalı şifreleme ve erişim kontrol sistemleri uygulaması yer almaktadır. Ayrıca öğrencilere veri etiği konusunda farkındalık eğitimi verilmesi, kullanıcı onayının alınması ve şeffaf veri politikalarının geliştirilmesi güven sorunlarını azaltma potansiyeline sahiptir.

Kültürel adaptasyon için, dijital değerlendirme sistemlerinin yerelleştirilmiş içerikle zenginleştirilmesi, çok dilli destek sunulması ve kültürel açıdan tarafsız soru yapıları içermesi önerilmektedir. Özellikle farklı bölgelerde uygulanan dijital sınavlar için, içeriğin bölgesel uzmanlar tarafından yeniden değerlendirilmesi gerektiği vurgulanmaktadır.

Erişilebilirlik sorunlarına bir çözüm olarak düşük bant genişliğine sahip sınav yazılımlarının, çevrimdışı sınav modüllerinin ve mobil uyumlu değerlendirme araçlarının geliştirilmesi kritik öneme sahiptir. Ayrıca, kamu politikası düzeyinde, dijital uçurumu kapatmak için evde internet desteği, okul tabanlı dijital merkezlerin açılması ve donanım hibeleri gibi kapsayıcı uygulamalar da tartışılmaktadır. Diğer önemli stratejiler arasında video yerine sesli soru formatları, sınavlar arasında mikro molalar ve duyusal farkındalık egzersizlerinin entegrasyonu yer almaktadır. Tüm bu kavramsal unsurların içerik analizi bağlamında nasıl bir araya geldiğini göstermek için aşağıda kavramsal bir harita sunulmuştur.



Şekil 7. Dijital ölçme uygulamalarına yönelik eleştiriler ve gelişim önerilerinin kavramsal haritası

Bu kavramsal harita, eleştiriler ve önerilen çözümler arasındaki karşılıklı ilişkileri görselleştirerek dijital değerlendirme sistemlerine ilişkin gelişen bakış açılarını ortaya koymaktadır. Haritadaki her eleştiri, farklı bağlamlardan türetilen önerilerle eşleştirilmektedir. Örneğin, "veri güvenliği" kavramı, "şifreleme yetersizliği" ve "erişim kontrol sorunları" alt başlıklarıyla daha da ayrıntılı olarak ele alınmış ve bu sorunları ele almak için

"veri etiği eğitimi" ve "hibrit modeller" gibi stratejiler önerilmiştir.

Dijital değerlendirme uygulamalarının hızla yaygınlaşması eleştirel refleksleri zorunlu kılmıştır. Eğitimde, teknolojinin adaleti, güvenliği, erişilebilirliği ve kültürel duyarlılığı, işlevselliği kadar hayati önem taşımaktadır. Bu bağlamda, içerik analizleri, dijital değerlendirme sistemlerinin daha kapsayıcı, etik temelli ve öğrenme odaklı hale gelmesi için rehberlik sağlamaktadır. Özellikle Müslüman toplumların eğitim sistemlerinde, dijital değerlendirmenin insani değerlerle ve pedagojik hedeflerle uyumlu olması gerektiği unutulmamalıdır. Çünkü değerlendirme yalnızca başarıya değil, aynı zamanda insanın anlam arayışına da rehberlik eder.

4. Tartışma

Dijital değerlendirme sistemlerinin eğitim ortamlarında giderek yaygınlaşması, ölçme ve değerlendirme süreçlerini yalnızca teknik bir dönüşüme uğratmakla kalmamış; aynı zamanda pedagojik yaklaşımlar, değerlendirme felsefesi ve öğrenme süreçlerinin doğası üzerinde de köklü değişimler yaratmıştır. Bu çalışma kapsamında elde edilen bulgular, dijital değerlendirme literatüründe dile getirilen eleştirilerin ve gelişim önerilerinin belirli uygulama yönelimleri etrafında kümelendiğini ortaya koymaktadır. Bu durum, dijital değerlendirme süreçlerinin rastlantısal ya da dağınık uygulamalar bütünü olmaktan ziyade, belirli kavramsal ve işlevsel yönelimler doğrultusunda yapılandığını göstermektedir.

Araştırmanın en dikkat çekici bulgularından biri, veri güvenliğinin literatürde en baskın sorun alanı olarak öne çıkmasıdır. Bu bulgu, öğrenme analitiği ve dijital değerlendirme alanında yapılan önceki çalışmalarla tutarlılık göstermektedir. Özellikle veri odaklı uygulama yönelimleri kapsamında değerlendirilen bu durum, dijital değerlendirme sistemlerinin yalnızca öğrenme çıktıları üretmekle kalmayıp aynı zamanda büyük ölçekli veri üretim ve işleme süreçlerini de içermesinden kaynaklanmaktadır. Ferguson (2019) ve Prinsloo ve Slade (2017) tarafından vurgulandığı üzere, öğrenci verilerinin etik kullanımı, şeffaflık ve veri yönetimi politikaları dijital değerlendirme sistemlerinin güvenilirliğinde belirleyici bir rol oynamaktadır. Bu çalışmada elde edilen bulgular da bu yaklaşımları desteklemekte ve veri güvenliği konusunun dijital değerlendirme sistemlerinin sürdürülebilirliği açısından kritik bir uygulama yönelimi olarak ele alınması gerektiğini ortaya koymaktadır.

Çalışmada öne çıkan bir diğer önemli bulgu, dijital değerlendirme araçlarının kültürel bağlamla ilişkili sınırlılıklarıdır. Literatürde sıklıkla dile getirildiği gibi, dijital değerlendirme araçlarının büyük ölçüde Batı merkezli pedagojik yaklaşımlar doğrultusunda geliştirilmiş olması, farklı kültürel bağlamlarda uygulanabilirlik sorunlarını beraberinde getirmektedir. JISC (2020) tarafından da vurgulandığı üzere, değerlendirme araçlarının kültürel olarak uyarlanmaması durumunda ölçme sonuçlarının geçerliliği zayıflayabilmektedir. Bu çalışmanın bulguları, kültürel duyarlılık temelli uygulama yönelimlerinin dijital değerlendirme süreçlerinde daha fazla önem kazanması gerektiğini göstermektedir.

Erişim eşitsizlikleri bağlamında elde edilen sonuçlar da literatürdeki dijital uçurum tartışmalarıyla paralellik göstermektedir. Luckin vd. (2016), dijital teknolojilere erişimdeki farklılıkların öğrenme ve değerlendirme süreçlerinde eşitsizlik yarattığını belirtmektedir. Bu çalışmada da benzer biçimde, dijital değerlendirme uygulamalarının farklı sosyoekonomik koşullara sahip öğrenciler için eşit fırsatlar sunmadığı görülmektedir. Bu durum, eşitlik ve erişilebilirlik temelli uygulama yönelimlerinin dijital değerlendirme sistemlerinin tasarımında merkezi bir konuma yerleştirilmesi gerektiğini ortaya koymaktadır.

Bununla birlikte, öğrencilerin dijital refahı ve öğrenme deneyimi üzerindeki etkiler de önemli bir tartışma alanı olarak öne çıkmaktadır. Uzun süreli ekran kullanımı, yoğun çevrimiçi değerlendirme süreçleri ve sınırlı geri bildirim mekanizmaları öğrencilerin bilişsel yükünü artırabilmekte ve motivasyonlarını olumsuz etkileyebilmektedir. Bearman vd. (2020) tarafından da ifade edildiği gibi, dijital değerlendirme süreçlerinin öğrenme psikolojisiyle uyumlu biçimde tasarlanması gerekmektedir. Bu çalışma, refah ve öğrenme deneyimi odaklı uygulama yönelimlerinin dijital değerlendirme sistemlerinde daha sistematik biçimde ele alınması gerektiğini ortaya koymaktadır.

Geri bildirim mekanizmalarına ilişkin bulgular da dikkat çekicidir. Literatürde dijital değerlendirme sistemlerinin çoğunlukla otomatik puanlama süreçlerine odaklandığı, ancak öğrenmeyi destekleyen geri bildirim mekanizmalarının sınırlı kaldığı ifade edilmektedir. Nicol (2021), etkili geri bildirimin öğrenme sürecinin ayrılmaz bir parçası olduğunu vurgulamaktadır. Bu bağlamda, geri bildirim temelli uygulama yönelimlerinin güçlendirilmesi, dijital değerlendirme sistemlerinin öğrenme odaklı hale getirilmesi açısından kritik bir gereklilik olarak ortaya çıkmaktadır.

Genel olarak değerlendirildiğinde, bu çalışmanın bulguları dijital değerlendirme sistemlerinin çok boyutlu bir yapı içerisinde ele alınması gerektiğini açık biçimde göstermektedir. Pedagojik, etik, kültürel ve teknolojik boyutların birbirinden bağımsız değil, aksine birbirini tamamlayan unsurlar olarak değerlendirilmesi gerekmektedir. Bu bağlamda, uygulama yönelimleri kavramı, dijital değerlendirme literatüründe dağınık halde bulunan yaklaşımların sistematik biçimde analiz edilmesine olanak tanıyan önemli bir kavramsal çerçeve sunmaktadır. Dijital değerlendirme sistemlerinin yalnızca teknolojik yenilikler olarak değil, öğrenme süreçlerini şekillendiren çok boyutlu yapılar olarak ele alınması gerekmektedir. Bu çalışma, dijital ölçme ve değerlendirme uygulamalarının belirli uygulama yönelimleri çerçevesinde yeniden yorumlanmasının, alandaki teorik ve uygulamalı çalışmalara önemli katkılar sağlayacağını ortaya koymaktadır.

5. Sonuç ve Öneriler

Bu çalışma, dijital ölçme ve değerlendirme uygulamalarına ilişkin akademik literatürde dile getirilen eleştirileri ve gelişim önerilerini sistematik biçimde analiz ederek bu alanın belirli uygulama yönelimleri etrafında yapılandığını ortaya koymuştur. Elde edilen bulgular, dijital değerlendirme sistemlerinin yalnızca teknolojik araçlar olarak ele alınmasının yetersiz olduğunu; bu sistemlerin pedagojik, etik, kültürel ve psikometrik boyutların etkileşimi içinde değerlendirilmesi gerektiğini açık biçimde göstermektedir.

Araştırma sonuçları, dijital değerlendirme literatüründe öne çıkan sorun alanlarının başında veri güvenliği, erişim eşitsizliği, kültürel uyumsuzluk, sınırlı geri bildirim mekanizmaları ve öğrenci refahı ile ilişkili risklerin geldiğini göstermektedir. Özellikle veri güvenliği ve algoritmik şeffaflık, dijital değerlendirme sistemlerinin güvenilirliğini doğrudan etkileyen temel bileşenler olarak öne çıkmaktadır. Bunun yanı sıra, dijital değerlendirme uygulamalarının büyük ölçüde düşük ve orta düzey bilişsel becerilere odaklanması, bu sistemlerin üst düzey düşünme becerilerini ölçme kapasitesinin sınırlı kaldığını göstermektedir.

Bu çalışma kapsamında geliştirilen kavramsal çerçeve, dijital değerlendirme uygulamalarını veri odaklı, kültürel duyarlılık temelli, eşitlik ve erişilebilirlik temelli, refah odaklı ve geri bildirim temelli uygulama yönelimleri altında sınıflandırarak literatürde dağınık halde bulunan yaklaşımları bütüncül bir yapıda yeniden organize etmektedir. Bu yönüyle çalışma, dijital değerlendirme alanına yalnızca betimleyici değil, aynı zamanda açıklayıcı ve yapılandırıcı bir katkı sunmaktadır.

Elde edilen bulgular ayrıca, dijital değerlendirme sistemlerinin tasarımında pedagojik hedeflerin çoğu zaman teknolojik olanakların gerisinde kaldığını göstermektedir. Bu durum, teknoloji odaklı değil öğrenme odaklı değerlendirme yaklaşımlarının geliştirilmesi gerektiğini ortaya koymaktadır. Özellikle yapay zekâ destekli değerlendirme sistemlerinin yaygınlaşmasıyla birlikte, bu sistemlerin etik, pedagojik ve psikometrik geçerliliklerinin daha sistematik biçimde ele alınması zorunlu hale gelmektedir. Çalışma, dijital ölçme ve değerlendirme uygulamalarının sürdürülebilir, kapsayıcı ve pedagojik açıdan anlamlı hale getirilebilmesi için çok boyutlu ve yönelim temelli bir yaklaşımın benimsenmesi gerektiğini ortaya koymaktadır. Bu bağlamda önerilen uygulama yönelimleri çerçevesi, hem araştırmacılar hem de uygulayıcılar için analitik bir değerlendirme aracı sunmakta; dijital değerlendirme sistemlerinin tasarımı, uygulanması ve değerlendirilmesi süreçlerinde kullanılabilecek kavramsal bir rehber işlevi görmektedir.

Bu çalışma kapsamında elde edilen bulgular doğrultusunda, dijital ölçme ve değerlendirme uygulamalarının daha etkili, kapsayıcı ve sürdürülebilir hale getirilebilmesi için çok boyutlu iyileştirme alanlarının dikkate alınması gerekmektedir. Öncelikle, dijital değerlendirme sistemlerinde veri güvenliği ve etik standartların güçlendirilmesi büyük önem taşımaktadır. Öğrenci verilerinin korunmasına yönelik şeffaf veri politikalarının geliştirilmesi, kullanıcı onayına dayalı veri işleme süreçlerinin yaygınlaştırılması ve kurum düzeyinde veri etiği standartlarının oluşturulması bu alandaki temel gereksinimler arasında yer almaktadır.

Bununla birlikte, dijital değerlendirme araçlarının farklı kültürel bağlamlara uyum sağlayabilecek şekilde yeniden tasarlanması gerekmektedir. Özellikle çok kültürlü eğitim ortamlarında kullanılan standart değerlendirme araçlarının yerelleştirilmiş içeriklerle desteklenmesi, dilsel ve kültürel uyarlama çalışmalarının artırılması ve bağlama duyarlı ölçme araçlarının geliştirilmesi, değerlendirme süreçlerinin geçerliliğini ve güvenilirliğini artıracaktır.

Dijital teknolojilere erişimdeki eşitsizliklerin azaltılması da önemli bir diğer öncelik alanıdır. Bu bağlamda, düşük bant genişliğine uygun yazılımların geliştirilmesi, çevrimdışı değerlendirme seçeneklerinin sunulması ve mobil uyumlu ölçme araçlarının yaygınlaştırılması gibi uygulamalar, dijital değerlendirme süreçlerinin daha kapsayıcı hale gelmesine katkı sağlayacaktır. Aynı zamanda kamu politikaları düzeyinde yapılacak altyapı yatırımları ve dijital eşitliği destekleyen stratejiler, bu sürecin sürdürülebilirliğini güçlendirecektir.

Öğrenci refahını merkeze alan değerlendirme yaklaşımlarının geliştirilmesi de dijital ölçme süreçlerinin kalitesini artıracaktır. Uzun süreli ekran kullanımının ve yoğun çevrimiçi değerlendirme uygulamalarının öğrenciler üzerindeki bilişsel ve duygusal etkileri dikkate alınmalı; esnek zamanlı sınav uygulamaları, mikro değerlendirme etkinlikleri ve öğrenme sürecini destekleyen alternatif değerlendirme yöntemleri yaygınlaştırılmalıdır.

Ayrıca, dijital değerlendirme sistemlerinde geri bildirim mekanizmalarının niteliğinin artırılması gerekmektedir. Otomatik puanlama odaklı sistemlerin ötesine geçilerek öğrencilerin öğrenme süreçlerini destekleyen kişiselleştirilmiş ve anlamlı geri bildirim sistemlerinin geliştirilmesi, öğrenmenin kalıcılığı ve öğrencilerin öz düzenleme becerileri açısından kritik bir rol oynayacaktır. Dijital ölçme ve değerlendirme uygulamalarının pedagojik temellerinin güçlendirilmesi ve öğretmenlerin bu süreçteki rolünün yeniden tanımlanması gerekmektedir. Öğretmenlere yönelik dijital değerlendirme okuryazarlığı eğitimlerinin artırılması, uygulama yönelimleri çerçevesinde farklı değerlendirme stratejilerinin tanıtılması ve bu süreçlerin mesleki gelişim programlarıyla desteklenmesi, dijital değerlendirme sistemlerinin daha etkin kullanılmasına katkı sağlayacaktır. Ayrıca, gelecekte yapılacak araştırmalarda uygulama yönelimleri temelinde karşılaştırmalı, deneysel ve uzunlamasına çalışmaların gerçekleştirilmesi, alandaki kuramsal ve uygulamalı bilgi birikimini daha da zenginleştirecektir.

Ethics Committee Approval: I declare that this study did not contain any elements requiring ethical permission in any part of the data collection and writing process, and therefore did not require ethical committee approval.
Informed Consent: I declare that informed consent was not required at any stage of the data collection and writing process for this study.
Peer-review: Externally peer-reviewed.
Author Contributions: The author carried out all aspects of the study, including conceptualization, design, data collection, analysis, and manuscript preparation.
Conflict of Interest: The authors have no conflicts of interest to declare.
Financial Disclosure: The authors declared that this study has received no financial support.
Use of Artificial Intelligence: During the literature review, artificial intelligence tools (Consensus: AI for Research and NotebookLM) were used, particularly for searching foreign sources. Paid/pro versions of the AI tools were used. The author is fully responsible for the content of the article.

References

Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education, 32*(2), 159–181. https://doi.org/10.1080/02602930600801928

Barrett, H. C. (2007). Researching Electronic Portfolios and Learner Engagement: The REFLECT Initiative. *Journal of Adolescent & Adult Literacy, 50*(6), 436–449. https://doi.org/10.1598/JAAL.50.6.2

Bearman, M., Dawson, P., Ajjawi, R., Tai, J., & Boud, D. (2020). Re-imagining university assessment in a digital world. *Assessment & Evaluation in Higher Education, 45*(5), 785–798.

Bennett, R. (2015). The changing nature of educational assessment. Review of Research in Education.

Bennett, R., Dawson, P., Bearman, M., Molloy, E., & Boud, D. (2017). How technology shapes assessment design.

Biggs, J. B., & Tang, C. S. (with Society for Research into Higher Education). (2011). Teaching for quality learning at university: What the student does (4th edition). McGraw-Hill/Society for Research into Higher Education/Open University Press.

Biggs, J. B., Biggs, J. B., & Collis, K. F. (1982). Evaluating the quality of learning: The SOLO taxonomy (structure of the observed learning outcome). Academic Press.

Bozkurt, A., & Sharma, R. C. (2020). Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. https://doi.org/10.5281/ZENODO.3778083

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (Third edition). Sage.

Darling-Hammond, L., Zielezinski, M. B., & Goldman, S. (2020). Using technology to support at-risk students' learning. Alliance for Excellent Education.

De Ayala, R. J. (2022). The theory and practice of item response theory (Second edition). The Guilford Press.

Deardorff, D. K. (2011). Assessing intercultural competence. *New Directions for Institutional Research, 2011*(149), 65–79. https://doi.org/10.1002/ir.381

Ferguson, R. (2019). Ethical challenges for learning analytics. Journal of Learning Analytics, 6(3), 25–30.

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education, 57*(4), 2333–2351. https://doi.org/10.1016/j.compedu.2011.06.004

Haladyna, T. M., & Rodriguez, M. C. (2013). Developing and Validating Test Items (0 edn). Routledge. https://doi.org/10.4324/9780203850381

Henderson, M., Selwyn, N., & Aston, R. (2017). What works and why? Student perceptions of 'useful' digital technology in university teaching and learning. *Studies in Higher Education, 42*(8), 1567–1579. https://doi.org/10.1080/03075079.2015.1007946

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial Intelligence in Education.

Ifenthaler, D., Mah, D.-K., & Yau, J. Y.-K. (2019). Utilising Learning Analytics for Study Success: Reflections on Current Empirical Findings. In D. Ifenthaler, D.-K. Mah, & J. Y.-K. Yau (Eds), Utilizing Learning Analytics to Support Study Success (pp. 27–36). Springer International Publishing. https://doi.org/10.1007/978-3-319-64792-0_2

JISC. (2020). Learning analytics code of practice. JISC.

Kizilcec, R. F., & Lee, H. (2022). Algorithmic fairness in education. In W. Holmes & K. Porayska-Pomsta, The Ethics of Artificial Intelligence in Education (1st edn, pp. 174–202). Routledge. https://doi.org/10.4324/9780429329067-10

Laurillard, D. (2012). Teaching as a design science: Building pedagogical patterns for learning and technology. Routledge.

Looney, J. (2011). Integrating Formative and Summative Assessment: Progress Toward a Seamless System? (OECD Education Working Papers No. 58; OECD Education Working Papers, Vol. 58). https://doi.org/10.1787/5kghx3kbl734-en

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. (2016). Intelligence Unleashed: An Argument for AI in Education. Pearson.

McKenney, S., & Reeves, T. C. (2019). Conducting educational design research (Second edition). Routledge, Taylor & Francis Group. https://doi.org/10.4324/9781315105642

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). Qualitative data analysis: A methods sourcebook (Fourth edition). SAGE.

Nicol, D. (2021). The power of internal feedback: Exploiting natural comparison processes. *Assessment & Evaluation in Higher Education, 46*(5), 756–778.

Penetrating the Fog: Analytics in Learning and Education. (n.d.). EDUCAUSE Review. Retrieved 4 August 2025, from https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education

Prinsloo, P., & Slade, S. (2017). An elephant in the learning analytics room: Ethics and student privacy. *Journal of Learning Analytics, 4*(2), 24–29.

Redecker, C. (2017). European framework for digitally competent educators.

Redecker, C., & Johannessen, Ø. (2013). Changing assessment – towards a new assessment paradigm using ICT. European Journal of Education.

Selwyn, N. (2019). Should Robots Replace Teachers? AI and the Future of Education.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. International Journal of Medical Education, 2, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tekin, H. (2019). *Eğitimde ölçme ve değerlendirme* (10. baskı). Yargı Yayınevi.

Tondeur, J., Scherer, R., Baran, E., Siddiq, F., Valtonen, T., & Sointu, E. (2019). Teacher educators as gatekeepers:

Preparing the next generation of teachers for technology integration in education. *British Journal of Educational Technology, 50*(3), 1189–1209. https://doi.org/10.1111/bjet.12748

University of Southampton, UK, Alruwais, N., Wills, G., & Wald, M. (2018). Advantages and Challenges of Using e-Assessment. *International Journal of Information and Education Technology, 8*(1), 34–37. https://doi.org/10.18178/ijiet.2018.8.1.1008

Weiss, D. J., & Kingsbury, G. G. (1984). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement, 21*(4), 361–375. https://doi.org/10.1111/j.1745-3984.1984.tb01040.x

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 3–14. https://doi.org/10.1016/j.stueduc.2011.03.001

Wiliam, D. (2018). Embedded formative assessment (Second edition). Solution Tree Press.

Williamson, B. (2017). Big Data in Education.

Williamson, B., & Piattoeva, N. (2019). Objectivity as standardization in data-scientific education policy, technology and governance. *Learning, Media and Technology, 44*(1), 64–76. https://doi.org/10.1080/17439884.2018.1556215