Copyright © 2026 The Author(s)
Didaktika: Jurnal Kependidikan Vol. 15 No. 2 Mei 2026

Computerized Adaptive Testing (CAT) as an Instrument for Learning Assessment in Elementary Schools: A Systematic Review of Implementation Challenges and Solutions

Siti Zainab[1] , Shaleh[2]
[1,2]Universitas Islam Negeri Sunan Kalijaga Yogyakarta, Indonesia

[1] 24204081003@student.uin-suka.ac.id
DOI: https://doi.org/10.58230/27454312.4026

Abstract

Technology-based assessments are increasingly relevant in the digital era, but the implementation of Computerized Adaptive Testing (CAT) at the elementary school level is still lacking systematic studies, especially in Indonesia. This study aims to describe the development of CAT implementation, identify key challenges, and synthesize effective solutions based on national and international literature. The method used is systematic literature review (SLR) referring to the PRISMA 2020 protocol through four stages: identification, screening, eligibility, and inclusion, with searches on the Google Scholar, ERIC, Scopus, and SINTA databases for the 2021–2026 range. Of the 309 identified literature, 15 articles met the inclusion criteria and were analyzed using thematic analysis and narrative synthesis. The results show a significant increase in publication from one article (2021) to five articles (2025). Implementation challenges include three dimensions: technical in the form of network infrastructure and hardware limitations; pedagogical in the form of low digital literacy of students and teachers due to the elimination of ICT subjects; and institutions in the form of a lack of supporting systemic policies. Evidence-based solutions include infrastructure strengthening, intensive computer training, development of CAT systems based on Item Response Theory (IRT), and adaptive platform integration. It is concluded that resistance to CAT is more extrinsic-infrastructural than psychological. Controlled experimental research and longitudinal studies in the Indonesian context are highly recommended to strengthen the ecological validity of these findings.
Keywords: Digital Assessment, Computerized Adaptive Testing, Item Response Theory, Elementary School, Implementation Challenges

Introduction

The transformation of educational evaluation practices within the digital era has created substantial opportunities for developing more precise, adaptive, and efficient psychometric instruments. Among the most groundbreaking advancements in this domain is Computerized Adaptive Testing (CAT), a technology-driven assessment framework that autonomously calibrates item difficulty in correspondence with each individual examinee's demonstrated competency level. Unlike conventional static testing instruments that present a uniform set of questions to all respondents regardless of their ability, CAT employs algorithmic mechanisms grounded in Item Response Theory (IRT) to engage in real-time, dynamic item selection, thereby ensuring that every participant encounters tasks aligned with their actual proficiency. Research has demonstrated that a CAT system utilizing an IRT approach can effectively assess student learning outcomes with high levels of practicality and measurement precision, particularly when the system terminates once a satisfactory ability estimate is achieved (Huda et al., 2024). Furthermore, variable-length CAT configurations have been shown to reduce overall test length

by approximately 30% while sustaining high measurement accuracy, with adaptive testing exhibiting lower estimation errors and greater reliability compared to conventional fixed-length evaluations (Ogunjimi et al., 2021). This capacity to minimize the quantity of administered items without compromising psychometric quality substantially mitigates cognitive fatigue among examinees, thereby yielding more valid and dependable assessment data.

A number of empirical findings strengthen the urgency of developing CAT at the basic education level. Research conducted by (Li et al., 2023) It shows that computer-based adaptive tests are able to estimate students' abilities more accurately than teachers, even in the population of students with special needs in inclusive schools. In the context of primary school, the balance between test duration and measurement accuracy is a crucial factor, given the limitations of early childhood students' concentration that cannot be ignored. Meanwhile, (Mauritius, 2022) in their study of 28,485 students in grades two through six elementary schools in China proved that the system *Cognitive Diagnostics* CAT (CD-CAT) significantly improves the accuracy of reading ability classification without adding to the excessive test load. From the Indonesian context, through a qualitative phenomenological approach, various real obstacles in the implementation of CAT are identified, including the limitations of educators' knowledge and skills, technical constraints of infrastructure, and the incompatibility of the CAT system with the essay-based question format that is still predominantly used (Image, 2025).

Although the potential of CAT as a learning assessment instrument has gained considerable recognition, there is a gap (*Gap*) striking research in the available literature. First, the majority of studies on CAT are still centered on secondary and higher education, and are carried out in developed countries with well-established technological infrastructure. Studies that specifically raise the context of primary schools, especially in developing countries such as Indonesia, are still very minimal (Akhtar et al., 2023). In fact, learners at the elementary school level have unique cognitive development characteristics, being at a concrete operational stage according to Piaget's theory, so adaptive assessment approaches designed for adults cannot simply be applied without fundamental adjustments. Second, most of the *systematic review* that has been published in the CAT domain only touches on the technical and psychometric dimensions, without integrating the pedagogical, institutional, and policy dimensions that actually determine the success of implementation in the field (Ebenbeck & Gebhardt, 2022). Third, research on the development of CAT instruments designed specifically for the context of subjects in primary schools, taking into account the suitability of local cultural and curricular contexts, is hardly found in reputable national or international literature (Ebenbeck et al., 2024).

The novelty of this research comes from three aspects that complement each other. Thethematically, this study exclusively focuses on the context of elementary schools as educational units with different characteristics and challenges from other levels. Methodologically, the research uses a *systematic review* approach that integrates findings from the national and international literature in a balanced manner. In terms of contribution, this study not only documents previous findings, but also produces a synthesis map that can be used as the basis for the formulation of contextual and evidence-based implementation solutions. Thus, this research is expected to be able to fill the knowledge gap that has not been thoroughly answered in the technology-based assessment literature at the basic education level.

This study addresses three principal objectives. First, it seeks to construct a comprehensive description of the developmental trajectory of *Computerized Adaptive Testing* (CAT) as an instrument for learning assessment in elementary schools, synthesized from a broad corpus of national and international scholarly literature. Second, the study endeavors to identify and examine the predominant challenges encountered during CAT implementation at the primary education level, encompassing *technical* and infrastructural constraints, *pedagogical*

Copyright © 2026 The Author(s)

dimensions and educator competency gaps, as well as institutional and regulatory frameworks. Third, this investigation aims to evaluate and document empirically validated solutions that have demonstrated measurable effectiveness in addressing the aforementioned implementation barriers within primary school contexts, as substantiated by findings drawn from preceding empirical research.

This research is expected to make a real contribution to two main dimensions. Theoretically, this study enriches the treasures of psychometric science and educational technology by providing a comprehensive synthesis of knowledge about CAT in the context of primary schools, while opening up space for more in-depth and contextual advanced empirical research. Practically, the findings of this study can be used by teachers and school principals as an initial guide in considering the adoption of the CAT system in their respective educational units. For software and educational technology developers, the results of this study can be a reference in designing a more user-friendly system for elementary school-age students. Meanwhile, for policy makers, this research can be a scientific basis in drafting regulations and programs to strengthen assessment technology infrastructure that is equitable, inclusive, and equitable throughout Indonesia.

Method

This research uses a design *Systematic Literature Review* (SLR), which is a research approach that systematically identifies, selects, evaluates, and synthesizes findings from various studies relevant to the topic being studied. This approach was chosen because it is able to produce a comprehensive, structured, and replicable synthesis of evidence by other researchers. The entire reporting process of this systematic review refers to the guidance *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) 2020, as developed by (Page et al., 2021), which consists of 27 reporting items covering four main stages, namely *Identification*, *Screening*, *Eligibility*, and *Inclusion*. It should also be emphasized that in the context of education in Indonesia, *Computerized Adaptive Testing* (CAT) is interpreted and implemented institutionally through the *Computer-Based National Assessment* (ANBK), which is a national evaluation program organized by the Ministry of Education, Culture, Research, and Technology to measure the quality of education units at the primary and secondary levels through three main instruments, namely *Minimum Competency Assessment* (AKM), *Character Survey*, and *Learning Environment Survey* (Hastuti et al., 2025). Thus, the study of CAT in this study also inherently includes an examination of the practices and challenges of ANBK at the elementary school level as a concrete form of the implementation of a computer-based adaptive assessment system in Indonesia (Hastuti et al., 2025).

The subjects in this study are not individuals or respondents, but scientific articles as units of analysis. The research population includes all national and international articles that cover the topic *Computerized Adaptive Testing* (CAT) and ANBK in the context of learning assessments, especially at the basic education level, which are issued in the period of 2021 to 2026. The search process is carried out systematically through several reputable databases, including Google Scholar, ERIC (*Education Resources Information Center*), Scopus, and SINTA (*Science and Technology Index*) for national sources. The search uses a combination of keywords in English, including: *"Computerized Adaptive Testing"*, *"CAT assessment"*, *"Adaptive Testing Elementary School"*, *"CAT implementation challenges"*, *"Item Response Theory Primary Education"*, and *"Adaptive Assessment Technology"*. The keywords in Indonesian used include: "computerized adaptive test", "elementary school CAT", "Computer-Based National Assessment", "elementary school ANBK", "implementation of basic education CAT", and

"learning ANBK challenges". The keywords are combined using the *Boolean* AND and OR to expand and focus search coverage (Marya Ulfha 1, 2025).

Data collection was carried out through two stages of selection. The first stage is screening based on titles and abstracts, while the second stage is *full-text review* based on predetermined inclusion and exclusion criteria. Inclusion criteria include: articles published between 2021–2026, in Indonesian or English, in the form of empirical research articles or reviews that discuss CAT and ANBK in the context of basic education, and available in full-text versions. Articles are excluded if they discuss CAT outside of the educational context, do not include analyzable research results, or are duplicates of studies that have been identified. In order to minimize selection bias, the screening process was carried out independently by two researchers and the differences in assessment were resolved through consensus discussion, as recommended in the PRISMA 2020 protocol.

The data analysis technique applied is thematic analysis (*Thematic Analysis*) and narrative synthesis (*Narrative Synthesis*). Thematic analysis was used to identify the main patterns, categories, and themes that emerged from all included articles, especially related to the challenges and solutions of CAT and ANBK implementation in elementary schools. The data extraction process is carried out using a form that contains information about the author's name, year of publication, research country, study design, objectives, main findings, as well as reported challenges and solutions. Each article was assessed for quality using an assessment instrument tailored to the diversity of study designs involved. (Promise, 2025) affirms that a synthesis approach that combines thematic analysis with bias risk assessment consistently results in more valid and reliable conclusions in systematic reviews in the field of education. The findings of all studies are then synthesized narratively to produce a comprehensive picture of the developments, challenges, and solutions of CAT and ANBK implementation in primary schools, as well as applied in a systematic study that uses a similar approach in mapping CAT research trends globally.

Results

Results of PRISMA-based Literature Screening

The literature selection process in this *systematic review* research is carried out using *the PRISMA* (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) diagram as presented in the Figure above. The *identification* stage resulted in 309 article identified from various *databases*, then as many as 102 duplicate records were eliminated before the screening process began. At the *screening* stage, 207 journals were screened and 95 of them were excluded based on the criteria that had been set, leaving 112 journals that were then sought to be taken in *full-text*. However, 56 journals could not be taken, so only 56 journals were assessed for eligibility *assessment*. From this stage, a number of journals were again excluded based on specific reasons, namely *Reason 1* (n = 21) and *Reason 2* (n = 20). Finally, at the *included* stage, as many as 15 journals were declared eligible and included in a systematic review of the implementation of *Computerized Adaptive Testing* (CAT) in elementary schools.

Copyright © 2026 The Author(s)

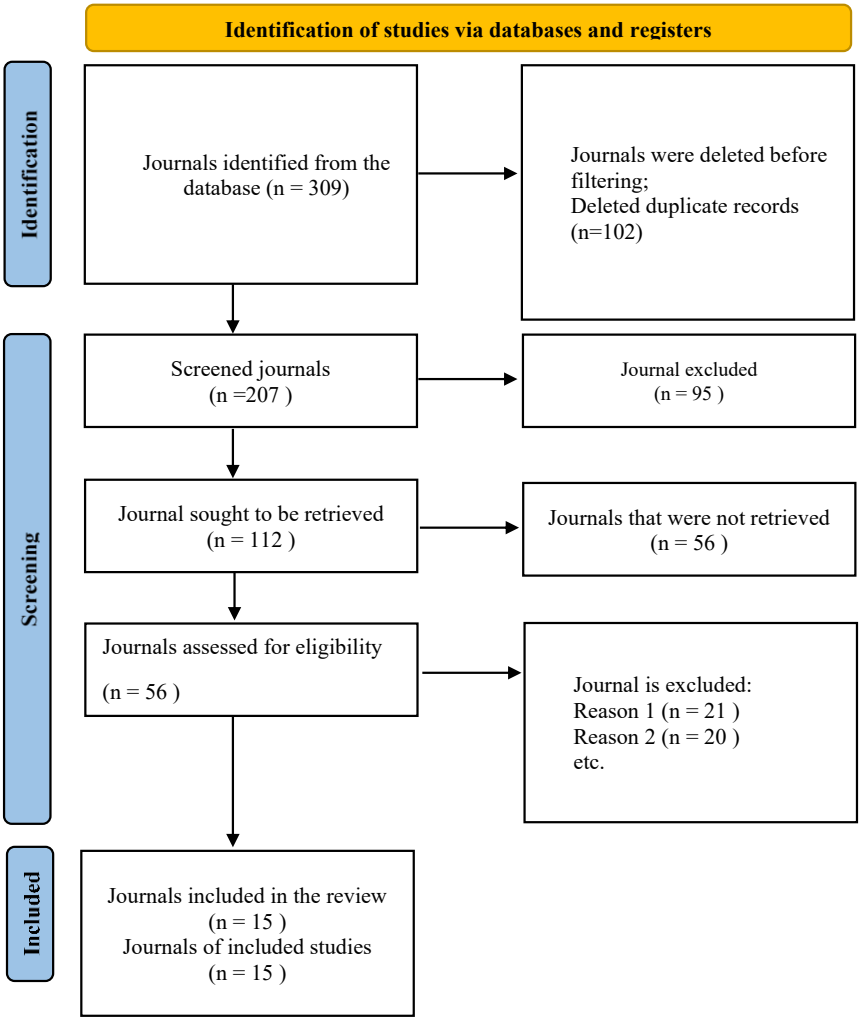


Figure 1. Flowchart PRISMA

Research Synthesis Results of 15 relevant Articles and Journals

Table 1. Synthesis of relevant research literature

No	Article Title	Author & Year	Research Focus	Computerized Adaptive Testing (CAT)	Learning Assessment in Elementary School	CAT Implementation Challenges and Solutions	Implications
1	Analysis of the Implementation of the Computer-Based National Assessment (ANBK) at Madrasah Ibtidaiyah Negeri 9 Hulu Sungai Utara	(Setiyowati, 2022)	Examine the process of implementing ANBK comprehensively starting from the stages of preparation, implementation, to post-assessment at MIN 9 HSU	It has not yet implemented CAT explicitly, but computer-based ANBK is the initial foundation towards the development of digital adaptive assessments	ANBK is carried out independently and online as a form of computer-based assessment at the basic level with structured procedures	Central server disruptions are a major obstacle; The solution is to strengthen network infrastructure and the readiness of a competent personnel team	The readiness of human resources, infrastructure, and participant coaching greatly determines the success of the implementation of computer-

							based assessments in elementary schools
2	Analysis of the Implementation of Computer-Based National Assessment (ANBK) in Elementary Schools	(Bella & Syahnaidi, 2022)	Describe the readiness and obstacles to the implementation of ANBK at SDN Puger from the preparation stage to the implementation process	Computer-based assessment is the first step in the introduction of a digital evaluation system that can be developed towards CAT	The implementation of ANBK faces real obstacles in the form of limited devices, unstable networks, and low student understanding of the material	The solutions implemented include borrowing laptops by teachers, adding wifi, and organizing additional lessons based on AKM questions	Schools need to build a comprehensive digital ecosystem so that computer-based assessments can be scaled up towards adaptive systems
3	Analysis of the Implementation of ANBK at SDN Tamansari I in Improving the Quality of Graduates	(Suryaman, 2025)	Analyzing the implementation of ANBK at SDN Tamansari I as an effort to improve the quality of elementary school graduates	ANBK which is carried out independently and online shows the potential for development towards a computer-based adaptive assessment system	Assessment is carried out through three structured stages, namely pre, implementation, and post-assessment in the elementary school environment	Central server disruptions are a recurring challenge; The solutions include improving infrastructure readiness and technical training of proctors and technicians	Good standardization of ANBK procedures provides a solid foundation for the development of adaptive and data-based assessment systems in primary schools
4	Analysis of the Implementation of ANBK at Senurus State Elementary School, Central Lombok	(Ningsih et al., 2025)	Analyzing the implementation of ANBK in rural elementary schools with limited access to technology and internet networks	The use of ChromeBooks and the ExamBrowser app demonstrates the readiness of the basic infrastructure that can support future CAT implementations	The implementation of online ANBK in remote schools proves that digital assessments in elementary schools can be carried out even in the midst of limitations	Technical obstacles such as server and device unstable disruptions; Recommendations include infrastructure improvements and ongoing technical training	Schools in remote areas need infrastructure support and technical training as a prerequisite for the implementation of computer-based adaptive assessments
5	The Impact of the Elimination of ICT Subjects in K-13 in ANBK Activities	(Saidah et al., 2023)	Examining the impact of the removal of ICT subjects on the quality of student participation	The lack of computer literacy due to the absence of ICT lessons has a direct impact on students'	The elimination of ICT from the curriculum reduces students' readiness to take digital	The main challenge is the low digital competence of students and teachers; The solution is to systematically	Students' digital competencies are a fundamental prerequisite that must be met before

Copyright © 2026 The Author(s)

			and ANBK results	ability to take computer-based assessments including CAT	assessments, including ANBK at the primary education level	reintegrate digital literacy into the curriculum	the CAT system can be implemented effectively
6	Towards Digital Excellence through Microsoft Office Skills Training for Grade 5 Students as ANBK Support	(Waziana et al., 2024)	Describe basic computer skills training activities for 5th grade elementary school students in preparation for the implementation of ANBK	The introduction of basic computer skills into an important initial investment to prepare students for computer-based assessments including CAT	Computer skills training for 5th grade elementary school students aims to increase students' readiness to face digital ANBK	The limitation of computer knowledge due to the absence of ICT subjects is a major challenge; The solution is hands-on practical training in the classroom	The integration of technology training from an early stage in elementary schools is a crucial strategy to support the successful implementation of computer-based adaptive assessment
7	Optimizing Educational Assessment: The Practicality of Computer Adaptive Testing (CAT) with an Item Response Theory (IRT) Approach	(Huda et al., 2024)	Develop a web-based CAT system using the IRT model 3-PL approach to measure student learning outcomes adaptively	The CAT system developed is able to adjust the difficulty level of the questions based on the participants' responses using the IRT 3-PL algorithm automatically	Web-based CAT systems have the potential to be applied in primary schools as a more precise and efficient learning assessment instrument	The technical complexity of development is a challenge; The system was evaluated with a very high practicality score of 90 students as a research sample	IRT-based CAT has proven to be practical and effective; its application in elementary school requires simplifying the interface and calibrating age-appropriate question items
8	Computer Basics Training in Preparation for Welcoming the Implementation of ANBK for Students of SD Negeri 218 Congkoe	(Farolai, 2022)	Facilitating 5th grade elementary school students in preparing themselves to face ANBK through practical basic computer training	Mastery of basic computer operations is a mandatory prerequisite competency before students can take computer-based assessments including CAT	Computer training for elementary school students is carried out to improve students' operational skills ahead of the implementation of ANBK	The challenge is that students are not familiar with computers at all; The direct practicum method was proven to significantly improve the skills of 12 students	Intensive basic computer training in elementary school is an essential foundation for the successful implementation of CAT-based digital assessments in the future

9	Design of a Website-Based Computer Assisted Test (CAT) System for School Exams (Case Study: SDN Pengasinan IV)	(Anugrah, 2022)	Designing and implementing a website-based CAT system for the implementation of school exams at SDN Pengasinan IV using a prototype model	The website-based CAT system was designed with UML and tested directly in elementary schools, demonstrating significant benefits in the efficiency of exams and instant assessments	The implementation of CAT at SDN Pengasinan IV proves that a computer-based exam system can increase the effectiveness and efficiency of assessment in elementary schools	The main challenge is designing a user-friendly system; The solution is in the form of a gradual prototype development by involving teachers and students as users	A website-based CAT system that has been successfully implemented in one elementary school can be adopted by other schools as a model for implementing efficient digital assessments
10	ROAR-CAT: Rapid Online Assessment of Reading Ability with Computerized Adaptive Testing	(Anya et al., 2025)	Develop and validate an online CAT instrument to efficiently measure children's single word reading ability without supervision	ROAR-CAT uses an IRT-based CAT algorithm that has been proven to increase test efficiency by 40%; 75 CAT questions are equivalent to 125 random questions	Validation on 485 students in grades 1–8 showed that the 3-minute ROAR-CAT resulted in a high correlation with an individual reading assessment of 5–15 minutes	Challenges in the form of consistency of cross-group item parameters; The solution is an in-depth IRT analysis and the removal of 6 items that function differently between groups	ROAR-CAT is a model for the development of online adaptive assessments that bridges research and practice, highly relevant to be applied in primary schools
11	Computerized Adaptive Testing (CAT)-Based Interactive Test System	(Sanatang, 2021)	Developed a CAT-based interactive test system for Testing Centre institutions using the Waterfall model SDLC method	The CAT system built uses the theory of two-parameter response of logistics and is able to measure the ability of participants adaptively and efficiently	The CAT system developed has the potential to be adapted for the elementary school context as a more accurate and standardized measurement of student abilities	Challenges in building an easy-to-use interface; The results of the evaluation showed a level of ease of administration of 95% and a user-friendly aspect of 89%	Interactive CAT systems are proven to be excellent in effectiveness and ease of use; adaptation for elementary school requires adjusting the difficulty level of the question item
12	Students' Acceptance on Computer-Adaptive Testing for Achievement Assessment	(Goto et al., 2023)	Investigating the acceptance rate of elementary and junior high school students in Japan to CAT	Piloted CAT does not exacerbate students' exam anxiety; students did not show a more negative	Direct research was conducted in Japanese elementary and junior high schools, proving that	Challenges in the form of concerns about exam anxiety in students with the aim of avoiding poor performance;	CAT can be carefully introduced into the primary school assessment system taking

Copyright © 2026 The Author(s)

	in Japanese Elementary and Secondary School		as an instrument for learning achievement assessment	attitude towards the CAT than the traditional fixed-item test	CAT can be accepted as an instrument for assessing learning achievement at the basic education level	CAT has been proven not to aggravate the condition	into account students' admissions and their learning objectives
13	Digital-Based PAI Learning Evaluation Transformation: Quizizz Media Optimization as an Adaptive Measurement Tool	(Scott, 2025)	Examining the use of the Quizizz application as an adaptive measurement tool in the evaluation of PAI learning through a literature review approach	Quizizz as an adaptive evaluation platform is able to adjust the difficulty level of questions based on participants' abilities, in line with the basic principles of CAT	The Quizizz platform is proven to increase student participation, provide instant feedback, and support cognitive and affective assessments in elementary schools	The challenge is in the form of maintaining the essence of affective assessment in digital format; The solution is the optimization of Quizizz's interactive and fun adaptive features	Digitizing evaluation through adaptive platforms such as Quizizz is an effective strategy towards the implementation of CAT-based assessments in primary schools
14	Two-Tier Test Based on Everyday Life Activity Stimulus to Improve the Numeracy Ability of AKM Elementary School Students	(Hapsari, 2025)	Proving the quality of two-tier numeracy assessment instruments based on stimulus daily activities in improving the numeracy achievement of elementary school students	The quality analysis of question items using IRT through the Partial Credit Model (PCM) supports the development of bank items that can be integrated into the CAT system	The instrument was tested on 442 elementary school students in Kulon Progo Regency; Numeracy achievement increased by 16.01% after the use of the developed instrument	The challenge is in the form of comprehensive psychometric validation; The solution is the combined use of Aiken, EFA, and IRT formulas to ensure the quality of the instrument	Valid and reliable IRT-based numeracy instruments are an important foundation for the development of CAT bank items for assessment in primary schools
15	Computerized Adaptive Assessment of Understanding of Programming Concepts in Primary School Children	(Hogenboom et al., 2022)	Develop and validate the Computerized Adaptive Programming Concepts Test (CAPCT) to measure the understanding of programming concepts in elementary school children	CAPCT uses IRT as the basis of an adaptive system with 4,489 questions from 93,341 students aged 4–13 years; 75% of variance in the difficulty of the question is explained by	CAPCT is implemented directly in elementary schools through the Math Garden online training system, proving that CAT is feasible to be applied at the primary education level	Challenges in the form of large-scale bank item development and psychometric validation across age groups; The solution is massive data collection and system robustness tests	IRT-based CAT developed for elementary schools has proven to be robust and efficient; Teachers can use it to identify students' level of understanding individually

	the characteristics of the item

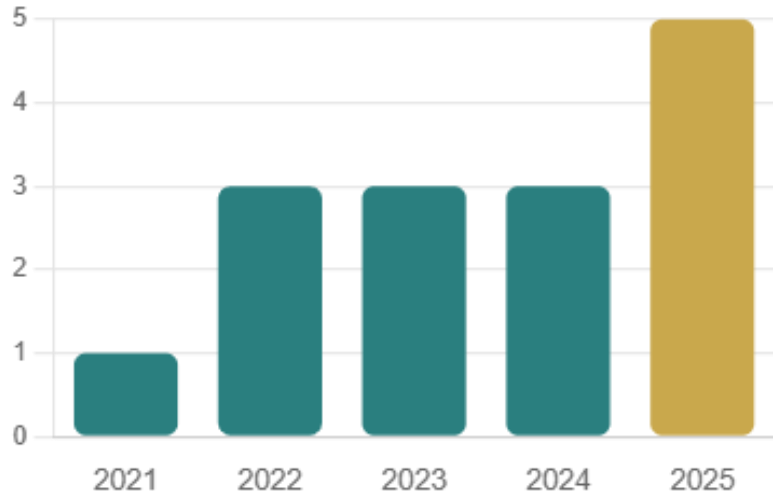


Figure 2. Publication Distribution by Year (2021–2025)

Figure 2 presents the pattern of publication distribution from 15 journals studied in the range of 2021 to 2025. It can be seen that in 2021 there was only one publication, then increased to three publications in 2022, 2023, and 2024 respectively. The highest surge occurred in 2025 with five publications, indicating that academic interest in *Computer-Based Assessment* and *Computerized Adaptive Testing* (CAT) is growing rapidly. This increasing trend is in line with the acceleration of digital transformation in the primary education sector, including the implementation *of the Computer-Based National Assessment* (ANBK) which encourages researchers to explore technology-based assessment approaches that are more adaptive, efficient, and relevant to the learning needs of the 21st century.

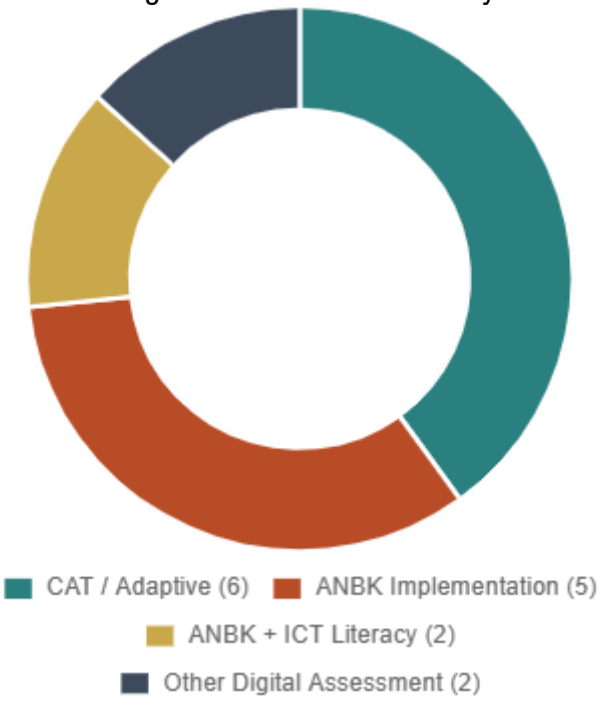


Figure 3. Journal Study Focus Proportions (Donut Diagram)

Copyright © 2026 The Author(s)

Figure 3 illustrates the distribution of study focus correlations from 15 journals analyzed using *a doughnut chart*. The study themed *Computerized Adaptive Testing* (CAT) dominated with a proportion of 40% or six journals, followed by the ANBK implementation study of 33% or five journals. Meanwhile, studies that integrate ANBK with *Information and Communication Technology* (ICT) literacy and other *digital assessment* studies are 13% or two journals each. This distribution pattern shows that the development of the CAT system based on *Item Response Theory* (IRT) is the main focus of the latest research, as well as indicating a strong relevance between CAT studies and efforts to improve the quality of assessments at the elementary school level.

Table 2. Mapping the Characteristics of 15 Synthesis Journals

No.	Author (Year)	Study Focus	Method	Key Contributions
1	Setiyowati et al. (2022)	ANBK	Descriptive Qualitative	The pre-implementation–post-ANBK stage at MIN 9 HSU went smoothly
2	Bella (2023)	ANBK	Case Studies	Infrastructure & network constraints in elementary schools; Laptop Borrowing & WiFi Solutions
3	São Paulo (2025)	ANBK	Descriptive Qualitative	ANBK SDN Tamansari I according to the procedure; Day 2 server outage
4	Ningsih et al. (2025)	ANBK	Case Studies	ANBK in remote areas; Recommendations for improving technical infrastructure
5	Saidah et al. (2023)	ANBK+TIK	Case Studies	The elimination of ICT K-13 has an impact on the low ANBK score
6	Waziana et al. (2024)	ANBK+TIK	PKM / Training	MS Office 5th grade training improves pre-ANBK computer literacy
7	Huda et al. (2024)	CAT	R&D (4-D)	3-PL IRT-based CAT system; very practical at SMK Multimedia
8	Farolai & Nurjannah (2022)	ANBK	Live Practicum	Basic computer training grade 5 improves operational skills
9	Rahmadhi Anugrah et al. (2024)	CAT	Prototype	SDN Salinization IV web CAT system: instant assessment & test efficiency
10	Ma et al. (2025)	CAT	IRT Experiment	ROAR-CAT is 40% faster efficient; Valid for students in grades 1–8
11	Sanatang & B (2021)	CAT	SDLC Waterfall	2-PL CAT IRT System; Admin Usability Score 95% & Testee 89%
12	Goto et al. (2023)	CAT	Experimental Surveys	Japanese elementary and high school students' acceptance of positive CAT; does not increase anxiety
13	Magfirah et al. (2025)	Digital Assessments	Literature Review	Quizizz as an adaptive PAI tool; Instant feedback increases participation

14	Hapsari & Labib (2025)	AKM Numeracy	IRT (PCM)	Two-tier instruments + IRT; 16% post-test increase in numeracy
15	Hogenboom et al. (2022)	CAT	Adaptive IRT	CAPCT: 93K children; 75% variance explained characteristics of programming items

Table 2 presents a comprehensive mapping of the 15 journals used as synthesis material in this *systematic review*. Each journal is identified based on author, year of publication, focus of study, methodological approach, and major contribution. Studies with a focus on *Computerized Adaptive Testing* (CAT) generally use *the Research and Development* (R&D) or *Item Response Theory* (IRT) approach, while ANBK implementation studies mostly use descriptive qualitative methods and *case studies*. The diversity of these methods reflects the rich academic perspectives that support the development of computer-based assessment instruments. This table also shows a correlation between the study of technical implementation in the field and the study of adaptive system development which is more theoretical and psychometrics-based.

Discussion

Development of CAT Implementation in Elementary Schools

Synthesis of 15 selected literature through the protocol *PRISM* reveals that the implementation of *Computerized Adaptive Testing* (CAT) at the elementary school level is experiencing a significant development trajectory, although it is not evenly distributed globally and nationally. The trend of publications showing a surge from one article in 2021 to five articles in 2025 confirms the increasing academic attention to technology-based adaptive assessments in primary education. Historically, the foundation of CAT implementation in Indonesia has been built through policies *Computer-Based National Assessment* (ANBK), which although not yet fully adaptive, has created an initial digital ecosystem that is a prerequisite for the development of a real CAT system. (Anya et al., 2025), collectively documented that the implementation of ANBK in a variety of primary school contexts, from urban to remote rural areas, has provided valuable institutional experience in the management of national-scale digital assessments.

On the system development side, a number of research has moved further towards the actual construction of CAT. (Anugrah, 2022) successfully designed and piloted a CAT based system *Copyright* © at SDN Pengasinan IV using the *Prototype*, proving that the implementation of CAT directly in elementary schools is not just a theoretical concept, but can be realized operationally. At the international level, (Farolai, 2022) develop *Computerized Adaptive Programming Concepts Test* (CAPCT) which was validated in 93,341 students aged 4–13 years, while Ma et al. (2025) introduced ROAR-CAT which was proven to be able to increase measurement efficiency by up to 40% compared to conventional tests. These two international studies confirm that CAT is not only psychometrically superior, but also practically feasible to be applied to the elementary school-age student population.

Main Challenges of CAT Implementation in Elementary Schools

Three dimensions of challenges are consistently identified from this synthesis of literature: technical and infrastructural, pedagogical and competence, and institutional and policy. From a technical aspect, central server disruptions, internet network instability, and hardware limitations are obstacles that have been repeatedly reported. Document the conditions at SDN Puger where teachers must lend personal laptops for the sake of continuous assessment, while (Bella & Syahnaidi, 2022) underlining the fragility of digital infrastructure in

Copyright © 2026 The Author(s)

schools located in remote areas. This condition reflects the digital divide (*Digital divide*) which is still very real and is a structural barrier to equitable CAT implementation.

From the pedagogical dimension, the biggest challenge lies in the low *Digital Literacy* both students and teachers. (Goto et al., 2023) empirically shows that the elimination of Information and Communication Technology (ICT) subjects from the 2013 Curriculum is directly correlated with the decline in the quality of student participation in ANBK, which implies how important it is to be ready for digital competence as a foundation before the CAT system can be optimally functioned. This argument proves that intensive basic computer training is able to significantly improve the operational skills of 5th grade students. Meanwhile, from an institutional aspect, the absence of systemic policies that integrate digital literacy into the curriculum, as well as the lack of technical training for *Proctor* and technicians, is a structural problem that cannot be overcome individually by schools (Setiyowati, 2022).

Proven Effective Solutions in Previous Research

A variety of evidence-based solutions (*evidence-based solutions*) has been identified and documented in the literature reviewed. At the technical level, strengthening network infrastructure, procurement of adequate equipment, and training *Proctor* is an intervention that is consistently recommended by (Hogenboom et al., 2022). At the pedagogical level, (Hapsari, 2025) Proving the effectiveness of skills training *Microsoft Office* for grade 5 students in improving pre-ANBK digital readiness, while the integration of adaptive platforms such as *Quizizz* Reviewed by (Scott, 2025)Offering a Solution-Based *Gamification* that have been proven to increase participation and provide *Feedback* instant to students.

At the more complex level of system development, the CAT system is based on *Item Response Theory* (IRT) model *3-Logistic Parameters* (3-PL) can be developed with a very high level of practicality, with an evaluation score of 90 students confirming its feasibility of use. (Ningsih et al., 2025) Reporting levels *Usability* The two-parameter IRT-based CAT system of logistics reaches 95% for the ease of administration aspect, indicating that the interface *User-friendly* is a design component that cannot be ignored. Furthermore, (Saidah et al., 2023) demonstrate that the development of *Bank Item* through IRT-based psychometric validation using *Partial Credit Model* (PCM) was able to increase the numeracy achievement of elementary school students by 16.01%, emphasizing the crucial role of question item quality in the success of the CAT system.

Interpretation of Results and Comparison with Other Studies

Overall, the findings of this synthesis confirm that the implementation of CAT in primary schools is multidimensional and cannot be reduced to technical issues alone. This is in line with the theoretical framework *Technology Acceptance Model* (TAM) which states that the adoption of technology is simultaneously influenced by the perception of ease of use and the perception of usefulness. (Huda et al., 2024) specifically proved that elementary school students in Japan showed positive acceptance of CAT and did not experience increased exam anxiety, a finding that contradicts the common assumption that computer-based assessments can create additional stress for young learners. These findings reinforce the argument that resistance to CAT is more rooted in extrinsic factors such as insufficient infrastructure than intrinsic psychological factors of students.

Compared to research at the higher education or secondary school level that has been more mature in implementing CAT, research in elementary schools is still in the pilot phase. However, the contribution of the research synthesized in this study provides a *comprehensive roadmap for policy developers and education practitioners to design contextual and sustainable CAT implementation strategies at the primary school level.*

Implications, Limitations, and Further Research Directions

Practically, these findings imply the need for a reorientation of curriculum policies that place digital literacy not as an optional subject, but as a cross-curricular competency that is integrated from the beginning of the classroom. Theoretically, this study contributes to strengthening the *CAT implementation framework* in the context of basic education, which has been dominated by the perspective of higher education.

However, this study has a number of limitations that need to be transparently acknowledged. First, of the 309 literature identified, only 15 met the inclusion criteria, so the representativeness of the findings needs to be confirmed through follow-up studies with wider database coverage. Second, the dominance of descriptive qualitative methods in the study of the implementation of ANBK limits the ability to generalize findings statistically. Third, the heterogeneity of research contexts between schools in Japan, Indonesia, and the Netherlands makes it difficult to make direct cross-contextual comparisons. For further research, it is recommended to conduct a *randomized controlled trial* that directly compares the effectiveness of CAT with conventional tests in elementary school students in the Indonesian context, as well as longitudinal research that measures the long-term impact of CAT implementation on students' learning outcomes and intrinsic motivation.

Conclusion

A systematic review of 15 selected literature through *the PRISMA protocol* revealed that the implementation of *Computerized Adaptive Testing* (CAT) in primary schools is in a promising but challenging phase of multidimensional research. Technically, the limitations of network infrastructure and hardware are a persistent structural barrier, especially in remote areas. Pedagogically, the low *digital literacy of* students and teachers—which was exacerbated by the elimination of ICT subjects from *the 2013 Curriculum*—is a competency prerequisite that has not been met equally. However, a number of *evidence-based solutions* have been identified, including infrastructure strengthening, intensive computer training, development of CAT systems based on *Item Response Theory* (IRT), and adaptive platform integration. This study contributes to strengthening *the* framework for the implementation of CAT in the context of basic education, which has been dominated by the perspective of higher education, while emphasizing that resistance to CAT is more extrinsic-infrastructural than psychological. However, the generalization of findings needs to be done carefully considering the limited number of literature samples, the dominance of descriptive qualitative methods, and the heterogeneity of cross-border research contexts. Controlled experimental research and longitudinal studies in the Indonesian context are highly recommended to strengthen the ecological validity of these findings.

References

Akhtar, H., Silfiasari, Vekety, B., & Kovacs, K. (2023). The Effect of Computerized Adaptive Testing on Motivation and Anxiety: A Systematic Review and Meta-Analysis. *Assessment*, *30*(5), 1379–1390. https://doi.org/10.1177/10731911221100995

Anugrah, A. R. (2022). *Perancangan Sistem Computer Assisted Test (CAT) Berbasis Website Untuk Ujian Sekolah (Studi Kasus: SDN Pengasinan IV)*.

Anya, W., Adam, M., Halford, R., Burkhardt, A. K., Kanopka, K., Chou, C., Domingue, B. W., & Yeatman, J. D. (2025). *ROAR - CAT : Rapid Online Assessment of Reading ability with Computerized Adaptive Testing*.

Bella, S., & Syahnaidi, M. F. (2022). *Analisis Pelaksanaan Asesmen Nasional Berbasis Komputer*

Copyright © 2026 The Author(s)

(Anbk) Di Sekolah Dasar. *11*(2), 226–234.

Ebenbeck, N., Bastian, M., Mühling, A., & Gebhardt, M. (2024). *Duration versus accuracy — what matters for computerised adaptive testing in schools ? February*, 3443–3453. https://doi.org/10.1111/jcal.13074

Ebenbeck, N., & Gebhardt, M. (2022). *Simulating computerized adaptive testing in special education based on inclusive progress monitoring data*. *November*. https://doi.org/10.3389/feduc.2022.945733

Farolai, N. (2022). *Pelatihan Dasar-Dasar Komputer Sebagai Persiapan Menyambut Pelaksanaan Asesmen Nasional Berbasis Komputer Pada Siswa SD Negeri 218 Congkoe*. *3*(1).

Goto, T., Kano, K., & Shiose, T. (2023). *Students ' acceptance on computer-adaptive testing for achievement assessment in Japanese elementary and secondary school*. *July*, 1–12. https://doi.org/10.3389/feduc.2023.1107341

Hapsari, W. P. (2025). *Two-Tiers Test Berbasis Everyday Life Activity Stimulus Untuk Meningkatkan Kemampuan Numerasi Akm Siswa Sekolah Dasar*. *9*(1), 129–156. https://doi.org/10.26811/didaktika.v9i1.1764

Hastuti, N. S., Sarifah, I., & Wahyuni, L. D. (2025). *Penskoran Politomus untuk Optimalisasi Asesmen Pembelajaran Matematika di Era Digital*. 493–505.

Hogenboom, S. A. M., Hermans, F. F. J., Maas, H. L. J. Van Der, Hogenboom, S. A. M., Hermans, F. F. J., & Der, H. L. J. Van. (2022). Computerized adaptive assessment of understanding of programming concepts in primary school children. *Computer Science Education*, *32*(4), 418–448. https://doi.org/10.1080/08993408.2021.1914461

Huda, A., Irfan, D., Hendriyani, Y., & Sukmawati, M. (2024). *Optimizing Educational Assessment : The Practicality of Computer Adaptive Testing (CAT) with an Item Response Theory (IRT) Approach*. *8*(March).

Imawan, O. R. (2025). *The challenges of implementing computerized adaptive testing in Indonesia*. *12*(2), 124–144. https://doi.org/10.20448/jeelr.v12i2.6677

Li, Y., Huang, C., & Liu, J. (2023). Diagnosing Primary Students' Reading Progression: Is Cognitive Diagnostic Computerized Adaptive Testing the Way Forward? *Journal of Educational and Behavioral Statistics*, *48*(6), 842–865. https://doi.org/10.3102/10769986231160668

Magfirah, I. (2025). *Transformasi Evaluasi Pembelajaran PAI berbasis Digital: Optimalisasi Media Quizizz sebagai Alat Ukur Adaptif*. *17*(1), 435–444. https://doi.org/10.47435/al-qalam.v17i1.3949

Marya Ulfha 1. (2025). *Asesmen Diagnostik dalam Pembelajaran Berdiferensiasi di Sekolah Dasar: Tinjauan Literatur Sistematik Tahun (2021-2025)*. *8*, 1115–1125.

Maulani, M. R. (2022). *Implementasi Item Response Theory Model Three-Parameter Logistics Pada Aplikasi Computerized Adaptive Test*. *16*(1), 1–9.

Ningsih, R. P., Hidayati, N., & Mustari, M. (2025). *Analisis Pelaksanaan Program Asessmen Nasional Berbasis Komputer (ANBK) di SD Negeri Senurus , Lombok Tengah*. *10*, 1920–1925.

Ogunjimi, M. O., Ayanwale, M. A., Oladele, J. I., Daramola, D. S., Jimoh, I. M., & Owolabi, H. O. (2021). Simulated evidence of computer adaptive test length: Implications for high stakes assessment in Nigeria. *Journal of Higher Education Theory and Practice*, *21*(2), 202–212.

Page, M. J., Mckenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-wilson, E., Mcdonald, S., … Moher, D. (2021). *The PRISMA 2020 statement : an updated guideline for reporting systematic reviews Systematic reviews and Meta-Analyses*. https://doi.org/10.1136/bmj.n71

Prasetya, M. G. A. (2025). *A Study on the Role of Item Response Theory in the Development of Computerized Adaptive Testing*. *3*(8), 322–336.

Saidah, J. N., Mahmudah, F. N., & E, D. M. N. (2023). *Dampak Dari Penghapusan Mata Pelajaran TIK Pada K-13 Dalam Kegiatan ANBK (Asismen Nasional Berbasis Komputer)*. *3*, 10308–10321.

Sanatang. (2021). *Sistem Tes Interaktif Berbasis Computerized Adaptive Testing (CAT)*. *4*(1), 30–34.

Setiyowati, H. (2022). *Analisis Pelaksanaan Asesmen Nasional Berbasis Komputer (Anbk) Di Madrasah Ibtidayah Negeri 9 Hulu Sungai Utara*. *6*(3), 803–818. https://doi.org/10.35931/am.v6i3.1086

Suryaman, M. (2025). *Analisis Pelaksanaan Asesmen Nasional Berbasis Komputer (Anbk) Di Sekolah Dasar Negeri Tamansari I Dalam Meningkatkan Mutu Lulusan*. *6*(10), 1473–1491.

Waziana, W., Sari, D. N., Saputra, R. H., & Kusnadi, D. (2024). *Pelatihan Keunggulan Digital Keterampilan Microsoft Office Pada Pelajar Kelas 5 Sekolah Dasar Sebagai Penunjang Anbk*. *3*.