Pertanika J. Soc. Sci. & Hum. 34 (S4): 35 - 66 (2026)

PERTANIKA JOURNALS

SOCIAL SCIENCES & HUMANITIES

Journal homepage: http://www.pertanika.upm.edu.my/

From Learning to Assessment: Development and Validation of a REACT-Based Digital Numeracy Literacy Instrument

Radeni Sukma Indra Dewi[1]*, Riska Pristiani[1], Chatree Faikhamta[2], Risky Setiawan[3], Jaka Nugraha[4], Muhammad Idris Effendi[5], Mohammad Yusuf Randy[1], and Henzi Haryanto[1]

[1]Department of Basic Education, School of Post Graduate Studies, Malang State University, 65145 Malang, East Java, Indonesia
[2]Science Education, Faculty of Education, Kasetsart University, 10900 Bangkok, Thailand
[3]Department of History Education, Faculty of Social, Law, and Political Sciences, Universitas Negeri Yogyakarta, 55281 Yogyakarta, Indonesia
[4]Primary School Teacher Education, Faculty of Teacher Training and Education, Pamulang University, 15417 South Tangerang, Banten, Indonesia
[5]Vocational Education, Graduate School, Malang State University, 65145 Malang, East Java, Indonesia

ABSTRACT

This study aimed to develop and validate a REACT (Relating, Experiencing, Applying, Cooperating, and Transferring)-based numeracy literacy assessment instrument for Indonesian elementary school students. The study employed the Four-D (Define, Design, Develop, and Disseminate) development model. During the development stage, the instrument was evaluated by five expert validators specialising in mathematics education and educational linguistics. A pilot study involving 30 elementary school students was conducted to assess item clarity and preliminary psychometric quality, followed by a field test with 120 students to examine the instrument's validity, reliability, discrimination power, and difficulty level. Content validity analysis showed high validity, with Aiken's V values ranging from 0.85 to 0.88 overall. The field testing demonstrated that most items satisfied acceptable psychometric criteria and effectively differentiated students across varying levels of numeracy competence. Unlike conventional numeracy assessments that primarily emphasise procedural competence, the proposed instrument operationalises contextual problem solving through the REACT framework by integrating experiential learning, collaborative reasoning, and knowledge transfer into the assessment process. Consequently, the

ARTICLE INFO

Article history:
Received: 13 July 2025
Accepted: 30 July 2026
Published: 25 September 2026

DOI: https://doi.org/10.47836/pjssh.34.S4.03

E-mail addresses:
radenisukmaindradewi.pasca@um.ac.id (Radeni Sukma Indra Dewi)
feductf@ku.ac.th (Chatree Faikhamta)
dosen03052@unpam.ac.id (Jaka Nugraha)
idris.tugas@gmail.com (Muhammad Idris Effendi)
yusufrandy123@gmail.com (Mohammad Yusuf Randy)
henzihariyanto@gmail.com (Henzi Haryanto)
* Corresponding author

© Universiti Putra Malaysia Press

instrument offers a valid and practical diagnostic tool for assessing contextual numeracy literacy while extending current educational assessment practices toward more authentic competency-based evaluation in elementary education.

Keywords: Development, instrument, literacy and numeracy, REACT learning strategy

INTRODUCTION

Early numeracy literacy has become one of the most essential competencies in contemporary education because it provides the cognitive foundation for mathematical reasoning, problem-solving, and informed decision-making throughout life (Bailey & Bulotsky-Shearer 2022; Blunch 2014; Mtika & Abbott 2023). Beyond mastering arithmetic procedures, numeracy literacy enables learners to interpret quantitative information, evaluate evidence, and apply mathematical reasoning across authentic contexts encountered in education, society, health, and everyday life (Carpenter et al., 2017; Jayaraman et al., 2018). Consequently, developing numeracy competence from the early years of schooling has become a strategic priority in many educational systems (Carpenter et al., 2017; McNaughton et al., 2011; Nandyala et al., 2018). Empirical studies consistently demonstrate that children who acquire strong numeracy competencies during primary education achieve higher academic performance in subsequent educational stages and exhibit greater confidence in solving unfamiliar problems (Fauzan et al., 2024; Rohmah et al., 2022). Likewise, the successful development of numeracy literacy is strongly influenced by the interaction between school learning experiences and family support. Previous studies have shown that parental expectations, home numeracy activities, and early mathematical interactions substantially contribute to children's mathematical achievement and long-term learning trajectories (Bonifacci et al., 2022; Lestari et al., 2022; Martini & Sénéchal, 2012). These findings collectively suggest that numeracy literacy should not be viewed merely as a curriculum outcome but as a multidimensional competence that develops through continuous interaction between learners, instructional practices, and authentic learning environments.

Despite its recognised importance, improving numeracy literacy remains a significant challenge in elementary education. Numerous instructional innovations, including project-based learning, ethnomathematics, realistic mathematics education, digital learning applications, and interactive mathematical comics, have demonstrated positive effects on students' mathematical understanding and engagement (Aini et al., 2024; Fauzan et al., 2024; Kustantina et al., 2022; Utama et al., 2023). These approaches emphasise contextual learning by connecting mathematical concepts with students' daily experiences, thereby promoting deeper conceptual understanding rather than procedural memorisation. Nevertheless,

evidence indicates that many elementary students continue to experience considerable difficulties when confronted with contextual mathematical problems requiring reasoning, interpretation, and flexible application of mathematical knowledge (Situmorang & Sinaga, 2022; Zulfayani et al., 2023). The decline in students' numeracy performance following the COVID-19 pandemic further illustrates the vulnerability of existing instructional and assessment practices in supporting sustainable numeracy development (Fatonah et al., 2023; Marthyane Pratiwi et al., 2024). These persistent challenges suggest that improving instructional approaches alone may be insufficient if assessment practices remain unable to capture the broader competencies required for meaningful numeracy literacy.

The growing emphasis on contextual learning has also transformed expectations regarding educational assessment. Contemporary assessment is no longer expected merely to determine whether students obtain correct answers but also to provide evidence of how learners construct, apply, and transfer mathematical knowledge across different situations (Hijriani et al., 2022). International assessment frameworks, particularly the Programme for International Student Assessment (PISA), have substantially shifted the measurement of numeracy from procedural calculation toward the application of mathematics in authentic real-world contexts. Similarly, Indonesia's Minimum Competency Assessment (AKM) was introduced to encourage schools to assess higher-order reasoning rather than routine computational skills (Hidayah et al., 2021b; Humairah et al., 2024). These developments represent important advances in educational assessment because they recognise that mathematical competence extends beyond mastery of algorithms. However, both frameworks primarily conceptualise numeracy as an individual's cognitive achievement demonstrated through responses to standardised contextual tasks. While contextual scenarios are incorporated into the assessment items, the assessment process itself remains predominantly outcome-oriented, focusing on the accuracy of individual performance rather than on the learning dimensions through which mathematical understanding is constructed and subsequently transferred to unfamiliar situations.

Consequently, although AKM and PISA have significantly contributed to advancing contextual numeracy assessment, several conceptual limitations remain when these frameworks are considered from the perspective of experiential learning. Existing numeracy assessments generally emphasise contextual problem solving as the final manifestation of competence but provide limited opportunities to examine how learners relate prior knowledge to new mathematical concepts, construct understanding through direct experience, collaborate with peers during mathematical reasoning, and transfer acquired knowledge to novel situations (Rohendi, 2019; Via & Abadi 2021). These learning processes constitute essential characteristics of

authentic mathematical competence, yet they are rarely operationalised as measurable constructs within conventional assessment instruments. As a result, students may demonstrate correct procedural or contextual solutions without necessarily exhibiting the experiential, collaborative, and transferable competencies required for meaningful mathematical literacy. This limitation becomes particularly relevant in elementary education, where mathematical understanding is actively constructed through interaction, exploration, and guided social learning rather than through individual performance alone.

The limitations of conventional assessment become increasingly evident when viewed alongside current theories of learning. Traditional paper-based assessments generally evaluate the products of learning after instruction has occurred, assuming that contextual performance adequately represents students' underlying mathematical competence. However, learning theories increasingly emphasise that meaningful understanding develops through active engagement with authentic experiences, collaborative inquiry, and repeated opportunities to apply knowledge across multiple contexts. Consequently, assessment instruments that focus exclusively on individual cognitive outcomes may underestimate important dimensions of mathematical competence developed during the learning process itself. This conceptual mismatch highlights the need for assessment approaches capable of capturing not only whether students can solve contextual problems but also whether they are able to connect prior knowledge with new situations, learn through experience, collaborate in constructing mathematical understanding, and transfer acquired knowledge to unfamiliar contexts. Addressing these dimensions requires assessment frameworks that move beyond traditional measurement paradigms toward more authentic representations of contextual numeracy competence.

This study does not simply develop another numeracy instrument; rather, it reconceptualises contextual numeracy assessment by embedding the REACT dimensions as measurable assessment constructs, thereby extending existing assessment frameworks beyond procedural and contextual mathematical performance toward experiential, collaborative, and transferable competencies. The study fills this gap by creating and testing a REACT-based numeracy literacy tool that is appropriate for Indonesian elementary school students. More and more people are talking about using REACT (Relating, Experiencing, Applying, Cooperating, and Transferring) to facilitate people learning how to read and write numbers. This model focuses on active learning and getting students involved, which facilitates them connecting math ideas to real-life situations and deepen their understanding and use of math skills. The REACT framework has been shown to facilitate people understanding math concepts and how to use them better, which is necessary for developing numeracy literacy (Faqih et al., 2022; Utama et al.,

2023). Internationally, frameworks such as the OECD PISA numeracy assessment emphasise the application of mathematics in real-world contexts, aligning with the 'Applying' and 'Transferring' dimensions of REACT. Furthermore, instrument validation draws from established psychometric principles, which inform our approach to content validity and reliability testing.

LITERATURE REVIEW

Traditional static assessments often impose high extraneous cognitive load, failing to measure students' true capabilities (Sweller, 2022). In contrast, contextual strategies like REACT align with generative learning principles, which have been proven to foster deeper understanding by actively engaging students in meaning-making processes rather than rote memorisation (Fiorella, 2023). Unlike traditional numeracy tests that often emphasise procedural fluency, the present instrument is explicitly designed around the REACT framework, which integrates real-world contextualisation, collaborative problem-solving, and knowledge transfer into the assessment structure. This distinguishes it from existing tools such as the Indonesian National Assessment (AKM), which, while measuring literacy and numeracy, does not systematically embed experiential and cooperative learning phases within the test items themselves.

The need for more authentic assessment has encouraged educational researchers to reconsider the relationship between learning theory and assessment design. Rather than functioning solely as instruments for measuring learning outcomes, contemporary assessments are increasingly expected to represent the cognitive and social processes through which knowledge is constructed. This perspective is consistent with Cognitive Load Theory, which argues that effective learning occurs when instructional and assessment activities minimise unnecessary cognitive demands while directing learners' cognitive resources toward meaningful schema construction (Sweller, 2022). Likewise, Generative Learning Theory emphasises that deep understanding develops when learners actively organise, integrate, and apply new information instead of merely recalling previously memorised knowledge (Fiorella, 2023). These perspectives suggest that assessment should move beyond verifying whether students can produce correct answers and instead capture the cognitive processes involved in constructing, applying, and transferring mathematical knowledge. Consequently, assessment instruments should be conceptually aligned with learning theories that emphasise active knowledge construction rather than passive knowledge reproduction.

Among contextual learning approaches, the REACT framework—consisting in Relating, Experiencing, Applying, Cooperating, and Transferring—offers a particularly relevant conceptual foundation for integrating learning and assessment. Existing studies have consistently demonstrated that REACT improves students' conceptual understanding, mathematical reasoning, motivation, and

engagement because mathematical ideas are introduced through meaningful experiences before being applied to authentic situations (Faqih et al., 2022; Utama et al., 2023). Each REACT dimension represents a distinct learning process. Relating encourages learners to connect new mathematical concepts with prior knowledge and familiar situations; Experiencing emphasises exploration through authentic activities; Applying requires learners to utilise mathematical concepts to solve contextual problems; Cooperating promotes collaborative reasoning and shared problem solving; and Transferring challenges students to extend acquired knowledge to unfamiliar situations. Collectively, these dimensions reflect how mathematical competence develops progressively through experience, interaction, application, and adaptation rather than through isolated procedural performance.

Despite its demonstrated effectiveness as an instructional strategy, previous studies have almost exclusively positioned REACT as a pedagogical approach for improving mathematics learning outcomes. Research has primarily examined its influence on students' achievement, motivation, classroom engagement, and conceptual understanding, while considerably less attention has been devoted to how REACT may inform the design of educational assessment instruments. Consequently, the potential of REACT to serve as a conceptual framework for assessment remains largely underexplored. This distinction is theoretically important because instructional models and assessment frameworks pursue different objectives. Instructional models describe how learning should occur, whereas assessment frameworks define which competencies should be represented, operationalised, and measured. Transforming REACT from an instructional strategy into an assessment framework therefore requires translating each learning dimension into observable assessment constructs capable of producing valid and reliable evidence regarding students' contextual numeracy competence.

This study argues that the contribution of REACT extends beyond supporting instruction by providing a multidimensional conceptual structure for numeracy assessment. Conventional numeracy instruments generally evaluate students' ability to interpret information, perform mathematical procedures, and solve contextual problems through individual responses. Although these competencies are fundamental, they provide only a partial representation of mathematical literacy because they primarily assess the outcomes of learning rather than the processes through which mathematical understanding is constructed. In contrast, a REACT-based assessment conceptualises contextual numeracy as an integrated construct encompassing students' ability to connect prior knowledge with new situations (Relating), develop understanding through authentic experiences (Experiencing), apply mathematical reasoning to meaningful contexts (Applying), construct solutions collaboratively (Cooperating), and transfer

acquired knowledge across unfamiliar situations (Transferring). Consequently, the dimensions represented within the assessment are broadened from procedural and contextual performance toward experiential, collaborative, and transferable mathematical competence.

This reconceptualisation also has important implications for contextual problem-solving assessment. Existing assessments such as AKM and PISA primarily infer students' reasoning ability from their responses to contextual tasks, whereas the present framework intentionally embeds the learning dimensions that support successful contextual reasoning into the design of the assessment itself. Rather than treating context merely as a background for mathematical problems, the proposed instrument operationalises context as a multidimensional learning environment in which students activate prior experiences, construct mathematical meaning, collaborate when appropriate, apply concepts, and demonstrate transfer of knowledge. Consequently, contextual numeracy is assessed not only as the ability to obtain correct answers in realistic situations but also as the capacity to employ mathematical knowledge through authentic cognitive and social processes. This perspective extends existing approaches to contextual assessment by integrating learning-oriented competencies directly into the measurement framework.

Accordingly, the theoretical contribution of this study lies in extending the conceptual foundation of contextual numeracy assessment. Instead of viewing assessment solely as a mechanism for evaluating mathematical performance, the proposed instrument positions assessment as a representation of how mathematical competence develops through experience, collaboration, contextual application, and knowledge transfer. In doing so, this study bridges learning theory and educational measurement by operationalising the REACT dimensions as measurable assessment constructs rather than treating them exclusively as instructional principles. This conceptual extension contributes to educational assessment by offering an alternative framework for developing numeracy instruments capable of capturing broader dimensions of authentic mathematical competence that have received limited attention in conventional assessment models.

Although previous studies have demonstrated the effectiveness of contextual learning approaches in improving students' mathematical achievement, most investigations have concentrated on evaluating learning outcomes rather than examining how contextual learning principles can inform the design of educational assessment instruments. Studies involving project-based learning, ethnomathematics, interactive mathematics media, and REACT have consistently reported improvements in students' numeracy performance, conceptual understanding, and learning motivation (Aini et al., 2024; Fauzan et al., 2024; Kustantina et al., 2022; Muliani et al., 2022; Utama et al., 2023). However, these

studies generally position assessment as a means of documenting instructional effectiveness rather than as an integral component of the learning framework itself. Consequently, the assessment instruments employed in previous research largely continue to measure conventional cognitive outcomes, providing limited evidence regarding how students relate prior knowledge to new mathematical concepts, develop understanding through authentic experiences, collaborate in mathematical reasoning, and transfer acquired knowledge across unfamiliar situations. This indicates that the conceptual relationship between contextual learning theory and educational assessment remains insufficiently developed within the current literature.

A second gap concerns the operationalisation of contextual numeracy assessment. Existing assessment frameworks, including nationally implemented assessments such as AKM and internationally recognised assessments such as PISA, have substantially advanced the measurement of contextual mathematical reasoning by emphasising authentic problem-solving situations. Nevertheless, contextualisation within these assessments is primarily represented through realistic problem contexts rather than through multidimensional learning processes. In other words, students' mathematical competence is inferred from their responses to contextual tasks, while the experiential, collaborative, and transferable dimensions through which contextual understanding develops remain largely outside the scope of measurement. Consequently, current assessment practices provide valuable evidence regarding students' mathematical performance but comparatively limited information regarding broader competencies that characterise authentic numeracy learning. This unresolved issue becomes particularly important in elementary education, where mathematical understanding develops through continuous interaction among prior experience, exploration, collaboration, application, and knowledge transfer.

Building upon these limitations, the present study proposes that the REACT framework offers an opportunity to extend the conceptual foundation of contextual numeracy assessment. Rather than employing REACT solely as an instructional strategy, this study operationalises its five dimensions as measurable assessment constructs that guide the development of assessment items and scoring indicators. From this perspective, contextual numeracy is conceptualised not only as the ability to solve mathematical problems situated in realistic contexts but also as the capacity to connect prior experiences with new concepts, construct understanding through authentic engagement, collaborate in mathematical reasoning, apply concepts meaningfully, and transfer knowledge across different situations. This reconceptualisation shifts the role of assessment from merely documenting learning outcomes toward representing the multidimensional competencies involved in meaningful mathematical learning.

Accordingly, the theoretical contribution of this study extends beyond

the development of a psychometrically valid assessment instrument. Conceptually, this research contributes to the field of educational assessment by integrating learning-oriented competencies into the measurement of contextual numeracy. Whereas conventional numeracy assessments predominantly emphasise procedural fluency, mathematical reasoning, and contextual problem solving, the proposed framework broadens the construct of numeracy assessment by explicitly incorporating experiential engagement, collaborative reasoning, and knowledge transfer as measurable dimensions of competence. By operationalising these dimensions within the assessment design, the study establishes a closer alignment between contemporary learning theories and educational measurement, thereby extending existing perspectives on authentic numeracy assessment. Consequently, the contribution of this study lies not only in producing a reliable assessment instrument but also in proposing an alternative conceptual model for assessing contextual numeracy in elementary education.

The novelty of this study therefore resides in two complementary aspects. First, unlike previous studies that primarily investigated REACT as a pedagogical strategy to improve mathematics instruction, this study reconceptualises REACT as a framework for educational assessment by translating its five learning dimensions into measurable assessment constructs. Second, while existing numeracy instruments generally evaluate contextual mathematical performance through individual cognitive responses, the proposed instrument embeds experiential learning, collaborative reasoning, authentic application, and knowledge transfer within the assessment structure itself. Consequently, the instrument extends existing approaches to contextual numeracy assessment by providing a more comprehensive representation of mathematical competence that reflects both cognitive performance and learning-oriented processes. This dual contribution differentiates the present study from previous instrument-development research and positions it within the broader discourse on authentic educational assessment.

Based on these considerations, this study aims to develop and validate a REACT-based numeracy literacy assessment instrument for elementary school students in Indonesia using the Four-D development model. Specifically, the study seeks to examine the content validity, construct quality, and psychometric properties of the proposed instrument while evaluating its capacity to represent the multidimensional competencies embedded within the REACT framework. Through this approach, the study is expected to provide both a practically applicable assessment instrument for elementary mathematics education and a conceptual contribution to the continuing development of contextualised educational assessment.

METHODOLOGY

The 4D development method is used in this study. It has four steps: define, design,

develop, and disseminate. In the first step, we looked at the needs and problems that elementary school students have with their math and reading skills. The first step was to look over the curriculum and other relevant materials to get a sense of what numeracy and literacy skills needed to be learned. At this point, the students' traits were also looked at so that the development of the instrument could be based on what they knew and needed. To make sure the instrument was useful and made sense, they also looked at math topics that people use in their daily lives.

The next step is design, where the instrument's design starts based on the needs analysis that was done in the define stage. The REACT (Relating, Experiencing, Applying, Cooperating, and Transferring) learning strategy is used to arrange the complex multiple-choice questions that make up this instrument according to the indicators of numeracy literacy. The design process includes making an instrument lattice that outlines the test's purpose, content, and structure. Each question is meant to measure students' ability to think critically and solve problems in real life. The instrument is made so that children in elementary school can easily and clearly understand and answer the questions. The development stage involved constructing a model of the tool and getting it approved by five math education specialists. Validators then looked at the prototype of the tool numerous times and got comments from experts to make it more accurate. Content validity was quantified using Aiken's V formula. A score of $V > 0.80$ was considered valid. A small pilot test was also done with 30 elementary school pupils to assess how clear the instrument was and make sure it could be utilised properly in a math class. The results of this pilot test were used to improve the tool so that it could be used by more people.

At the distribute stage, the modified tool was tested again on a larger number of 120 students to check if it was trustworthy and valid in real life. We utilised SPSS to analyse the results of this pilot test, look at factors like the validity of the content, the dependability of the results, the ability to tell the difference, and the difficulty of the questions. With this analysis, each thing on the instrument is better so that it provides consistent and reliable measurement results. There is a lot of information available for each step of the 4D development technique. This means that other scientists can do the same things and get the same findings. We had professionals check it out and did field tests to make sure that the instrument can be used properly in different types of schools. The REACT-based numeracy literacy tool that was produced as part of this study should be a good and effective technique to test primary school pupils' arithmetic skills in line with national education requirements. Each stage in the 4D development method is represented in more detail in Figure 1.

Sample Size Justification

The sample sizes employed in this study were determined based on methodological recommendations for educational instrument

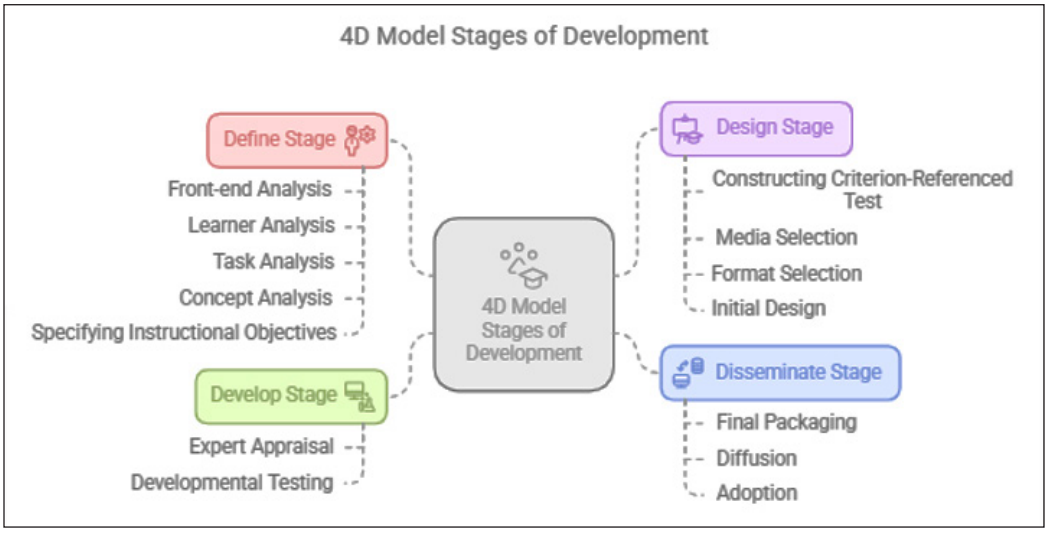

Figure 1. 4D model stages of development

development rather than statistical hypothesis testing. During the content validation stage, the instrument was evaluated by five expert validators consisting of specialists in mathematics education, educational assessment, and language. This number is considered adequate for content validity evaluation using Aiken's V because a panel of between five and ten experts is generally regarded as sufficient to provide stable and reliable judgments regarding item relevance, clarity, and representativeness (Aiken, 1985; Lynn, 1986).

A pilot study involving 30 elementary school students was subsequently conducted to examine the clarity, readability, comprehensibility, and preliminary functioning of the assessment of items. Pilot samples of approximately 20-30 participants are commonly recommended in instrument development studies to identify ambiguous wording, administration problems, and initial psychometric characteristics before large-scale implementation (Johanson & Brooks, 2010).

For the field-testing phase, the instrument was administered to 120 elementary school students. This sample size satisfies commonly accepted recommendations for preliminary psychometric evaluation, including item validity, internal consistency, discrimination index, and difficulty index analyses. Methodological literature on educational measurement indicates that samples exceeding 100 participants are generally considered adequate for item analysis and reliability estimation during the development of educational assessment instruments (DeVellis, 2017). Therefore, the selected sample sizes were considered methodologically appropriate for each phase of the Four-D development model because each stage served different objectives, namely expert judgment for content validation, pilot testing for instrument refinement, and field testing for psychometric evaluation.

Ethical Considerations

This study was conducted in accordance with ethical principles governing educational research involving children. Prior to data collection, permission to conduct the study was obtained from the participating elementary school and the relevant educational authorities. Parents or legal guardians of all participating students were informed about the objectives, procedures, and voluntary nature of the study, and informed consent was obtained before students participated in the research. In addition, verbal assent was sought from the students to ensure that participation was based on their willingness and understanding. Participants were informed that they could withdraw from the study at any stage without academic consequences. To protect participants' privacy, all collected data was anonymised, treated confidentially, and used exclusively for research purposes. No personally identifiable information was included in the data analysis or publication of the findings.

RESULTS AND DISCUSSION

Define Stage

At the define stage in the development of REACT-based numeracy literacy instruments for primary school students, the first step is to analyse the needs and problems related to numeracy literacy. Based on the literature review and curriculum study, it was found that numeracy literacy is one of the essential competencies for 21st century students, which involves not only numeracy skills, but also critical thinking, analytical and problem-solving skills in everyday contexts. Unfortunately, the results of international studies such as PISA show that Indonesian students' numeracy literacy is still below average, signaling an urgent need to improve this skill at the basic education level.

Furthermore, a curriculum analysis was conducted to understand the core competencies that need to be achieved in numeracy literacy according to national standards. The curriculum underlines the importance of contextualised numeracy skills, which involve the application of mathematics in real life. Based on this analysis, common daily mathematics topics, such as basic number operations, measurement and data representation, were selected as the main focus in developing the instrument to make it relevant to students' lives and to enhance their critical and analytical thinking skills. Analysing student characteristics is also an important part of this stage. By understanding that elementary school students are in the concrete cognitive stage of development, the instruments designed should be simple and contextual, using language that is easy to understand and situations that are close to students' daily experiences. This method should facilitate students understanding the questions better and feel more like they are part of the learning process.

The define stage also included looking at the research on effective learning models to facilitate people becoming better at math and reading. The REACT (Relating, Experiencing, Applying, Cooperating, and Transferring) strategy was chosen because

it lets students connect math ideas to real-life situations, work together with friends, and use what they know in new ways. People think that this strategy will make learning math and reading more interesting and useful for students. The define stage gives us a lot of information that we can use to make tools that fit the curriculum and the students' needs, and that also use new ways of teaching. It is hoped that this will facilitate making numeracy literacy tools that are useful, work well, and can fully measure what students can do.

Design Stage

During the design phase of making REACT-based numeracy literacy tools for elementary school students, a number of design steps are taken to create question grids and tools that meet the numeracy literacy indicators that were set during the defined phase. Initial display of REACT-Based Numeracy Literacy AKM Questions is shown in more detail in Figure 2. Making the instrument grid is the first step in the design process. The grid includes the goals, the skills that will be measured, the question of indicators, and the types of questions that will be used. Each question was based on one of the five parts of the REACT strategy: relating, experiencing, applying, cooperating, and transferring. This made sure that the test not only measured basic math skills but also critical thinking skills and the ability to use them in real life.

In this design process, the questions are made in the form of complex multiple choices that require students to analyse and interpret information before choosing an answer. Complex multiple choices were chosen because they are able to provide a higher challenge, so students need to think more deeply and be actively involved in the problem-solving process. Each question includes daily life contexts that are close to students' experiences, such as situations in families, schools, and neighborhoods, so that students feel more connected to the material being tested. To facilitate instrument testing and maintain consistency, the initial draft of the instrument was organised into several forms of questions with varying difficulty levels, which included questions with low, medium, and high difficulty levels. This variation aims to measure students' abilities comprehensively, from basic skills to more in-depth analytical abilities. The questions were then categorised by difficulty level so that the results of the instrument evaluation could cover a wide spectrum of student abilities.



Figure 2. Initial display of REACT-based numeracy literacy AKM questions

Also, the instructions for how to answer the questions are written in simple, clear language that is appropriate for elementary school students. Each clue is set up and visualised in Figure 3 so that students can easily understand what they need

to do, and each problem has pictures or real-life examples that facilitate students understand the problem's context, and each problem is accompanied by illustrations or real examples that support students' understanding of the context of the problem which is visualised in Figure 4. Students can better understand the problem and be more motivated to answer it when they can see it in pictures.

Students can better understand the problem and be more motivated to answer it when they can see it in pictures.

Students will get a notification on their app screen if they answer the question correctly or incorrectly, as shown in Figure 5.

The outcome of this design stage is a tool in the form of an initial prototype that has a range of numeracy and literacy questions that follow the REACT method and are organised according to the grids that have been made. We are ready to test this prototype again in the next step, with the goal of making a tool that is valid, reliable, and useful for measuring elementary school students' math and reading skills.

Stage of Development

At the development stage of the research on making REACT-based math and reading tools for elementary school students, the tools that were designed at the design stage are then made into prototypes and tested to make sure they work, are valid, and reliable. Prototyping, expert validation, and limited student testing are some of the important steps that happen at this stage. The first thing that needs to be done at this stage is to make instrument prototypes based on the grids and question designs that have been chosen.

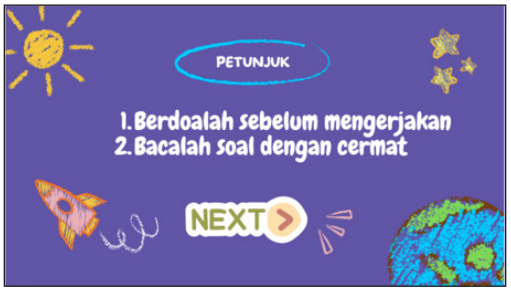

Figure 3. Display of problem-solving instructions



Figure 4. Display of REACT-based numeracy literacy AKM questions and text

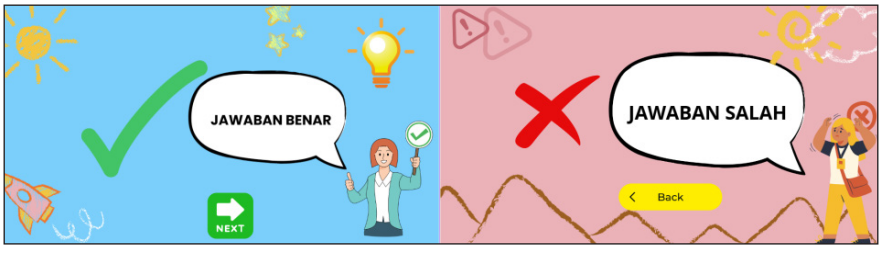

Figure 5. Incorrect and correct answer display

The prototype was made up of complicated multiple-choice questions with pictures and situations that were important to students' everyday lives. The tool was made to include all the parts of the REACT strategy, such as making connections to students' lives, using the strategy in real life, and working together to solve problems. The prototype also has questions of different levels of difficulty, from easy to hard, so that students with different levels of understanding can use it.

Furthermore, validation was carried out by mathematics education experts and linguists to assess the suitability and accuracy of the instrument. Experts provided input related to language, context suitability, and the ability of the questions to measure numeracy literacy in accordance with the desired indicators. Content validity was quantified using Aiken's V formula. A score of $V > 0.80$ was considered valid. From the validation results, some suggestions for improvement were obtained, especially in terms of language clarity and illustrations on some questions to make it easier for students to understand. After improvement, the instrument prototype was reviewed until it met high validity standards and was considered suitable for field testing. The results of mathematics education expert validation are presented in following Table 1.

Based on Table 1, the overall validity score is 0.85 (High Validity). However, based on qualitative feedback, experts suggested minor revisions, such as adding more contextual questions relevant to students' daily lives and increasing the variety of difficulty levels to better measure different depths of understanding. The results of language and linguistic expert validation are presented in following Table 2.

Table 2 shows a high validity score of 0.88 from language experts. Qualitative feedback indicated that while the grammar is precise, compound sentences should be simplified to suit the cognitive level of elementary students. As per an article by Boateng et al. (2021)), the numeracy literacy instrument tested and verified according to REACT is now ready for use in learning as

Table 1
Results of mathematics education expert validation

Validation Aspect	Aiken's V Index	Category
Content Compatibility with the Curriculum	0.85	Valid
Clarity of Language	0.90	Valid
Compatibility with Numeracy Literacy Indicators	0.88	Valid
Accuracy of Difficulty Level	0.80	Valid
Consistency of Assessment	0.83	Valid
Creativity and Innovation of Questions	0.78	Medium Validity
Suitability with REACT Strategy	0.82	Valid
Cognitive Level Appropriateness	0.89	Valid
Contextual Relevance	0.89	Valid
Total Average Validity	0.85	High Validity

Table 2
Results of language and linguistic expert validation

Validation Aspect	Aiken's V Index	Category
Clarity and Precision of Language	0.84	Valid
Use of Math Terms	0.88	Valid
Language Appropriateness for Age	0.92	Valid
Readability of Questions	0.85	Valid
Consistence in Language Style	0.87	Valid
Grammatical Appropriateness	0.89	Valid
Simplicity and Straightforwardness of Language	0.86	Valid
Total Average Validity	0.88	High Validity

this stage of development is finalised. The final prototype of the site was then polished based on the constrained trial results and ready to be implemented in a wider context in the disseminate phase. The results revealed that the results of the develop phase reveal that this instrument has been highly polished and constructed, thus it is an efficient approach to measure the numeracy and literacy skills of students in elementary school who are educated in an environment that is relevant to their own lives.

Stage of Dissemination

At the disseminate stage of the development process of the REACT-based numeracy literacy tool for primary school children, the enhanced tool that can be obtained from the results of earlier trials is subsequently used on a larger group of children to test its validity and reliability in an enlarged context. The goal at this stage is to make the tool measure students' math and reading skills consistently and accurately in a more representative environment. Word gets around by using the tool in a sample of students across different schools and backgrounds to gather representative results. In this application, the instrument is utilised to test the understanding of students on numbers and how they are able to use them in real life, for instance, solving problems in a project-based learning environment. How the instrument is administered is organised in a way in which all the students have an equal chance to try answering the questions. The teacher in charge is also directed to make sure there is proper use of the instrument according to procedures set.

We obtained and analysed data on instrument test scores in SPSS to get an empirical description of how valid and reliable the instrument is. We performed a reliability analysis to find out if the findings were the same for all of the questions in the instrument. We also carried out a validity analysis to validate that each question, in fact, measures aspects of numeracy literacy with respect to the competencies desired. The analysis shows that the tool is consistent because coefficient values are in the normal range. It indicates that the tool consistently measures students' capability. A statistical test was also performed to ascertain how

diverse and challenging the questions were. We also reworded and asked more rigorous questions which were either too simple or too hard to discriminate against. Such an examination ensured that the instrument had a satisfactory mix of questions that could discriminate among students according to ability level and at the correct difficulty level. Such items that are underperformed are flagged and, if necessary, reworded to improve the instrument.

The instrument is then ready to be used as a primary school pupil in mathematics and reading assessment once it is disseminated and interpreted. Especially in the instance of REACT-based education, this is true. The instrument was also shown to teachers so they can provide comments and see how it could be implemented in everyday learning. Teachers who used the instrument on their class lessons testified that the REACT-based problem-solving method facilitateed students to understand and utilise mathematics principles in real life and rendered them more participative and motivated learners. Generally, the result of the disseminate stage shows that this REACT-based numeracy literacy tool is a valuable and trustworthy method for identifying how good the students are at mathematics. Thus, this instrument is now ready to be utilised by numerous individuals to screen numeracy and literacy skills among basic education students. This will improve the quality of education more relevant to the real-life needs of the students. To screen this, REACT-based instrument for numeracy and literacy, researchers perform several tests such as the following:

Validity Test

Validity test results are shown in Table 3. Based on the validity test data of 30 questions tested on 20 students, each question was assessed using Aiken's V method to determine its suitability with the numeracy literacy indicators measured. The Aiken's V value obtained illustrates the validity level of each question, with the following criteria: questions with a score above 0.85 are considered highly valid, indicating that the question is very suitable for the purpose of measurement and does not require improvement. Some questions, such as question numbers 6, 11, 21, and 26, fall into this category, showing validity scores above 0.90 and are considered high quality. Overall, the results of this validity test provide guidance to maintain the questions that have met the standard and improve the questions that are below the threshold so that the overall instrument can provide accurate and reliable results in measuring students' numeracy literacy skills.

Reliability Test

Reliability test results are shown in Table 4. Based on the results of the reliability test on 30 questions tested on 20 students, the reliability value of each question was measured using Cronbach's Alpha coefficient to assess internal consistency. Questions with reliability values above 0.85 are considered highly reliable, indicating that these questions have very high consistency. These questions are recommended to be revised in order to improve the reliability of measurement in the instrument as a whole.

Table 3
Validity test results

Question Number	Validity (Aiken's V)	Description	Question Number	Validity (Aiken's V)	Description
1	0.78	Valid	16	0.80	Valid
2	0.65	Quite Valid	17	0.66	Quite Valid
3	0.82	Valid	18	0.73	Valid
4	0.60	Needs Revision	19	0.65	Quite Valid
5	0.87	Valid	20	0.78	Valid
6	0.90	Very Valid	21	0.91	Very Valid
7	0.75	Valid	22	0.58	Needs Revision
8	0.62	Quite Valid	23	0.85	Valid
9	0.85	Valid	24	0.69	Quite Valid
10	0.68	Quite Valid	25	0.81	Valid
11	0.92	Very Valid	26	0.93	Very Valid
12	0.77	Valid	27	0.76	Valid
13	0.55	Needs Revision	28	0.64	Quite Valid
14	0.88	Valid	29	0.82	Valid
15	0.70	Valid	30	0.67	Quite Valid

Table 4
Reliability test results

Question Number	Reliability (Cronbach's Alpha)	Description	Question Number	Reliabilitas (Cronbach's Alpha)	Description
1	0.82	Reliable	16	0.81	Reliable
2	0.76	Quite Reliable	17	0.72	Quite Reliable
3	0.84	Reliable	18	0.82	Reliable
4	0.69	Needs Revision	19	0.70	Quite Reliable
5	0.88	Very Reliable	20	0.83	Reliable
6	0.91	Very Reliable	21	0.90	Very Reliable
7	0.80	Reliable	22	0.66	Needs Revision
8	0.74	Quite Reliable	23	0.85	Reliable
9	0.85	Reliable	24	0.73	Quite Reliable
10	0.77	Quite Reliable	25	0.87	Reliable
11	0.89	Very Reliable	26	0.92	Very Reliable
12	0.83	Reliable	27	0.81	Reliable
13	0.68	Needs Revision	28	0.75	Quite Reliable
14	0.86	Reliable	29	0.83	Reliable
15	0.79	Quite Reliable	30	0.78	Quite Reliable

Overall, the results of this reliability test provide an overview of which questions are consistent and can be used, and which questions require revision to improve the quality of the instrument in measuring students' numeracy literacy.

Internal consistency reliability was evaluated at the instrument level using Cronbach's Alpha coefficient rather than at the individual item level. The overall Cronbach's Alpha coefficient obtained for the instrument exceeded the commonly accepted threshold of 0.85, indicating satisfactory internal consistency among the assessment items. Item-level evaluation was conducted using corrected item-total correlation and item discrimination analyses to determine the contribution of each item to the overall scale consistency.

Discriminating Power Test

Discriminating power test results are shown in Table 5. Based on the results of the differentiating power test on 30 questions tested on 20 students, the differentiating power value is obtained which demonstrates the extent to which each question is able to distinguish students with high and low abilities. Questions with a differentiating power value above 0.60, such as question numbers 5, 6, 11, 21, and 26, are categorised as demonstrate proficiency in differentiating student ability levels. These questions demonstrate a very strong ability to identify differences in ability between students. Questions with differentiating power values between 0.40 to 0.59, such as question numbers 1, 3, 7, and 23, are considered good. These questions have met the criteria

Table 5
Discriminating power test results

Question Number	Discrimination Index	Description	Question Number	Reliability (Cronbach's Alpha)	Description
1	0.45	Good	16	0.81	Reliable
2	0.30	Mediocre	17	0.72	Quite Reliable
3	0.52	Good	18	0.82	Reliable
4	0.25	Poor	19	0.70	Quite Reliable
5	0.60	Excellent	20	0.83	Reliable
6	0.65	Excellent	21	0.90	Very Reliable
7	0.40	Good	22	0.66	Needs Revision
8	0.28	Mediocre	23	0.85	Reliable
9	0.55	Good	24	0.73	Quite Reliable
10	0.32	Mediocre	25	0.87	Reliable
11	0.62	Excellent	26	0.92	Very Reliable
12	0.47	Good	27	0.81	Reliable
13	0.20	Poor	28	0.75	Quite Reliable
14	0.50	Good	29	0.83	Reliable
15	0.35	Mediocre	30	0.78	Quite Reliable

of being sufficient to differentiate between high and low ability students, so they can be used without significant revision.

Overall, the results of this differentiating power test provide information about the quality of the questions in differentiating students' abilities. Questions in the good and excellent categories can be used directly, while the ones with sufficient and insufficient differentiating power require adjustments in order to function optimally in measuring students' numeracy literacy skills.

Level of Difficulty Test

Levels of difficulty test results are revealed in Table 6. A test with thirty questions ranging in difficulty from extremely difficult to very easy was administered to twenty students. The results demonstrate the breadth of topics covered by the questions. Questions with difficulty scores below 0.40, such as 4, 8, 13, 17, 22, and 24, are deemed challenging. Only a small percentage of students correctly answer these questions, indicating that they are extremely difficult. The results of this difficulty make it evident which questions need to be revised or adjusted to better balance the easy and hard categories. Too-easy or too-hard questions must be changed in order for the assessment tool to fairly evaluate students' general proficiency. While questions that are too easy or too difficult should be revised, moderate-level questions can be used directly. Items with lower discrimination indices (e.g., Items 4, 13, 22) often require 'Transferring' knowledge to novel situations, a cognitively demanding REACT phase. This

Table 6
Results of the level of difficulty test

Question Number	Difficulty Level	Description	Question Number	Difficulty Level	Description
1	0.70	Medium	16	0.72	Medium
2	0.85	Easy	17	0.38	Difficult
3	0.55	Medium	18	0.63	Medium
4	0.40	Difficult	19	0.82	Easy
5	0.65	Medium	20	0.54	Medium
6	0.90	Easy	21	0.88	Easy
7	0.50	Medium	22	0.42	Difficult
8	0.35	Difficult	23	0.77	Medium
9	0.75	Medium	24	0.32	Difficult
10	0.80	Easy	25	0.69	Medium
11	0.45	Difficult	26	0.90	Easy
12	0.60	Medium	27	0.56	Medium
13	0.30	Difficult	28	0.36	Difficult
14	0.78	Easy	29	0.74	Medium
15	0.68	Medium	30	0.80	Easy

suggests that students struggle with applied transfer, highlighting an area for targeted instructional focus. Conversely, items aligned with 'Experiencing' and 'Applying' showed higher engagement and success rates, indicating that contextual familiarity enhances performance. Furthermore, the correlation between difficulty levels and REACT components reveals that tasks involving 'Cooperating' and 'Transferring' consistently presented greater challenges, whereas 'Relating' to familiar contexts generally yielded higher success.

Validity and Psychometric Quality of the REACT-Based Numeracy Literacy Instrument

The present study aimed to develop and validate a digital numeracy literacy assessment instrument grounded in the REACT (Relating, Experiencing, Applying, Cooperating, and Transferring) framework for elementary school students. The findings demonstrate that the developed instrument possesses satisfactory psychometric properties, including high content validity, acceptable item validity, appropriate levels of item difficulty and discrimination, and strong internal consistency. Collectively, these findings indicate that the instrument is capable of measuring contextual numeracy literacy in a valid and reliable manner while maintaining item characteristics appropriate for elementary learners. More importantly, the psychometric evidence suggests that the REACT framework can be operationalised into measurable assessment constructs without compromising measurement quality. This finding is particularly significant because previous studies have primarily investigated REACT as a pedagogical approach to improve mathematics learning, whereas limited attention has been devoted to its application as the conceptual foundation of educational assessment (Faqih et al., 2022; Utama et al., 2023). Consequently, the present study extends the existing literature by demonstrating that REACT is not only effective for organising classroom instruction but also suitable for guiding the systematic development of contextualised assessment instruments.

The high content validity obtained through expert evaluation indicates that the assessment of items adequately represents the intended dimensions of contextual numeracy embedded within the REACT framework. The consistent agreement among expert validators suggests that the developed indicators successfully translate abstract learning principles into observable assessment tasks appropriate for elementary school students. This finding supports previous studies emphasising that contextual mathematics learning becomes more meaningful when mathematical concepts are closely connected with students' daily experiences and authentic situations (Faqih et al., 2022; Muliani et al., 2022; Utama et al., 2023). However, the contribution of the present study differs from these earlier investigations in an important respect. Previous studies primarily examined how REACT enhances students' learning outcomes through instructional activities, whereas the current study demonstrates

that the same theoretical principles can be transformed into measurable assessment indicators. Thus, the role of REACT is expanded from an instructional model into a conceptual framework for educational assessment, enabling contextual numeracy competencies to be represented more systematically during the measurement process.

The item analysis further strengthens the quality of the developed instrument. The satisfactory item validity coefficients indicate that individual assessment items consistently contribute to measuring the overall construct of contextual numeracy literacy. Likewise, the combination of acceptable discrimination indices and predominantly moderate difficulty levels demonstrates that the instrument effectively differentiates students with varying levels of numeracy competence while avoiding items that are either excessively easy or excessively difficult. Such balanced psychometric characteristics are essential for educational assessments because they improve the accuracy of student classification and provide more dependable information regarding students' mathematical abilities. These findings are consistent with previous studies on numeracy instrument development, which emphasise that well-functioning assessment items should simultaneously demonstrate acceptable validity, discrimination, and difficulty characteristics to ensure accurate measurement of students' competencies (Hamdu et al., 2023; Muliani et al., 2022). The present results therefore confirm that contextualised assessment items can achieve satisfactory psychometric quality while maintaining authentic mathematical contexts.

The overall internal consistency of the developed instrument further supports its suitability for assessing elementary school students' numeracy literacy. High internal consistency indicates that the assessment of items collectively measures a coherent construct rather than functioning as isolated questions. This finding is particularly important because the instrument integrates five interrelated dimensions of the REACT framework, each representing different aspects of contextual learning. Despite measuring multiple learning processes, the obtained reliability evidence suggests that these dimensions operate cohesively as components of a broader contextual numeracy construct. This result strengthens the conceptual assumption that contextual numeracy literacy is multidimensional but internally integrated, encompassing students' abilities to relate prior knowledge, engage in authentic experiences, apply mathematical concepts, collaborate in problem solving, and transfer knowledge to unfamiliar situations. Consequently, the psychometric evidence supports not only the technical quality of the instrument but also the coherence of the underlying theoretical framework.

Beyond demonstrating acceptable measurement properties, the findings also provide evidence regarding the feasibility of integrating contextual learning theory into educational assessment design.

Traditionally, assessment development has focused primarily on measuring cognitive outcomes through procedural or contextual problem-solving tasks. In contrast, the present instrument was intentionally designed to represent broader learning-oriented competencies embedded within the REACT framework. This design approach allows assessment items to capture not only students' mathematical performance but also the cognitive processes associated with meaningful learning experiences. Such an approach aligns with the argument that authentic assessment should represent the processes through which knowledge is constructed rather than merely documenting the correctness of final answers. Therefore, the satisfactory psychometric performance observed in this study suggests that contextual learning principles can be operationalised into educational measurement without sacrificing technical assessment quality.

Taken together, these findings provide empirical support for the development of contextualised digital numeracy assessment instruments that integrate learning theory and educational measurement. While previous research has successfully demonstrated the instructional effectiveness of REACT-based learning environments (Faqih et al., 2022; Utama et al., 2023), the present study extends this body of knowledge by validating REACT as a viable framework for assessment development. Accordingly, the contribution of this study extends beyond the production of a valid and reliable digital instrument. It also demonstrates that contextual learning principles can serve as measurable assessment constructs, thereby strengthening the conceptual relationship between contextual pedagogy and educational assessment. This extension is particularly relevant for elementary mathematics education, where meaningful assessment should capture not only what students know but also how they construct, apply, and transfer mathematical understanding across authentic contexts.

REACT as a Framework for Contextualised Numeracy Assessment

One of the most important findings of this study is that the REACT framework can be extended beyond its conventional role as an instructional strategy to function as a conceptual framework for educational assessment. Previous studies have consistently reported that REACT improves students' mathematical achievement, conceptual understanding, problem-solving ability, and learning motivation because mathematical concepts are introduced through authentic experiences and meaningful contexts (Faqih et al., 2022; Utama et al., 2023). However, these investigations have primarily evaluated REACT from the perspective of instructional effectiveness, using assessment merely as an instrument for measuring learning outcomes. Consequently, relatively little attention has been devoted to examining how the conceptual dimensions of REACT can be operationalised into measurable assessment constructs. The present study addresses this gap by demonstrating that the five dimensions of REACT—Relating,

Experiencing, Applying, Cooperating, and Transferring—can be systematically translated into assessment indicators capable of measuring contextual numeracy competencies. Therefore, this study extends previous research by positioning REACT not only as a pedagogical approach but also as a framework for designing authentic educational assessment.

This conceptual shift has important implications for the way contextual numeracy is assessed. Conventional numeracy assessments generally emphasise students' ability to solve contextual mathematical problems through individual cognitive performance. Such assessments provide valuable information regarding procedural competence, mathematical reasoning, and problem-solving ability. Nevertheless, the findings of the present study suggest that contextual numeracy encompasses broader competencies than those represented by conventional cognitive assessments alone. Through the integration of REACT, the developed instrument captures students' abilities to connect prior experiences with new mathematical concepts, construct understanding through authentic activities, apply knowledge in meaningful situations, collaborate in mathematical reasoning, and transfer acquired knowledge to unfamiliar contexts. Consequently, contextual numeracy is represented not merely as the ability to produce correct answers but as the capacity to engage in the cognitive and social processes through which mathematical understanding is developed. This broader conceptualisation supports current perspectives emphasising that authentic assessment should capture meaningful learning processes in addition to learning outcomes.

The findings also provide an opportunity to reconsider the relationship between contextual learning and educational assessment. Existing studies involving contextual mathematics learning have generally assumed that improvements in contextual understanding are adequately reflected by students' performance on contextualised test items (Muliani et al., 2022; Utama et al., 2023). While this assumption remains valid, the present findings indicate that contextual assessment can be strengthened by explicitly incorporating the learning dimensions responsible for generating contextual understanding. Rather than treating context solely as the setting in which mathematical problems are presented, the proposed instrument operationalises context as a multidimensional learning environment in which students activate prior knowledge, engage with authentic experiences, collaborate with peers, apply mathematical reasoning, and demonstrate knowledge transfer. Accordingly, assessment becomes more closely aligned with the processes through which mathematical competence is developed, thereby strengthening the authenticity of contextual numeracy measurement.

Another important contribution of this study concerns its relationship with existing large-scale numeracy assessments, particularly Indonesia's Minimum

Competency Assessment (AKM) and the programme for International Student Assessment (PISA). Both assessment systems have significantly influenced contemporary numeracy assessment by emphasising the application of mathematical knowledge in authentic real-world situations. Their contextual orientation has encouraged schools to move beyond procedural computation toward mathematical reasoning and everyday problem solving. Nevertheless, contextualisation within these frameworks is primarily reflected in the design of realistic mathematical tasks, while the experiential and collaborative processes through which students construct contextual understanding remain largely outside the measurement framework. Consequently, these assessments predominantly evaluate the outcomes of contextual reasoning rather than the broader learning processes that support contextual mathematical competence.

The present study does not seek to replace existing assessment frameworks such as AKM or PISA. Instead, it complements and extends their conceptual orientation by embedding the REACT dimensions directly into assessment design. In the developed instrument, contextual situations serve not merely as realistic problem settings but as opportunities for students to demonstrate how they relate mathematical ideas to prior experiences, construct meaning through authentic engagement, collaborate in reasoning, apply mathematical concepts, and transfer understanding across different contexts. This multidimensional representation allows assessment to capture competencies that are less explicitly represented in conventional numeracy instruments. Consequently, the contribution of the proposed instrument lies not in introducing an alternative definition of numeracy, but in expanding how contextual numeracy can be represented and measured within elementary educational assessment.

These findings also respond to ongoing discussions regarding authentic assessment in elementary mathematics education. As educational systems increasingly emphasise higher-order thinking, meaningful learning, and contextual problem solving, assessment instruments are expected to provide richer evidence regarding students' learning processes rather than merely documenting achievement scores. The satisfactory psychometric performance obtained in the present study indicates that incorporating contextual learning dimensions into assessment design does not compromise measurement quality. Instead, integrating REACT provides additional diagnostic information regarding students' strengths and challenges across multiple dimensions of contextual numeracy. Such information may support teachers in identifying whether students experience difficulties in connecting prior knowledge, applying mathematical concepts, collaborating during problem solving, or transferring mathematical understanding to new situations. Therefore, the assessment has potential value not only as an evaluative instrument but also as a diagnostic tool for improving mathematics instruction.

From a theoretical perspective, the present findings contribute to the growing body of literature seeking stronger alignment between learning theory and educational assessment. Previous research has demonstrated that contextual learning environments improve mathematical understanding because they encourage students to construct knowledge actively through meaningful experiences (Faqih et al., 2022; Muliani et al., 2022; Utama et al., 2023). The present study extends this perspective by demonstrating that the same theoretical principles can also guide the development of assessment instruments capable of representing contextual learning processes. Consequently, the contribution of this study extends beyond instrument development itself. It proposes an alternative conceptual perspective in which educational assessment is viewed as an extension of learning theory rather than merely a mechanism for evaluating learning outcomes. Such integration strengthens the conceptual foundation of authentic numeracy assessment and contributes to the continuing evolution of educational assessment frameworks designed for elementary mathematics education.

Theoretical Contribution, Practical Implications, Limitations, and Future Directions

Beyond establishing the psychometric adequacy of the developed instrument, the present study offers a broader contribution to the field of educational assessment by strengthening the conceptual relationship between contextual learning theory and educational measurement. Previous studies concerning REACT have consistently demonstrated its effectiveness in improving mathematics learning outcomes, conceptual understanding, and student engagement (Faqih et al., 2022; Muliani et al., 2022; Utama et al., 2023). Nevertheless, these investigations have predominantly viewed REACT as an instructional strategy designed to facilitate classroom learning. The present study extends this perspective by demonstrating that the conceptual principles underlying REACT can also guide the development of assessment instruments capable of representing multidimensional contextual numeracy competencies. Accordingly, this study contributes to educational assessment by proposing that authentic assessment should not merely evaluate mathematical performance but should also represent the learning processes through which mathematical competence is constructed, applied, and transferred. This conceptual extension broadens the existing understanding of contextual numeracy assessment and provides an alternative perspective for future instrument development research.

Another important contribution concerns the conceptualisation of contextual numeracy itself. Conventional numeracy assessments generally conceptualise numeracy as students' ability to interpret quantitative information, apply mathematical reasoning, and solve contextual problems individually. While these competencies remain fundamental, the findings of this

study suggest that meaningful numeracy development also depends upon learners' abilities to connect mathematical concepts with prior experiences, construct understanding through authentic activities, collaborate during mathematical reasoning, and transfer acquired knowledge across different situations. By operationalising these dimensions into measurable assessment indicators, the proposed instrument expands the representation of contextual numeracy beyond procedural accuracy and contextual reasoning alone. Consequently, the developed assessment provides a more comprehensive representation of mathematical competence that aligns more closely with contemporary perspectives emphasising meaningful, authentic, and student-centered learning.

The findings likewise have important implications for classroom assessment practices in elementary education. Teachers frequently rely on assessment instruments that primarily report whether students answer mathematical questions correctly, providing limited diagnostic information regarding the learning processes underlying students' performance. In contrast, the REACT-based instrument developed in this study enables teachers to obtain richer evidence regarding different aspects of contextual numeracy competence. Because the assessment items are systematically designed around the dimensions of Relating, Experiencing, Applying, Cooperating, and Transferring, teachers are able to identify more specifically which aspects of contextual mathematical learning require additional instructional support. For example, students may demonstrate adequate procedural calculation skills while experiencing difficulty transferring mathematical knowledge to unfamiliar situations or collaborating effectively during contextual problem solving. Such diagnostic information may assist teachers in designing more targeted instructional interventions and strengthening the alignment between assessment and classroom learning activities.

The digital implementation of the assessment instrument also provides practical advantages for elementary mathematics education. Digital assessment facilitates more efficient administration, scoring, and data management while reducing teachers' administrative workload. In addition, digital platforms provide opportunities for immediate feedback and systematic monitoring of students' numeracy development across different learning dimensions. These characteristics are particularly relevant as schools increasingly integrate digital technologies into teaching and assessment practices. Therefore, the present study contributes not only to contextual numeracy assessment but also to the broader movement toward technology-enhanced assessment capable of supporting evidence-based instructional decision making. The integration of contextual learning principles with digital assessment thus represents a promising direction for improving both the quality and efficiency of elementary mathematics assessment.

Despite these contributions, several limitations should be acknowledged when

interpreting the findings. First, the field testing was conducted with elementary school students from a limited geographical area, which may restrict the generalisability of the instrument to other educational contexts with different curricular, cultural, or demographic characteristics. Second, although the instrument demonstrated satisfactory content validity and internal consistency, the psychometric evaluation focused primarily on preliminary instrument development procedures, including item validity, discrimination, difficulty, and reliability analyses. Additional validation using larger and more diverse samples would further strengthen the evidence supporting the instrument's measurement properties. Third, the present study evaluated the psychometric quality of the instrument but did not investigate its long-term impact on students' learning outcomes or teachers' assessment practices following sustained classroom implementation. Consequently, the reported findings should be interpreted within the scope of instrument development rather than instructional effectiveness.

Building upon these limitations, several directions for future research emerge. Future studies may examine the applicability of the REACT-based assessment across different grade levels, regions, and educational systems to evaluate the consistency of its psychometric properties under more diverse learning conditions. Further investigations may also employ more advanced psychometric techniques, such as exploratory or confirmatory factor analysis and item response theory, to examine the dimensional structure of the instrument and strengthen construct validation. In addition, longitudinal studies investigating how REACT-based assessment supports the development of contextual numeracy over time would provide valuable evidence regarding its educational impact. Finally, integrating adaptive digital assessment features, automated feedback, and learning analytics may further enhance the diagnostic capabilities of the instrument while supporting more personalised mathematics learning experiences for elementary school students.

Overall, the findings of this study demonstrate that contextual learning principles embodied within the REACT framework can be successfully translated into a psychometrically sound digital assessment instrument for elementary numeracy literacy. More importantly, the study illustrates that educational assessment can move beyond measuring isolated mathematical outcomes toward representing the broader cognitive, experiential, collaborative, and transferable processes through which mathematical competence develops. By integrating contextual learning theory with educational measurement, the proposed instrument contributes both practically and theoretically to the continuing development of authentic numeracy assessment. Consequently, the present study not only provides an empirically validated assessment instrument for elementary mathematics education but also offers an alternative conceptual direction for future research concerning contextualised educational assessment.

CONCLUSION

This study investigates how a numeracy literacy tool was developed based on the REACT (Relating, Experiencing, Applying, Cooperating, and Transferring) learning strategy, which is particularly created for Indonesian primary school students. In 21st-century education, being numerate and literate, such as being critical, analysing, and problem-solving, is extremely important. The tool was developed in four stages: define, design, develop and validate, and finally, publish it. Each stage checks that the tool is consistent with the national curriculum, feasible in the students' lives, and can facilitate them to improve their mathematics abilities.

Development results indicate that the REACT approach can be used to facilitate students better understanding mathematical concepts by connecting what they learn in math class to life's situations and activities. The test is comprised of difficult multiple-choice questions that assess students' overall knowledge, including the ability to critically analyse and use what they learn to apply in real life. The validity and reliability of the tool were confirmed by math and language teachers and were highly found to be reliable. The test can be conducted in basic education environments. Field tests were also carried out in some schools to acquire real data that would confirm that the instrument is valid and reliable in measuring the ability of the students.

Practically, this instrument serves not only as an assessment tool but also as a diagnostic aid for teachers to identify specific student difficulties in the 'Applying' and 'Transferring' stages of learning. For future research, it is recommended to conduct a longitudinal study to analyse whether the use of this REACT-based instrument significantly improves students' long-term retention of numeracy concepts compared to traditional static assessments. Additionally, further development could integrate adaptive algorithms to tailor the question of difficulty automatically based on student responses.

ACKNOWLEDGEMENT

We would like to thank the Institute for Research and Community Service (LP2M), State University of Malang, for supporting and facilitating funding for this KBK Internal Research activity through the UM Research Fund in 2024.

REFERENCES

Aini, V., Hidayat, T., Kusnadi, K., Williams, C., & Hadibarata, T. (2024). Analysis numeracy literacy skills of high school students in biodiversity material based on minimum competency assessment questions. *Jurnal Pendidikan IPA Indonesia, 13*(1), 128-136. https://doi.org/10.15294/jpii.v13i1.49265

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. Educational and Psychological Measurement, 45(1), 131-142. https://doi.org/10.1177/0013164485451012

Bailey, J., & Bulotsky-Shearer, R. J. (2022). Learning behaviors mediate associations between parent-reported peer play skills and literacy and mathematics skills for low-income preschool children. *Early Education and Development, 33*(3), 523-540. https://doi.org/10.1080/10409289.2021.1904625

Barham, A. I., Ihmeideh, F., Al-Falasi, M., & Alabdallah, A. (2019). Assessment of first-grade students' literacy and numeracy levels and the influence of key factors. *International Journal of Learning, Teaching and Educational Research, 18*(12), 174-195. https://doi.org/10.26803/ijlter.18.12.11

Blunch, N. H. (2014). Literacy and numeracy skills and education sector reform: Evidence from Ghana. *Education Economics, 22*(2), 209-235. https://doi.org/10.1080/09645292.2011.597954

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2021). Best practices for developing and validating scales for health, social, and behavioral research: A primer. Frontiers in Public Health, 8, Article 550299. https://doi.org/10.3389/fpubh.2020.550299

Bonifacci, P., Trambagioli, N., Bernabini, L., & Tobia, V. (2022). Home activities and cognitive skills in relation to early literacy and numeracy: Testing a multifactorial model in preschoolers. *European Journal of Psychology of Education, 37*(3), 681-705. https://doi.org/10.1007/s10212-021-00528-2

Carpenter, C. R., Kaphingst, K. A., Goodman, M. S., Lin, M. J., Melson, A. T., & Griffey, R. T. (2017). Feasibility and diagnostic accuracy of brief health literacy and numeracy screening instruments in an urban emergency department. *Hepatology, 176*(3), 139-148. https://doi.org/10.1002/hep.30150.Ductular

DeVellis, R. F. (2017). Scale development: Theory and applications (4th ed.). SAGE.

Faqih, M. I., Adriyani, Z., & Listiyani, L. (2022). Advanced chatbot development to improve student literacy and numeracy skills. *Phenomenon: Jurnal Pendidikan MIPA, 12*(2), 206-215. https://doi.org/10.21580/phen.2022.12.2.14186

Fatonah, N., Permana, J., & Syaodih, E. (2023). Improving numeracy literacy skills of elementary school students through the Kampus Mengajar program policy. *Jurnal Cakrawala Pendas, 9*(2), 298-308. https://doi.org/10.31949/jcp.v9i2.4597

Fauzan, A., Harisman, Y., Yerizon, Suherman, Tasman, F., Nisa, S., Sumarwati, Hafizatunnisa, & Syaputra, H. (2024). Realistic mathematics education (RME) to improve literacy and numeracy skills of elementary school students based on teachers' experience. *Infinity: Journal of Mathematics Education, 13*(2), 301-316. https://doi.org/10.22460/infinity.v13i2.p301-316

Fiorella, L. (2023). Making sense of generative learning. *Educational Psychology Review, 35*, Article 50. https://doi.org/10.1007/s10648-023-09769-7

Hamdu, G., Hadiana, D., Sylvia, N., Apipatunnisa, I., & Yulianto, A. (2023). Measuring changes of students' conceptual understanding of literacy and numeracy in natural science by using Rasch model. *Jurnal Ilmiah Sekolah Dasar, 7*(3), 489-497. https://doi.org/10.23887/jisd.v7i3.59943

Handayani, N. N. L. (2023). Determination of realistic mathematics education on problem solving with numeracy literacy covariables. *Jurnal Pendidikan dan Pengajaran, 56*(2), 299-309. https://doi.org/10.23887/jpp.v56i2.63288

Hidayah, I. R., Kusmayadi, T. A., & Fitriana, L. (2021a). E-book based on local wisdom to improve students' numeracy skill: Is it effective? In *Proceedings of the International Conference of Mathematics and Mathematics Education (I-CMME 2021)* (Vol. 597, pp. 142-147). Atlantis Press. https://doi.org/10.2991/assehr.k.211122.020

Hidayah, I. R., Kusmayadi, T. A., & Fitriana, L. (2021b). Minimum competency assessment (AKM): An effort to photograph numeracy. *Journal of Mathematics and Mathematics Education, 11*(1), 14. https://doi.org/10.20961/jmme.v11i1.52742

Hijriani, L., Siahaan, M. M. L., Simarmata, J. E., Missa, Y., Tampani, M., & Talan, N. (2022). Numeracy training: Applying Jarimatika method to elementary school students at Nunpene border area, East Nusa Tenggara. *Smart Society: Community Service and Empowerment Journal, 2*(1), 27-32. https://www.journal.foundae.com/index.php/smartsoc/index

Hikamudin, E., Riyadi, A. R., Aryanti, Peniasiani, D., Nuryani, P., & Gofur, R. (2023). Improving elementary school students' understanding of literacy and numeracy through digital applications. *Mimbar PGSD Undiksha, 11*(3), 462-467. https://doi.org/10.23887/jjpgsd.v11i3.64018

Humairah, H., Khasanah, L. A. I. U., & Riansyah, D. M. (2024). Analysis of the implementation program numeracy literacy for teacher working groups in building a literate culture. *Jurnal Pendidikan dan Pembelajaran Indonesia (JPPI), 4*(1), 158-165. https://doi.org/10.53299/jppi.v4i1.448

Jayaraman, J. D., Jambunathan, S., & Counselman, K. (2018). The connection between financial literacy and numeracy: A case study from India. *Numeracy, 11*(2), Article 5. https://doi.org/10.5038/1936-4660.11.2.5

Johanson, G. A., & Brooks, G. P. (2010). Initial scale development: Sample size for pilot studies. Educational and Psychological Measurement, 70(3), 394-400. https://doi.org/10.1177/0013164409355692

Kusmaharti, D., Pramulia, P., & Yustitia, V. (2023). Ethnomathematics comics of Al-Akbar Mosque Surabaya to improve numeracy and literacy in reading and writing. *International Journal of Multicultural and Multireligious Understanding, 10*(11), 66-72. https://doi.org/10.18415/ijmmu.v10i11.5118

Kustantina, V. A., Nuryadi, & Marhaeni, N. H. (2022). Improving the students' numerical literacy by using interactive mathematical comics on Pythagorean theorem. *Jurnal Ilmu Pendidikan Muhammadiyah Kramat Jati, 3*(1), 10-16.

Lestari, G. D., Widodo, & Yusuf, A. (2022). The role of parents in the development of numerical literacy in early childhood. *European Journal of Education and Pedagogy, 3*(5), 86-92. https://doi.org/10.24018/ejedu.2022.3.5.451

Lindberg, M., & Silvennoinen, H. (2018). Assessing the basic skills of the highly educated in 21 OECD countries: An international benchmark study of graduates' proficiency in literacy and numeracy using the PIAAC 2012 data. *Comparative Education, 54*(3), 325-351. https://doi.org/10.1080/03050068.2017.1403676

Lynn, M. R. (1986). Determination and quantification of content validity. Nursing Research, 35(6), 382-385. https://doi.org/10.1097/00006199-198611000-00017

Manten, A., le Roux, M., Geertsema, S., & Graham, M. (2020). An investigation into the early literacy skills of English second language learners in South Africa. *Australasian Journal of Early Childhood, 45*(2), 142-154. https://doi.org/10.1177/1836939120918504

Marthyane Pratiwi, I., Apriani, L., & Rifqi Mahmud, M. (2024). Remeasuring numeracy literacy skills: How is the students' skills post pandemic? *International Conference on Mathematics and Science Education*, April, 369-380. https://doi.org/10.18502/kss.v9i13.15937

Martini, F., & Sénéchal, M. (2012). Learning literacy skills at home: Parent teaching, expectations, and child interest. *Canadian Journal of Behavioural Science, 44*(3), 210-221. https://doi.org/10.1037/a0026758

McNaughton, C., Wallston, K. A., Rothman, R. L., Marcovitz, D. E., & Storrow, A. B. (2011). Short, subjective measures of numeracy and general health literacy in an adult emergency department. *Academic Emergency Medicine,*

18(11), 1148-1155. https://doi.org/10.1111/j.1553-2712.2011.01210.x

Mtika, P., & Abbott, P. (2023). Literacy and quality of life: A study of adults with poor literacy capabilities in Western Rwanda. *Compare, 00*(00), 1-17. https://doi.org/10.1080/03057925.2023.2254222

Muliani, M., Novita, N., Mellyzar, M., Pasaribu, A. I., & Fadli, M. R. (2022). Analysis of the characteristics of the ethnoscience-based numeracy test instrument using the Rasch model. *Jurnal Penelitian Pendidikan IPA, 8*(5), 2176-2183. https://doi.org/10.29303/jppipa.v8i5.2285

Nandyala, A. S., Nelson, L. A., Lagotte, A. E., & Osborn, C. Y. (2018). An analysis of whether health literacy and numeracy are associated with diabetes medication adherence. *Health Literacy Research and Practice, 2*(1), e15-e20. https://doi.org/10.3928/24748307-20171212-01

Rini, T. A., Cholifah, P. S., Nuraini, N. L. S., & Margetts, K. (2021). Readiness of elementary teachers in Minimum Competency Assessment: Teachers' competence in arranging literature and numeration tests. *Profesi Pendidikan Dasar, 8*(2), 156-169. https://doi.org/10.23917/ppd.v8i2.16157

Rohendi, D. (2019). Game-based multimedia for horizontal numeracy learning. *International Journal of Emerging Technologies in Learning, 14*(15), 159-170. https://doi.org/10.3991/ijet.v14i15.10679

Rohmah, A. N., Sutama, S., Hidayati, Y. M., Fauziati, E., & Rahmawati, L. E. (2022). Planning for cultivation numerical literacy in mathematics learning for Minimum Competency Assessment (AKM) in elementary schools. *Mimbar Sekolah Dasar, 9*(3), 503-516. https://doi.org/10.53400/mimbar-sd.v9i3.51774

Segers, E., Kleemans, T., & Verhoeven, L. (2015). Role of parent literacy and numeracy expectations and activities in predicting early numeracy skills. *Mathematical Thinking and Learning, 17*(2-3), 219-236. https://doi.org/10.1080/10986065.2015.1016819

Setiawan, R., Dewi, R. S. I., & Zainul, R. (2024). Development of growth mindset in higher education in support of the World-Class University. *International Journal of Evaluation and Research in Education, 13*(6), 3806-3817. https://doi.org/10.11591/ijere.v13i6.30145

Situmorang, A. S., & Sinaga, D. F. (2022). Analysis of numeracy literacy skills of students on the Pythagorean theorem material in Class VIII UPT SMP Negeri 7 Medan. *International Journal of Humanities Education and Social Sciences, 2*(1), 342-347. https://doi.org/10.55227/ijhess.v2i1.242

Sweller, J. (2022). The role of evolutionary psychology in our understanding of human cognition: Consequences for cognitive load theory and instructional procedures. *Educational Psychology Review, 34*(2), 2229-2241. https://doi.org/10.1007/s10648-021-09647-0

Utama, F., Zaenuri, Z., & Pranoto, Y. K. S. (2023). Students numerical literacy ability in problem-based learning with ethnomathematics nuances. *Prima: Jurnal Pendidikan Matematika, 7*(2), 98. https://doi.org/10.31000/prima.v7i2.8387

Via, Y., Tatag, S., & Abadi, A. (2021). The effect of mathematics self-efficacy on numeracy skills of prospective elementary school teachers. *Cypriot Journal of Educational Sciences, 16*(6), 3405-3417. https://doi.org/10.18844/cjes.v16i6.6590

Zulfayani, Z., Ariawan, R., Nufus, H., & Kafrina, K. (2023). Profile of students' numeracy literacy ability and self-efficacy in mathematics learning. *Mathematics Research and Education Journal, 7*(2), 25-35. https://doi.org/10.25299/mrej.2023.vol7(2).14381