Jurnal Elemen, 12(3), 760-777, July 2026
https:/doi.org/10.29408/jel.v12i3.34250

e- ISSN: 2442 - 4226

Development of a diagnostic assessment for measuring students' conceptual understanding of real numbers

Nur Qomariyah Nawafilah [1]*, Rizky Oktaviana Eko Putri [2]

[1] Faculty of Science and Technology, Islamic University of Lamongan, East Java, Indonesia
[2] Faculty of Education, Universiti Malaya, Malaysia

* Correspondence: nq.nawafil@unisla.ac.id

Abstract

This research was motivated by the limitations of assessment instruments capable of measuring students' initial conceptual understanding of mathematics related to real numbers, which is the basis of calculus. This study aimed to develop a valid, practical, and effective diagnostic assessment to measure students' conceptual understanding of mathematics. This research is developmental research with the Plomp model, which includes the stages of preliminary research, development or prototyping, and assessment of the product. Data were collected through interviews, observations, questionnaires, and tests and analyzed using quantitative descriptive analysis. The results of the validator assessment on the validation sheet indicate that the developed diagnostic assessment has a validity level of 91.8% (content aspect), 87.6% (construct aspect), and 92.6% (linguistic aspects). The practicality level was 70.85% (practical), as obtained from the questionnaire scores completed by the students. In terms of effectiveness, the developed assessment met the effectiveness criteria because it could identify areas of conceptual difficulty, differentiate levels of conceptual understanding, and provide meaningful information for decision-making. Its uniqueness lies in the development of a diagnostic assessment item design framework based on specific conceptual understanding indicators for real numbers in higher education and its assessment method.

Keywords: diagnostic assessment; mathematical concepts; real numbers; research and development

How to cite: Nawafilah, N. Q., & Putri, R. O. E. (2026). Development of a diagnostic assessment for measuring students' conceptual understanding of real numbers. *Jurnal Elemen, 12*(3), 760-777. https://doi.org/10. 29408/jel.v12i3.34250

Received: 3 March 2026 | Revised: 1 July 2026
Accepted: 29 July 2026 | Published: 6 August 2026

Introduction

Understanding mathematical concepts is one of the main foundations of learning mathematics, especially at the tertiary level of education. Internationally, various reports and study results show that low understanding of mathematical drafts remains an ongoing problem, resulting in difficulties in learning, particularly in continued study of calculus and mathematical analysis (Tall, 2013; Stewart et al., 2016; OECD, 2019). Students are often able to perform procedural calculations but experience difficulty in explaining the meaning of concepts, classifying mathematical objects, and connecting one draft with others (Skemp, 1978; Kilpatrick et al., 2001). These conditions show that learning mathematics in tertiary education still tends to emphasize procedural aspects compared to conceptual understanding.

Similar findings are evident at the national level. Several studies in Indonesia report that new students have a poor understanding of mathematical concepts, particularly foundational knowledge required for calculus, such as real numbers, algebraic operations, and functions (Aziz et al., 2025; Ruamba, 2025; Usman et al., 2025). Students often hold misconceptions about distinguishing between rational and irrational numbers, representing numbers on a number line, and interpreting absolute values and intervals (Hidayat & Prabawanto, 2018; Kidron, 2018; Rizos & Adam, 2022). These problems are crucial because weaknesses in the concept of real numbers directly affect understanding of limits, derivatives, and integrals at higher levels (Tall & Vinner, 1981; Zazkis & Sirotic, 2010).

The experts emphasize the importance of conceptual understanding as the core of learning mathematics meaningfully. Skemp (1978) distinguishes between instrumental and relational understanding, with the latter emphasizing the interconnectedness of concepts and the reasoning behind mathematical procedures. In line with that, Kilpatrick et al. (2001) place conceptual understanding among the main components of *mathematical proficiency*. According to Hiebert and Grouws (2007), a good conceptual understanding allows students to build flexible, adaptive, and durable knowledge. Therefore, measuring the ability to understand the draft of mathematical learning needs to be done appropriately from the beginning of learning.

One of the recommended approaches for identifying and understanding the mathematical beginning of a student is through diagnostic assessment. Diagnostic assessment aims to measure ability early, identify conceptual errors, and detect misconceptions held by participants before learning begins (Black & Wiliam, 2009; Guskey, 2018). According to Wiliam (2011), diagnostic assessment is not intended to evaluate end-of-study results but rather to provide in-depth information about readiness for learning and the structure of participants' understanding. Thus, diagnostic assessment still measures ability, but within a diagnostic framework for the beginning.

Various studies have used diagnostic assessments in mathematics learning. Some studies have developed diagnostic instruments to identify students' misconceptions in algebra, geometry, and calculus (Aristiawan, 2025; Caleon & Subramaniam, 2010; Treagust, 1988). Other studies show that diagnostic assessments are effective in mapping learning difficulties and supporting remedial instruction (Alfageh et al., 2024; Jiang et al., 2022; Pratiwi et al., 2022; Zhan et al., 2021). However, many large studies focus on student-centered school settings or on procedural measurement error, rather than specifically on students' mathematical understanding at the college level.

In the mathematics context of real numbers, previous research generally highlights students' misconceptions related to irrational numbers, decimal representations, and number lines, such as studies conducted by Zazkis (2005) as well as Rizos and Adam (2022). However, these studies primarily focused on identifying specific types of errors or misconceptions and have not yet produced diagnostic instruments that comprehensively measure students' mathematical conceptual understanding of real numbers. Some diagnostic instruments that have been developed also tend to emphasize right or wrong answers and procedural errors, so they have not been explicitly designed based on indicators of mathematical conceptual understanding, such as the ability to restate concepts, identify examples and non-examples, classify objects based on conceptual properties, and connect mathematical concepts (Hiebert & Lefevre, 1986; NCTM, 2014). Consequently, the information obtained from these instruments is still limited in revealing the structure of students' conceptual understanding in depth. Recent studies also emphasize the need for assessment instruments that are theoretically grounded and empirically validated to accurately capture students' conceptual understanding and misconceptions (Berger et al., 2019; Sahin et al., 2024).

These limitations indicate a research gap, namely the lack of a diagnostic assessment specifically developed based on a theoretical framework for mathematical conceptual understanding of real numbers at the college level. Theoretically, mathematical conceptual understanding is viewed as the ability to construct meaning, represent ideas, and link various concepts into a coherent knowledge structure (Hiebert et al., 1997; Kilpatrick et al., 2001). Therefore, this study offers a contribution through the development of a diagnostic assessment that integrates indicators of mathematical conceptual understanding into the item design. The developed instrument aims not only to identify correct or incorrect answers but also to identify areas of conceptual difficulty experienced by students and reveal their level of conceptual understanding. Thus, the novelty of this study lies in the development of a diagnostic assessment item design framework based on indicators of mathematical conceptual understanding specifically designed for real numbers in college and in its assessment method so that the developed assessment can identify areas of conceptual difficulty experienced by students and reveal their level of conceptual understanding.

Given the exposure, this research focused on developing and assessing a diagnostic measure to determine beginning students' initial level of mathematical conceptual understanding of real numbers. This research aimed to produce a valid, practical, and effective diagnostic assessment capable of revealing students' initial level of mathematical conceptual understanding. With the existence of this diagnostic assessment, it is hoped that lecturers will obtain accurate information about students' readiness to study calculus, so that calculus learning can be designed in a way that is more effective and appropriate for students' needs.

Methods

Research design

This study is a development study (*research and development*) that aimed to produce a diagnostic assessment product to measure the mathematical understanding of beginning students on the topic of real numbers. The development model used was the Plomp model, which consists of three main stages, namely *preliminary research*, *development or prototyping*, and the *assessment phase* (Plomp, 1997). This model was chosen because it provides a systematic framework for developing educational products. It allowed researchers to analyze needs, develop the diagnostic assessment instrument gradually, and conduct evaluation and revision based on validation and testing results, so that the diagnostic assessment instrument generated has adequate quality in terms of validity, practicality, and effectiveness.

Research procedures

Research procedures implemented through a three-stage main process in accordance with the Plomp model. At the *preliminary research* stage, a needs analysis was carried out through class observations and semi-structured interviews with two Calculus I lecturers and ten first-semester students. Interviews explored students' difficulties in understanding real numbers and lecturers' experiences in assessing students' prior knowledge. Research findings show that: (1) students experience difficulty in distinguishing rational and irrational numbers and in classifying subsets of real numbers; (2) assessments used in previous classes focused only on procedural skills and did not provide information about students' conceptual understanding; and (3) lecturers reported difficulty in identifying student misconceptions before introducing advanced calculus topics. These findings emphasize the need for diagnostic assessments specifically designed to measure students' conceptual understanding of real numbers.

At the *development or prototyping* stage, researchers compile an instrument grid based on indicators of ability to understand mathematical concepts, design diagnostic assessment items on real numbers, and prepare research-supporting instruments. Instrument development begins with compiling an assessment blueprint that maps each indicator of conceptual understanding against certain learning objectives and the concept of real numbers. Eight items were then developed, with two items representing each indicator of conceptual understanding: (1) restating a concept, (2) identifying examples and non-examples, (3) classifying objects based on conceptual properties, and (4) connecting mathematical concepts. Each item is designed to elicit students' conceptual reasoning rather than procedural calculations. The blueprint forms the basis for ensuring the representativeness of the content and alignment between the theoretical framework and the assessment items developed. The instruments developed were then validated by experts to evaluate content, construction, and linguistic aspects before being revised and used in the testing phase.

Table 1. Blueprint of the diagnostic assessment items

Indicator	Real Number Topic	Item Number
Restating a concept	Definition and characteristics of real numbers	1, 2
Identifying examples/non-examples	rational and irrational numbers, undefined expressions	3, 4
Classifying objects	Classification of real numbers into rational and irrational numbers	5, 6
Connecting concepts	Relationship between real numbers and the number line	7, 8

The *assessment phase* was carried out through limited trials and field trials. A limited trial was conducted to determine the practicality of the diagnostic assessment, while a field trial was conducted to assess the effectiveness of the diagnostic assessment in measuring students' ability to understand mathematical concepts related to real numbers.

Research subject

The subjects of this study were 30 first-semester students enrolled in Calculus I at the Faculty of Science and Technology, Universitas Islam Lamongan, East Java, Indonesia. The subjects were selected purposively to align with the research objectives: students who would be studying real numbers as a foundational concept in calculus. The field trial involved 30 students, which is considered adequate for preliminary testing and formative evaluation in educational design research aimed at obtaining initial evidence regarding practicality and diagnostic effectiveness rather than establishing definitive psychometric properties (Tessmer, 1993; Plomp & Nieveen, 2013).

Research instrument

1. Diagnostic test for understanding mathematical concepts. The instrument was developed based on four indicators of understanding mathematical concepts, namely restating a concept, identifying examples and non-examples, classifying objects by conceptual properties, and connecting one concept to another.
2. Expert validation sheet. Validation was carried out by three validators who have more than five years of experience in their respective fields. The content expert holds a doctorate in mathematics and has real analysis skills; the construct expert is a senior mathematics education lecturer who has research experience in the field of assessment and instrument development; and the linguist is an Indonesian language education lecturer with academic writing expertise. The experts independently evaluated the instrument and subsequently discussed any discrepancies in their judgments to reach consensus before the final revision of the instrument was conducted.
3. Practicality questionnaire. A practicality questionnaire was used to collect data on the ease of use of the diagnostic assessment, based on student responses. This questionnaire was administered to students after a limited pilot test of the diagnostic assessment.

Data analysis techniques

Data analysis in the study was conducted using qualitative and quantitative methods in accordance with the type of data obtained from each instrument in the study.

1. Validation Sheet Data Analysis

   Data from instrument validation results obtained from expert validation sheets were analyzed quantitatively using the average score for each aspect of the assessment. The scores resulting from the validator assessment were then converted into validity values using the following validity index formula:

   $$V = \frac{\sum r}{n \times c} \quad (1)$$

   Information:

   V= value instrument validity

   r= score given by the validator to each grain

   n= amount grains statement

   c= maximum score on the rating scale

   Calculation results were used to assess validity, then interpreted against the criteria for validity (Table 2). In addition, qualitative data in the form of suggestions and input from the validator were analyzed descriptively and used as a basis for revising the instruments.

Table 2. Criteria validity (Nurhandayani et al., 2022)

Interval (%)	Criteria
80 < V ≤ 100	Very Valid
60 < V ≤ 80	Valid
40 < V ≤ 60	Quite Valid
20 < V ≤ 40	Less Valid
0 < V ≤ 20	Very Less Valid

2. Practicality questionnaire data analysis

   Questionnaire data on practicality, gathered from student responses, were analyzed quantitatively using counts and percentages, with the formula for practicality use as follows:

   $$P = \frac{Total\ scores\ obtained\ from\ the\ practicality\ questionnaire}{The\ maximum\ possible\ score\ that\ can\ be\ obtained} \quad (2)$$

   Percentage value practicality gained, further categorized by practicality criteria (Table 3), for a practicality assessment level and developed diagnostics.

Table 3. Practicality Criteria

Interval (%)	Criteria
80 < P ≤ 100	Very Practical
60 < P ≤ 80	Practical
40 < P ≤ 60	Enough Practical
20 < P ≤ 40	Less practical
0 < P ≤ 20	Very Impractical

3. Diagnostic test data analysis
   The diagnostic test data were analyzed quantitatively to examine the effectiveness of the developed diagnostic assessment. First, the percentage of correct responses for each item and indicator was calculated to describe students' levels of conceptual understanding of real numbers and to identify areas that may indicate conceptual difficulties. Second, the discrimination index of each item was computed to determine the extent to which the instrument could differentiate between students with high and low levels of conceptual understanding. In this study, the diagnostic assessment was considered effective if it (1) could identify areas of conceptual difficulty; (2) could differentiate students' levels of conceptual understanding, as indicated by acceptable item discrimination indices; and (3) could provide meaningful information for instructional decision-making.

   The following is the formula for calculating the discrimination index.

$$D = P_U - P_L \tag{3}$$

   where:
   P_U = proportion of students in the upper group who answered correctly
   P_L = proportion of students in the lower group who answered correctly

Table 4. Interpretation of discrimination index (Ebel & Frisbie, 1991)

D	Category
≥ 0.40	Very good
0.30 – 0.39	Good
0.20 – 0.29	Fair
< 0.20	Poor

The present study focused primarily on content validity, practicality, and preliminary diagnostic effectiveness. Additional psychometric analyses, including item difficulty, internal consistency, and construct validity assessed via exploratory or confirmatory factor analysis, were beyond the scope of this study and should be addressed in future research with larger samples.

Results

The results of the research on the development of diagnostic assessments to measure students' understanding of real numbers are presented according to the stages of the Plomp development model, which include: Preliminary Research, Development or Prototyping, and Assessment Phase.

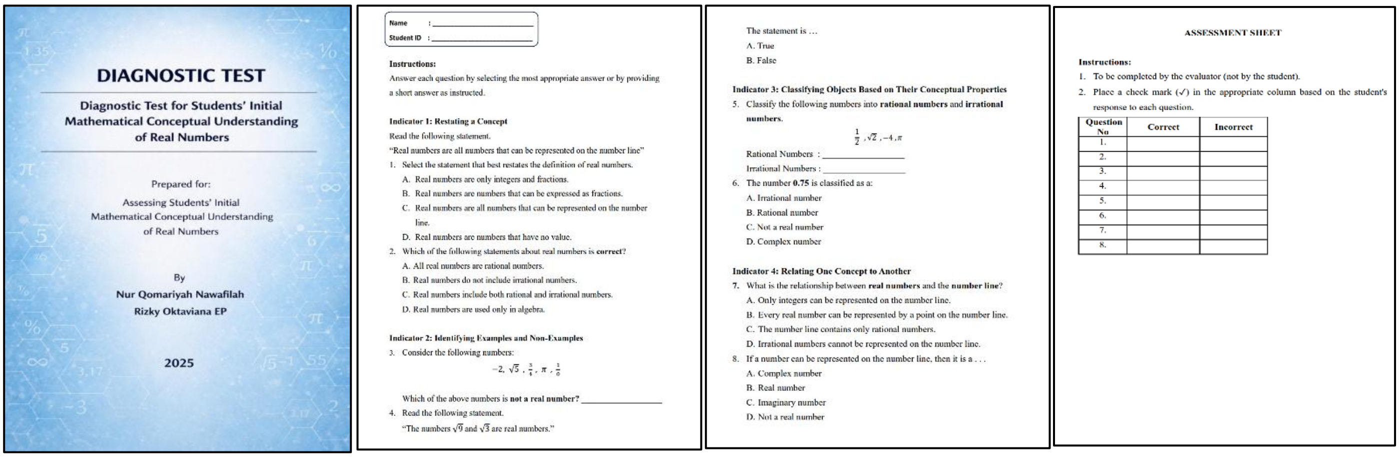
Preliminary research

At this stage, the analysis was conducted through observation and interviews with the lecturer, and students studying Calculus I and 1st-semester students. The results are: 1) First-semester students still experience difficulty understanding concepts based on real numbers, such as properties of real numbers, grouping of numbers, and relationships between real numbers and basic algebra, 2) Assessment during this period placed more emphasis on procedural and calculation abilities, so the ability to reveal students' mathematical understanding was not yet evident early on, 3) Although first-semester students have learned about real numbers at the high school level, this material was studied primarily in the context of practice questions and calculations, so basic concepts such as the classification of real numbers and their relationship to algebraic concepts have not been understood in depth. This indicated the need for the development of assessment diagnostics that are specially designed to measure students' understanding of draft mathematics before calculus is implemented.

Development or prototyping

Stage *development or prototyping* is a diagnostic assessment of stage design and development, based on results from *preliminary research*. At this stage, the researcher compiled a diagnostic assessment grid that refers to four indicators of mathematical understanding: (1) stating a concept, (2) identifying examples and non-examples, (3) classifying objects based on conceptual characteristics, and (4) connecting one draft with another draft.

Based on the grid, researchers developed a prototype diagnostic assessment in the form of a set of diagnostic questions on the material of real numbers. The questions were designed to reveal understanding among beginning students, with complexity levels customized for first-semester students.



(a) (b) (c) (d)

Figure 1. (a) Cover of diagnostic test; (b) Questions of Diagnostic test for indicators 1 and 2; (c) Questions of Diagnostic test for indicators 3 and 4; (d) Assessment sheet

Table 5. Example of items

Indicator	Sample Item	Expected Answer	Concept Assessed
Identifying examples and non-examples	From the following numbers: −2, √5, 3/4, π, 1/0, which is not a real number?	1/0	Understanding of real numbers and undefined expressions
Classifying objects	Classify 1/2, √2, −4, and π into rational and irrational numbers.	Rational: 1/2, −4, π; Irrational: √2	Classification of real numbers

Table 6. Scoring rubric for the diagnostic assessment

Indicator of Conceptual Understanding	Item Number	Score 1	Score 0
Restating a concept	1, 2	The student correctly restates the definition or properties of real numbers.	The student provides an incorrect response or demonstrates misconceptions about real numbers.
Identifying examples and non-examples	3, 4	The student correctly identifies examples and non-examples of real numbers.	The student incorrectly identifies examples and non-examples or demonstrates misconceptions.
Classifying objects based on conceptual properties	5, 6	The student correctly classifies numbers according to their properties.	The student incorrectly classifies numbers, indicating conceptual misunderstanding.
Connecting mathematical concepts	7, 8	The student correctly identifies relationships between real numbers and the number line.	The student fails to establish the relationships or demonstrates misconceptions.

Once the design assessment diagnostic was completed, the researcher prepared supporting instruments, such as validation sheets to assess the diagnostic's validity and practicality questionnaires to evaluate its practicality.

Assessment phase

Table 7. Validation results of mathematics expert (aspects content)

No	Component	%
1	Compliance material with draft real numbers	94
2	Compliance grains question with indicator understanding draft mathematical	91
3	Accuracy draft mathematically on every grain question	93
4	Representation indicator understanding draft mathematical	89
5	Compliance level difficulty question with characteristics student	92
	Average	**91.8**

As shown in Table 7, the validation results from the mathematics content expert indicate an average score of 91.8%, which falls into the very valid category. The highest score was obtained for the suitability of the material with the real number's topic (94%), while the lowest score was found in the representation of mathematical conceptual understanding indicators (89%). Overall, these results indicate that the content of the diagnostic assessment is highly relevant to the learning objectives and adequately represents the indicators of students' mathematical conceptual understanding.

Table 8. Validation results of mathematics education expert (aspect construct)

No	Component Evaluation Aspect Construct	%
1	Clarity construction question in represent understanding draft mathematical	86
2	Accuracy form question as instrument diagnostic assessment	85
3	Integration grains question in measure the same construct	89
4	Ability question reveals understanding draft mathematical beginning student	91
5	Consistency structure question with characteristics first semester students	87
	Average	**87.6**

Table 8 presents the validation results from the mathematics education expert regarding the construct aspect. The average validation score was 87.6%, indicating that the instrument is very valid in terms of its construct. The highest score was obtained for the instrument's ability to reveal students' initial mathematical conceptual understanding (91%), whereas the lowest score was related to the accuracy of the question format as a diagnostic assessment instrument (85%). These findings suggest that the assessment items have an appropriate structure and are capable of measuring the intended construct.

Table 9. Validation results of language expert (aspect linguistic)

No	Component	%
1	Compliance language used with correct Indonesian language rules	93
2	Correctness of use symbols and notations mathematics	95
3	Use simple and easy language understood student	90
4	Use communicative language	91
5	Clarity instructions and directions workmanship question	94
	Average	**92.6**

As presented in Table 9, the language expert awarded an average validation score of 92.6%, which is categorized as very valid. The highest score was obtained for the correctness of mathematical symbols and notation (95%), while the lowest score was related to the use of simple and easily understood language (90%). These results indicate that the assessment instrument uses clear, communicative, and linguistically appropriate language, making it suitable for first-semester university students.

Table 10. Recapitulation average validation results expert (content, construct, language)

No	Expert Validation	%	Criteria Interpretation
1	Content Aspect	91.8	Very Valid
2	Aspect Construct	87.6	Very Valid
3	Language Aspects	92.6	Very Valid

Table 10 summarizes the overall expert validation results. The content, construct, and language aspects obtained average scores of 91.8%, 87.6%, and 92.6%, respectively. All three aspects were classified as very valid, indicating that the developed diagnostic assessment instrument satisfies the required quality standards in terms of content relevance, construct appropriateness, and language clarity. Therefore, the instrument is considered feasible for implementation in measuring students' initial mathematical conceptual understanding of real numbers.

Furthermore, a trial was conducted to evaluate the product's practicality, using stages *of one-on-one* and *small-group settings*. *One-on-one* testing used 3 students as the sample, with the objective of identifying possible difficulties that may arise during product use. Meanwhile, *small-group* testing was conducted with 12 students. Product test results can be seen in Table 11, which indicates that the product meets the practicality criteria.

Table 11. One-to-one and small group results

No	Trials	%	Criteria
1	One to One	66.78	Practical
2	Small Group	74.92	Practical
	Average	**70.85**	**Practical**

The next stage involved a field trial with 30 first-year students enrolled in the Calculus I course at the Faculty of Science and Technology, Universitas Islam Lamongan. The field trial was conducted to examine the diagnostic effectiveness of the developed assessment by describing the levels of conceptual understanding of real numbers and identifying students' conceptual difficulties. The results are presented in Table 12.

Table 12. Percentage students who answer correct

Indicator	Question Number	Students who answered Correct		% per indicator
		Total	%	
Restating a Concept	1	26	86.67	75.01
	2	19	63.34	
Identifying Examples and Non-Examples	3	19	63.34	73.34
	4	25	83.34	
Classifying Objects Based on Conceptual Properties	5	20	66.67	71.67
	6	23	76.67	
Connecting One Concept with Another Concept	7	20	66.67	73.34
	8	24	80.00	
% average of all indicator				**73.34**

As shown in Table 12, the average percentage of correct responses across all indicators was 73.34%. However, this percentage was not interpreted as evidence of high achievement but rather as diagnostic information describing about students' level of conceptual understanding of real numbers.

The results also revealed variations in students' performance across the four indicators. The highest percentage was observed for the restating a concept indicator (75.01%), indicating that most students were able to recall and restate the definitions and properties of real numbers. In contrast, the lowest percentage was observed for the classifying objects based on conceptual properties indicator (71.67%), suggesting that students had greater difficulty distinguishing and classifying real numbers according to their defining properties.

These findings indicate that the developed assessment identified areas of conceptual difficulty, particularly in classifying objects by conceptual properties. Therefore, the results provide valuable information for lecturers to determine which concepts require further explanation, reinforcement, or remedial instruction before students proceed to more advanced topics in calculus.

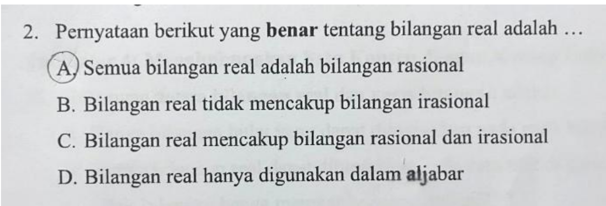
Table 13. Discrimination index of diagnostic test items

Question Number	Indicator	Discrimination Index (D)	Category
1	Restating a Concept	0.50	Very good
2	Restating a Concept	0.625	Very Good
3	Identifying Examples and Non-Examples	0.75	Very good
4	Identifying Examples and Non-Examples	0.625	Very Good
5	Classifying Objects Based on Conceptual Properties	0.625	Very Good
6	Classifying Objects Based on Conceptual Properties	0.75	Very good
7	Connecting One Concept with Another Concept	0.25	Fair
8	Connecting One Concept with Another Concept	0.75	Very good

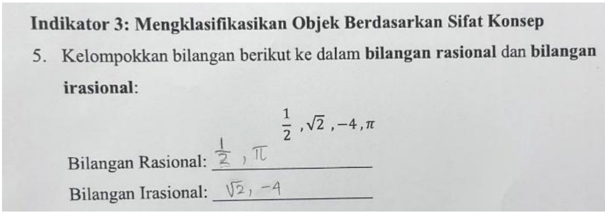
As shown in Table 13, most items demonstrated acceptable discrimination indices ($D \geq 0.30$), indicating that the items differentiated between students with high and low levels of conceptual understanding. This finding suggests that the developed assessment possesses adequate diagnostic capability because it can differentiate students according to their conceptual understanding of real numbers.

Taken together, the percentages of correct responses and the discrimination indices indicate that the developed diagnostic assessment (1) identified areas of conceptual difficulty, (2) differentiated students' levels of conceptual understanding, and (3) provided meaningful information to support instructional planning. Therefore, the developed assessment can be considered an effective diagnostic instrument for measuring students' conceptual understanding of real numbers.

Here are some examples of students' responses and their interpretations:


2. Pernyataan berikut yang **benar** tentang bilangan real adalah …
A. Semua bilangan real adalah bilangan rasional
B. Bilangan real tidak mencakup bilangan irasional
C. Bilangan real mencakup bilangan rasional dan irasional
D. Bilangan real hanya digunakan dalam aljabar

This response indicates that the student does not understand the hierarchical structure of the real number, therefore unable to restate a given concept.



Indikator 3: Mengklasifikasikan Objek Berdasarkan Sifat Konsep

5. Kelompokkan bilangan berikut ke dalam **bilangan rasional** dan **bilangan irasional**:

$\frac{1}{2}, \sqrt{2}, -4, \pi$

Bilangan Rasional: $\frac{1}{2}$, π

Bilangan Irasional: $\sqrt{2}$, -4

This response demonstrates the misconception that π is a rational number because it is often approximated by $\frac{22}{7}$. The student fails to distinguish between an exact representation and an approximation of a number. Furthermore, the student does not recognize that -4 is a rational number even though it can be expressed in the form $\frac{a}{b}$.

Discussion

This study aimed to develop assessment instruments that were valid, practical, and effective for measuring beginning students' understanding of the real numbers. Research results show that the developed instruments fulfilled the third criterion stated, providing a direct answer to the research question formulated in the introduction. The instrument was not only theoretically sound but also applicable in real learning contexts to identify the structure of beginning students' understanding.

The very high levels of validity across content, construct, and language indicate that the development instrument is in harmony with the theoretical understanding of concepts and representations of meaning (Hiebert & Lefevre, 1986; Kilpatrick et al., 2001). In particular, construct validity indicated that the items truly measure conceptual understanding, not merely procedural ability. This aligns with Wiliam's view (2011) that diagnostic assessment must be designed to reveal the structure of participants' knowledge. These findings are consistent with recent research by Parsinia and Rott (2026), who reported that a well-validated assessment of real-number knowledge should provide evidence of content validity, response processes, internal structure, and diagnostic precision. Their study likewise emphasizes that assessment instruments should move beyond measuring procedural performance to identifying students' conceptual strengths and weaknesses

From a practical standpoint, trial results indicated that the instrument was easy for students to understand and use. Revision, gradually moving from the one-to-one stage to the small group, shows improvement in instrument quality, supporting Plomp's (1997) development models, which emphasize the importance of formative evaluation at every stage of the prototype. These findings were consistent with Treagust's (1988) study, which emphasized that diagnostic instruments must be operational and responsive to users.

The effectiveness of the developed diagnostic assessment should not be interpreted solely in terms of the average percentage of correct responses. Rather, its effectiveness lies in its diagnostic capability to reveal variations in students' conceptual understanding, identify areas of conceptual difficulty, and provide information that can support instructional decision-making. This interpretation is consistent with recent studies emphasizing that diagnostic assessments should generate meaningful information about learners' conceptual understanding to support instructional planning rather than merely reporting achievement scores (Parsinia & Rott, 2026).

The findings showed that students' conceptual understanding of real numbers varied across the four indicators. In particular, the indicator assessing classification by conceptual properties yielded the lowest percentage of correct responses (71.67%). This suggests that students struggled to grasp the hierarchical relationships among subsets of real numbers and to identify the defining properties of each number set. Classifying real numbers requires relational understanding because students must coordinate several interconnected concepts rather than merely recall isolated definitions. Therefore, the lower performance on this indicator indicates that students' understanding of the real number system remains fragmented and has not yet developed into a coherent conceptual structure. This finding is consistent with Skemp's (1978) distinction between instrumental and relational understanding and supports Tall's (2013) argument that students often possess concept images that are not fully integrated. From an instructional perspective, these results highlight the need for learning activities that emphasize conceptual connections and classification tasks to strengthen students' understanding of the real number system.

Furthermore, the discrimination index analysis demonstrated that most items had good to very good discrimination power, with only one item showing fair discrimination. This finding indicates that the developed assessment was able to distinguish between students with high and low levels of conceptual understanding. Therefore, the instrument possesses adequate diagnostic capability and can provide meaningful information regarding students' conceptual strengths and weaknesses. This finding is consistent with recent psychometric studies indicating that discrimination indices are essential for evaluating the ability of assessment items to differentiate learners with different levels of ability and to ensure the diagnostic quality of educational assessments (Metsämuuronen, 2026; Ekawati et al., 2025).

From a theoretical perspective, these findings support the distinction proposed by Skemp (1978) between instrumental and relational understanding. Students tended to perform better on indicators involving the recall and restatement of concepts than on indicators requiring the classification and connection of concepts, suggesting that relational understanding is more difficult to achieve than instrumental understanding. This finding is also consistent with Tall (2013), who argued that students' conceptual understanding of mathematical ideas is often fragmented and not yet fully integrated.

Previous studies have developed diagnostic assessments to identify students' misconceptions and procedural errors in mathematics (Herliana et al., 2024; Zazkis, 2005). Rather than proposing a new approach to diagnostic assessment, the present study extends existing work by developing an instrument to diagnose first-year university students' conceptual understanding of real numbers, a foundational topic for learning calculus. The instrument operationalizes conceptual understanding through four indicators, namely restating concepts, identifying examples and non-examples, classifying objects by conceptual properties, and connecting mathematical concepts. Consequently, the instrument provides a more detailed profile of students' conceptual strengths and difficulties, which can inform instructional planning and remedial interventions.

From a practical perspective, the developed assessment can serve as an initial diagnostic tool for lecturers to identify students' conceptual difficulties in real numbers before introducing more advanced calculus topics. The diagnostic information generated by the instrument may facilitate more adaptive and evidence-based instructional planning. This implication is consistent with the principles of formative assessment, which emphasize the use of assessment evidence to adapt instruction to students' learning needs (Black & Wiliam, 2009). Recent studies further suggest that diagnostic and formative assessment can provide detailed evidence of students' conceptual understanding, enabling lecturers to design more targeted instructional interventions and improve conceptual learning (Fatih, 2026; Parsinia & Rott, 2026). Nevertheless, this study was conducted in a single institution and did not examine the impact of using the diagnostic information on subsequent learning outcomes. Future research should therefore investigate the implementation of the instrument in broader contexts and examine its contribution to improving students' learning outcomes in calculus and other mathematical domains.

Conclusion

This study developed a diagnostic assessment designed to measure first-year college students' conceptual understanding of real numbers, a foundational topic for learning calculus. Results indicate that the instrument exhibits acceptable levels of validity, practicality, and diagnostic effectiveness in the context of this study. More importantly, the assessment was able to identify areas of conceptual difficulty and differentiate students' levels of conceptual understanding across several indicators. Therefore, this instrument had the potential to provide instructors with meaningful information about students' initial understanding, misconceptions, and conceptual difficulties before more advanced calculus topics were introduced.

The diagnostic information generated by this instrument could support more adaptive and evidence-based instructional planning, particularly in designing learning activities and strengthening conceptual connections among real number concepts that underlie the learning of limits, derivatives, and integrals. However, these findings should be interpreted with caution because the instrument was evaluated using a relatively small sample from a single institution. Consequently, further psychometric evaluation involving a larger and more diverse population is needed to examine the instrument's stability, generalizability, and broader applicability before it can be recommended for wider implementation.

Acknowledgment

The authors would like to thank the Faculty of Science and Technology at Lamongan Islamic University for their support in conducting this research. Appreciation is also extended to the lecturers teaching Calculus I and to the students who participated in the instrument testing.

Declarations

Conflicts of Interest	:	The authors declare no conflict of interest.
Generative AI Statement	:	Generative AI tools, Grammarly were employed solely for language editing and minor phrasing enhancements. All conceptualization, analysis, and scholarly content were independently developed and verified by the authors.
Funding Statement	:	This work received no specific grant from any public, commercial, or not-for-profit funding agency.
Author Contributions	:	**Nur Qomariyah Nawafilah:** Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Writing – Original Draft, Writing – Review & Editing. **Rizky Oktaviana Eko Putri:** Supervision, Validation, Writing – Review & Editing.

References

Alfageh, D. H., York, C. S., Hodge-Zickerman, A., & Xie, Y. (2024). Elementary teachers' use of adaptive diagnostic assessment to improve mathematics teaching and learning: A case study. *International Electronic Journal of Mathematics Education, 19*(1), Article em0768. https://doi.org/10.29333/iejme/14190

Aristiawan, A., Wilujeng, H., Erwanto, M. R. Y., & Murtafiah, W. (2025). Development of a four-tier diagnostic test to detect middle school students' misconceptions in algebraic thinking. *Prima: Jurnal Pendidikan Matematika, 9*(2), 256–268. https://doi.org/10.31000/prima.v9i2.12700

Aziz, I. R., Harahap, N. S. R., & Siregar, R. N. (2025). Analysis of students' difficulties on understanding integral concepts. *Journal of Innovative Mathematics Learning, 8*(4), 891–904. https://doi.org/10.22460/jiml.v8i4.30135

Berger, S., Verschoor, A. J., Eggen, T. J. H. M., & Moser, U. (2019). Development and Validation of a Vertical Scale for Formative Assessment in Mathematics. *Frontiers in Education, 4*, 103. https://doi.org/10.3389/feduc.2019.00103

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5.

Caleon, I. S., & Subramaniam, R. (2010). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education, 32*(7), 939–961. https://doi.org/10.1080/09500690902890130

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.

Ekawati, A., Siswono, T. Y. E., & Lukito, A. (2025). Psychometric analysis of high school mathematics ability test. *Journal of Education Culture and Society, 16*(1), 503–519. https://doi.org/10.15503/jecs2025.2.503.519

Fatih, J. G. (2026). Enhancing mathematical conceptual understanding through formative assessment with conceptual scaffolding. *Journal of Applied Mathematics and Data Science, 1*(1), 38–49. https://doi.org/10.64845/jamds.v1i1.291

Guskey, T. R. (2018). *On your mark: Challenging the conventions of grading and reporting*. Solution Tree Press.

Herliana, H., Maison, M., & Syaiful, S. (2024). Development and implementation of a five-tier diagnostic test to identify student misconceptions on fractions: A significant step towards improving mathematics education. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi, 8*(2), 563–576. https://doi.org/10.22437/jiituj.v8i2.34159

Hidayat, W., & Prabawanto, S. (2018). Students' difficulties in understanding irrational numbers. *Journal on Mathematics Education, 9*(1), 23–34. https://doi.org/10.22342/jme.9.1.5072.23-34

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Information Age Publishing.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), Conceptual and procedural knowledge: The case of mathematics (pp. 1–27). Lawrence Erlbaum Associates.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., Olivier, A., & Wearne, D. (1997). *Making sense: Teaching and learning mathematics with understanding*. Heinemann.

Jiang, B., Li, X., Yang, S., Kong, Y., Cheng, W., Hao, C., & Lin, Q. (2022). Data-driven personalized learning path planning based on cognitive diagnostic assessments in MOOCs. *Applied Sciences, 12*(8), Article 3982. https://doi.org/10.3390/app12083982

Kidron, I. (2018). Students' conceptions of irrational numbers. *International Journal of Research in Undergraduate Mathematics Education, 4*(1), 94–118. https://doi.org/10.1007/s40753-018-0071-z

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academies Press.

Metsämuuronen, J. (2026). Dimension-corrected discrimination index. *Behaviormetrika, 53*(1), 163–203. https://doi.org/10.1007/s41237-025-00279-0

NCTM. (2014). *Principles to actions: Ensuring mathematical success for all*. National Council of Teachers of Mathematics.

Nurhandayani, E. F., Mulyono, D., & Yanto, Y. (2022). Pengembangan e-modul matematika materi barisan dan deret dengan pendekatan problem-based learning (PBL) kelas XI SMA [Development of a mathematics e-module on sequences and series using a problem-based learning (PBL) approach for Grade 11 Senior High School Students]. *Jurnal Pendidikan Matematika (Judika Education), 5*(2), 126–137. https://doi.org/10.31539/judika.v5i2.4588

OECD. (2019). *PISA 2018 results (Volume I): What students know and can do*. OECD Publishing. https://doi.org/10.1080/09500690902890130

Parsinia, M., & Rott, B. (2026). Constructing and validating a test to assess pre-service teachers' knowledge of real numbers. *Frontiers in Education, 11*, 1776551. https://doi.org/10.3389/feduc.2026.1776551

Plomp, T. (1997). Educational design: Introduction. In T. Plomp & N. Nieveen (Eds.), *Educational design research* (pp. 9–20). Netherlands Institute for Curriculum Development (SLO).

Plomp, T., & Nieveen, N. (Eds.). (2013). *Educational design research: Part A: An Introduction*. Enschede, The Netherlands: SLO.

Pratiwi, D., Suryadi, D., & Sabandar, J. (2022). The role of diagnostic assessment in improving students' mathematical understanding. *Journal of Physics: Conference Series, 1987*(1), 012034. https://doi.org/10.1080/09500690902890130

Rizos, I., & Adam, M. (2022). Mathematics students' conceptions and reactions to questions concerning the nature of rational and irrational numbers. *International Electronic Journal of Mathematics Education, 17*(3), Article em0686. https://doi.org/10.29333/iejme/11977

Ruamba, M. Y. (2025). Persistent hurdles: A systematic review of limit concept misconceptions in undergraduate calculus. *Plusminus: Jurnal Pendidikan Matematika, 5*(3, 495-508 https://doi.org/10.31980/plusminus.v5i3.3291

Sahin, M., Ozdemir, G., & Yilmaz, R. M. (2024). *Mapping design stages and methodologies for developing STEM concept inventories: A scoping review*. Frontiers in Education, 9. https://doi.org/10.3389/feduc.2024.1442833

Skemp, R. R. (1978). Relational understanding and instrumental understanding. *Arithmetic Teacher, 26*(3), 9–15.

Stewart, J., Redlin, L., & Watson, S. (2016). *Calculus: Early transcendentals* (8th ed.). Cengage Learning.

Tall, D. (2013). *How humans learn to think mathematically: Exploring the three worlds of mathematics*. Cambridge University Press.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics. *Educational Studies in Mathematics, 12*(2), 151–169. https://doi.org/10.1007/BF00305619

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education, 10*(2), 159–169. https://doi.org/10.1080/0950069880100204

Tessmer, M. (1993). *Planning and conducting formative evaluations*. London: Kogan Page.

Usman, Aiyub, & Hasbi. (2025). Students' conceptual understanding of limit of functions reviewed from mathematical beliefs. *Jurnal Elemen, 11*(3), 515-532. https://doi.org/10.29408/jel.v11i3.29144

Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press.

Zhan, P., Li, F., & Jiao, H. (2021). Editorial: Cognitive diagnostic assessment for learning. *Frontiers in Psychology, 12*, 806636. https://doi.org/10.3389/fpsyg.2021.806636

Zazkis, R. (2005). Representations of irrational numbers: A case of conceptual understanding. *Journal for Research in Mathematics Education, 36*(5), 406–426. https://doi.org/10.1080/00207390412331316951

Zazkis, R., & Sirotic, N. (2010). Representing and defining irrational numbers: Exposing the missing link. *CBMS Issues in Mathematics Education, 16*, 1–27.