Design and Formative Validation of a Guided Discovery Learning Model Integrated with Motivation to Reasoning and Proving Tasks

Selvy Musdalifah[1*], I Wayan Sudarsana[1], Sukayasa[2], Dasa Ismaimuza[2], Sutji Rochaminah[2], Nurhayadi[2]

[1] Tadulako University, MIPA (Mathematics), Palu, Indonesia.
[2] Tadulako University, FKIP (Mathematics Education), Palu, Indonesia.

Abstract: This study develops and evaluates an instructional model that integrates Guided Discovery Learning with Mathematical Reasoning and Proving (MRP) Tasks to support undergraduate students' construction of mathematical proofs in Real Analysis. The research focuses on two foundational criteria in instructional development, namely validity and practicality. Expert review involving five specialists produced high Aiken V and ICC values, indicating strong conceptual coherence, structural alignment, and reliability of judgment across components of the model. Practicality was examined through one to one evaluation, small group testing, and a field implementation involving lecturers and students. Across these stages, the model received high ratings for clarity of instructional flow, readability of tasks, feasibility of classroom enactment, and support for structured reasoning processes. The findings demonstrate that the model is both theoretically sound and operationally feasible, providing a coherent pedagogical trajectory that guides learners from exploration toward formal proof construction. This study contributes a validated instructional model and establishes a foundation for future research on its effectiveness and broader applicability in advanced mathematics learning.

Keywords: Guided discovery learning; Instructional design; Mathematical reasoning and proving; Model development; Real analysis instruction

Introduction

The ability to construct mathematical proofs is a central component of advanced mathematics education and constitutes the core of rigorous mathematical thinking. However, recent international studies show that students continue to experience substantial difficulties with proof construction. Persistent challenges include understanding abstract concepts, formulating coherent logical arguments, identifying essential information, distinguishing statements to be proven from justificatory steps, and expressing ideas using precise mathematical language (Arnesen et al., 2025; Risalah et al., 2022; Stylianides et al., 2024; Teledahl et al., 2025). Students also struggle to select appropriate proof strategies and to connect formal definitions with the inferential steps required in an argument (Hamdani et al., 2023).

These difficulties become more pronounced within the context of Real Analysis, a course that demands high levels of abstraction and rigorous manipulation of formal definitions such as limit, continuity, and closedness. Prior research shows that students are often uncertain about why a particular proof method is chosen and find it challenging to initiate and outline a proof when confronted with unfamiliar theorems (Chand, 2021; Pala et al., 2020). They also experience obstacles in employing formal definitions correctly (Herizal et al., 2024) and in selecting appropriate reasoning strategies (Pala et al., 2020). Even high-performing students frequently struggle to choose coherent pathways for proof construction. This conceptual demand makes Real

Analysis one of the most challenging courses even for students with strong prior academic preparation.

Guided Discovery Learning (GDL) is frequently recommended as a pedagogical approach that can support meaningful exploration of mathematical ideas. However, its effectiveness depends heavily on the presence of structured scaffolding that guides students through the transition from exploration toward generalization and formal justification (Alanazi et al., 2025; Hein et al., 2024). Without such support, discovery-based learning can become unproductive and may impose excessive cognitive load (S. Arifin et al., 2020). Parallel to this, research on Mathematical Reasoning and Proving (MRP) Tasks highlights the role of tasks that explicitly trigger conjecturing, generalizing, and justifying as a means to foster coherent argumentation structures (Demiray et al., 2023; Mata-Pereira et al., 2017). The quality of reasoning supported by MRP Tasks is strongly influenced by task design and teacher orchestration (Komatsu et al., 2022). Yet, existing studies on MRP remain largely local and focus primarily on task characteristics rather than integrating such tasks within a complete instructional system (Adleff et al., 2023; Castillo-Martínez et al., 2024).

A theoretical gap arises from the absence of models that structurally integrate GDL with MRP Tasks into a coherent instructional trajectory. Existing literature continues to treat GDL and MRP separately: GDL emphasizes exploratory activity but typically lacks mechanisms that connect exploration to formal proof construction. Systematic reviews of proof and proving show that this separation causes fragmentation between discovery activity and the development of formal mathematical arguments (Chen et al., 2025). MRP research, in contrast, focuses on micro-level reasoning opportunities but does not explain how such tasks should be orchestrated within a discovery-oriented pedagogical design (Valenta et al., 2024; Weingarden et al., 2023). Meanwhile, recent scholarship on scaffolding stresses that discovery learning is effective only when embedded within clearly structured instructional sequences (Jatisunda et al., 2020) although this insight has not been systematically integrated with the MRP literature (Vroom et al., 2024). Consequently, no existing model offers a systematic explanation of how GDL and MRP can be linked to support students' transition from exploration to formal proof.

An empirical gap also persists. Most instructional development studies stop at expert validation or limited implementation, providing insufficient evidence regarding both validity and practicality as core criteria in development research (A. A. Arifin et al., 2024; Hidayati et al., 2021). A robust model must be evaluated conceptually and tested for classroom feasibility.

This study addresses these gaps by developing a Guided Discovery Learning model supported by MRP Tasks specifically designed for Real Analysis instruction. The model is not merely conceptual; it is embodied in complete instructional tools including semester lesson plans, student worksheets containing MRP Tasks, and assessment rubrics. These are evaluated through expert judgment and staged practicality testing with lecturers and students in real classroom settings. Accordingly, the research addresses the following questions: What is the validity level of the Guided Discovery Learning model supported by MRP Tasks for Real Analysis instruction? What is the practicality level of the model for lecturers and students in classroom implementation?

This study contributes on two fronts. Theoretically, it provides the first structurally integrated model combining GDL and MRP Tasks within a unified pedagogical framework. Practically, it offers validated and classroom-ready instructional tools that may serve as a basis for broader proof-oriented interventions in advanced mathematics courses.

Method

Research Design

This study employed an Educational Design Research (EDR) approach following the development and validation orientation. This design was selected because the objective of the study was not to test instructional effectiveness but to develop, refine, and evaluate the conceptual and pedagogical soundness of an integrated instructional model that combines Guided Discovery Learning (GDL) with Mathematical Reasoning and Proving (MRP) tasks. EDR is particularly suitable for studies aiming to establish the theoretical robustness and practical feasibility of instructional innovations through iterative, evidence-based refinement cycles (McKenne et al., 2019; Plomp, 2013).

The research process followed three major phases. Preliminary Research involved reviewing the literature on proof learning, GDL, MRP, and scaffolding in advanced mathematics. This phase also included an analysis of learning difficulties in Real Analysis and the formulation of the conceptual framework guiding the integration of GDL with MRP tasks. Prototyping consisted of designing and iteratively refining instructional components—including the course outline, instructional materials, model book, GDL–MRP-based worksheets, and instructor's guide. Iterative evaluations were conducted through self-evaluation, expert review, one-to-one testing, and small group evaluation. Assessment focused on examining the practicality of the model through a field test in an actual Real Analysis course. This phase aimed to determine whether the

model was implementable, comprehensible, and instructionally workable under real classroom constraints.

Consistent with early stage EDR, the study did not assess instructional effectiveness. The focus was restricted to two essential indicators of instructional feasibility, namely validity and practicality, as recommended by design research standards (Nieveen et al., 2013).

Participants and Context

The study was conducted in an undergraduate mathematics education program at a state university. Participants were purposively selected to represent key stakeholders in the validation and implementation of the model. They included: Five expert validators, consisting of specialists in mathematics education, instructional design, and Real Analysis. The experts evaluated content validity, construct coherence, and logical consistency of the model and its accompanying materials. Undergraduate students enrolled in a Real Analysis course, who participated in one-to-one evaluation, small group evaluation, and the field test. Their responses provided direct evidence of the model's usability and clarity. The course instructor, who implemented the model during the field test and evaluated its practicality through observations and reflective assessments.

The involvement of these three stakeholder groups ensured multidimensional evidence of validity and practicality, reflecting both conceptual rigor and real-world applicability. All procedures adhered to institutional ethical guidelines, including informed consent and approval from the host department.

Research Instruments

Four main categories of instruments were employed:

Expert Validation Sheets

These instruments were designed to evaluate content relevance, construct coherence, logical consistency, and instructional quality of the model components. Validators rated items using an ordinal scale. Validity was quantified using Aiken's V, while interrater consistency was assessed using the Interclass Correlation Coefficient (ICC) (Ayre et al., 2014; Koo et al., 2016).

Student Practicality Instruments

A Likert-scale questionnaire assessed clarity, readability, ease of navigating the model, and usefulness of MRP tasks during one-to-one, small group, and field test stages. Items captured students' perceptions of both cognitive and procedural usability.

Instructor Practicality Instruments

These included an implementation observation sheet and a lecturer perception questionnaire. Observations covered model functioning, instructional sequence, engagement level, and operational reliability.

Qualitative Documentation

This supporting dataset included expert comments, observational field notes, and student narrative feedback. These qualitative sources enabled triangulation of quantitative findings. All instruments were validated by a panel of five experts. Aiken's V scores ranged from 0.62 to 0.88, and ICC values ranged from 0.721 to 0.867, indicating strong validity and acceptable-to-high reliability.

Data Collection Procedures

Data collection occurred across three iterative formative evaluation stages:

Expert Review

Five validators independently assessed the instructional components (course outline, materials, model book, worksheets, instructor guide) for conceptual coherence, alignment with Real Analysis content, and logical structure. Quantitative scores (Aiken's V and ICC) and qualitative comments informed subsequent revisions.

One-to-One and Small Group Evaluation

The one-to-one evaluation involved three students who examined readability, clarity, and procedural flow. The small group evaluation involved eight students assessing task sequencing, instructional timing, and overall learning experience. Both stages produced refinements to the prototype, particularly in layout, clarity of instructions, and cognitive sequencing of MRP tasks.

Field Test

The refined model was implemented in a Real Analysis course. The instructor completed observation sheets evaluating model function, instructional sequencing, student engagement, and reliability. Students completed a practicality questionnaire. Field notes were collected to document classroom dynamics and identify potential instructional constraints.

This stage served to determine the model's feasibility under authentic classroom conditions. The model was implemented in a Real Analysis class over several sessions. The course instructor completed an observation sheet assessing model functions, instructional syntax, engagement, and reliability. Students completed a practicality questionnaire, and field notes were collected to supplement quantitative

results. The purpose of this implementation was to evaluate the feasibility and enactment of the model in an authentic classroom environment.

Data Analysis
Validity Analysis

Aiken's V was used to assess expert agreement on item relevance, with scores ≥ 0.70 indicating acceptable validity (Esteban et al., 2024). ICC values provided evidence of rater consistency, with the threshold of 0.60–0.85 indicating moderate to high reliability (Koo et al., 2016).

Practicality Analysis

Practicality data were analyzed through (a) mean Likert scores, (b) the feasibility coefficient (κ), and (c) categorical interpretation of practicality levels (moderate, high, very high). Triangulation of student responses, instructor evaluations, and field notes was conducted to strengthen the credibility of the findings (Miles et al., 2014).

Result and Discussion

Expert review results showed that all instructional components achieved high levels of conceptual and structural validity. Aiken's V values ranged from 0.62 to 0.88 across the course outline, instructional materials, model book, GDL MRP based student worksheets, and instructor guide. These values fall within the high to very high category, indicating strong alignment between the model's structure and the epistemic requirements of Real Analysis.

The student worksheet and the instructor guide recorded the highest validity values, namely 0.88 and 0.81, which reflects strong logical coherence, representational clarity, and appropriate sequencing of reasoning tasks. The instructional materials obtained the lowest value, 0.62, yet this score remains within the high category and required only minor editorial revisions. A summary of expert validation results is presented in Table 1. This table supports the conclusion that the model components meet essential standards of content accuracy, construct clarity, and internal consistency.

Table 1. Expert Validation Results (Aiken's V)

Instrument	Validators (n)	Aiken's V	Validity Level
Semester Lesson Plan (SLP)	5	0.75	High
Teaching Materials	5	0.62	High
GDL–MRP Model Book	5	0.79	High
Student Worksheet (LKM)	5	0.88	Very High
Instructor's Guidebook	5	0.81	Very High

Interrater Reliability (ICC)

The reliability analysis using the Interclass Correlation Coefficient showed that ICC values ranged from 0.715 to 0.867 for the five instructional components. These values represent moderate to high reliability and confirm that expert judgments were stable and consistent.

The teaching materials and the instructor guidebook demonstrated the highest reliability values (ICC 0.867 and 0.841), indicating strong agreement among validators. In contrast, the course outline, the GDL MRP model book, and the student worksheet produced ICC values within the moderate reliability range (0.721, 0.715, and 0.783). Table 2 summarizes the ICC values obtained during instrument validation and provides additional evidence that the model components were evaluated consistently across independent reviewers.

Table 2. ICC Values for Instrument Validation

Instrument	Validators (n)	Aiken's V	Validity Level
Semester Lesson Plan (SLP)	5	0.721	Moderate
Teaching Materials	5	0.867	High
GDL–MRP Model Book	5	0.715	Moderate
Student Worksheet (LKM)	5	0.783	Moderate
Instructor's Guidebook	5	0.841	High

Model Practicality

Model practicality was examined across three iterative stages: one to one evaluation, small group evaluation, and field testing in an actual classroom. Across all stages, the results indicate consistently high levels of usability, clarity, and instructional operability.

One to One Evaluation

The one to one evaluation involved three students who examined the clarity of instructions, readability, flow of activities, and ease of engaging with the GDL MRP sequence. Practicality values ranged from moderate to very high, with an overall mean of 0.74, which falls within the high category.

Students reported that the model was easy to follow, beneficial for structuring their reasoning, and supportive of their understanding of Real Analysis concepts. Only one indicator, related to time management for completing exercises, fell within the moderate category. This finding informed minor revisions for subsequent prototypes. The results are presented in Table 3.

Table 3. Practicality in the One to One Evaluation

Aspect Assessed	κ	Category
The model provides guided and enjoyable learning	0.75	High
The model facilitates learning in Real Analysis	0.81	Very High
I experienced clear benefits from using the model	0.81	Very High
I felt motivated to engage in subsequent topics	0.61	High
Ease of performing roles according to model steps	0.75	High
Ease of collaborating to solve problems	0.75	High
Ease of reading symbols, expressions, and text	0.68	High
Completing tasks within the allocated time	0.53	Moderate
Layout does not hinder visibility	0.75	High
I like the design of the book	0.91	Very High
Average κ	0.74	High

Small Group Evaluation

Six students participated in the small group stage. The results again showed consistently high practicality, with an overall mean of 0.72. Several indicators demonstrated improvement relative to the one to one stage, especially in readability, clarity of layout, and alignment between task structure and intended reasoning processes.

Students indicated that the revised prototype offered clearer instructions, more intuitive visual design, and better alignment between exploration prompts and proof related tasks. As presented in Table 4, all aspects were consistently rated within the high to very high range, indicating that the revisions effectively improved the model's usability, particularly in readability and layout.

Table 4. Practicality in the Small Group Evaluation

Evaluated Aspect	κ	Category
The model is guided and enjoyable	0.68	High
The model facilitates Real Analysis learning	0.68	High
I experience benefits from using the model	0.68	High
I feel motivated to continue to the next material	0.68	High
Ease of performing roles according to the model steps	0.65	High
Ease of collaborating to solve problems	0.61	High
Timeliness in completing exercises	0.72	High
Readability of letters, numbers, words, and sentences	0.65	High
Layout does not interfere with visual comfort	0.81	Very High
Preference for the book's physical form	0.81	Very High
Average κ	0.72	High

Field Test

The field test assessed model practicality based on both instructor observations and student responses in a real class setting.

Instructor Observations

The instructor rated the model's functioning, instructional sequence, engagement, and operational reliability. The overall practicality value reached 0.84, which is categorized as very high. The model received perfect scores for engagement and reliability, indicating that it supported classroom dynamics well and did not introduce instructional burdens. Table 5 presents the results from two independent observers.

Table 5. Practicality from Instructor Observations (Field Test)

Assessed Dimension	Number of Items	Observer 1	Observer 2	κ	Category
Function	5	3	3	0.60	High
Syntax	20	14	16	0.75	High
Attractiveness	2	2	2	1.00	Very High
Operational reliability	2	2	2	1.00	Very High
Average κ				0.84	Very High

Student Responses in the Field Test

Student feedback in the field test reflected high practicality, with an overall feasibility value of 0.72. The highest scores were obtained for layout clarity and time management, both of which fell into the very high

category. Students reported that the model guided them through structured reasoning, supported their understanding, and facilitated transitions from exploration to justification. Table 6 summarizes the results.

Table 6. Student Practicality Responses (Field Test)

Statement	к	Category
The model provides a guided and enjoyable learning experience	0.61	High
The model facilitates learning in Real Analysis	0.71	High
I clearly understand my tasks and roles in each phase	0.72	High
I am motivated to engage with subsequent material	0.79	High
I experience no difficulty performing my assigned roles	0.74	High
I perceive meaningful benefits from using the model	0.61	High
I am motivated to continue with the next learning activities	0.72	High
Letters, symbols, and text are easy to read	0.75	High
The layout does not interfere with visual clarity	0.75	Very High
I can complete the exercises within the allocated time	0.82	Very High
Average к	0.72	High

Together, the results from the three stages indicate that the model is feasible for classroom implementation, understandable to students, and usable for instructors without requiring additional instructional resources.

Synthesis of Validity and Practicality

The integration of validity and practicality findings demonstrates that the GDL MRP model meets the two foundational requirements for early stage instructional development: theoretical soundness and operational feasibility. High Aiken's V values confirm that the model aligns with conceptual standards for reasoning and proving in advanced mathematics. Meanwhile, consistently high practicality results across evaluation stages show that the model can be implemented reliably in real teaching conditions.

This convergence suggests that the model is ready for broader implementation and empirical testing of instructional effectiveness in subsequent research. It establishes a coherent bridge between conceptual exploration in GDL and structured argumentation in MRP tasks, addressing gaps identified in recent literature and providing a rigorous foundation for future innovation in proof oriented mathematics instruction.

Discussion
Model Validity: Alignment with Theoretical Principles and Advanced Mathematics Learning

The high validity ratings across all components of the GDL MRP model demonstrate strong alignment between the model's conceptual structure and established theories in advanced mathematics learning. The phased design that includes exploration, guided inquiry, and the construction of formal arguments reflects core principles of scaffolded learning, where instructional support is gradually withdrawn as learners

develop autonomy in deductive reasoning (Murphy et al., 2025; Sun et al., 2023). Such sequencing is essential for supporting students' transition from procedural engagement toward independent proof construction in Real Analysis, a domain known for its emphasis on rigor and conceptual precision.

The integration of MRP Tasks further reinforces this theoretical alignment by operationalizing cognitive apprenticeship principles. Prior research highlights that learners must be exposed to expert processes of conjecturing, refining, justifying, and organizing proofs to bridge the gap between informal reasoning and formal argumentation (Simsek, 2020; Subarinah et al., 2021). The structure of MRP Tasks in this study mirrors these expert pathways, enabling students to articulate intermediate reasoning and gradually internalize proof strategies. The consistently high Aiken's V coefficients indicate that experts evaluated the model as conceptually coherent and suitably aligned with the epistemic demands of Real Analysis. In addition, stable ICC values (Ayre et al., 2014) confirm that the model components were interpreted consistently across evaluators, demonstrating clarity and operational precision in the instructional design.

Comparison with earlier development studies reveals a marked methodological distinction. Much existing research relies on single-round expert judgments or descriptive averages to assert validity (Elfizon et al., 2021; Oluwadele et al., 2024). Such approaches offer limited insight into internal coherence and seldom articulate how indicators correspond to theoretical assumptions, resulting in claims that are administrative rather than analytical. The present study challenges this tendency by demonstrating that rigorous validity requires hierarchical coherence, clearly defined conceptual boundaries, and explicit alignment between

theoretical principles and implementation mechanisms. In doing so, this study advances the methodological rigor of instructional model validation in advanced mathematics education.

Practicality of the Model: Alignment with Real Classroom Implementation

The practicality of the GDL MRP model is supported by consistently high outcomes across one-to-one evaluations, small group testing, and real classroom implementation. Lecturers' reports that a model's syntax is straightforward, manageable within realistic time constraints, and free from additional instructional burden align closely with established criteria for classroom usability in advanced mathematics instruction (Estrada Molina et al., 2022). This procedural clarity allowed instructors to orchestrate guided inquiry phases effectively while maintaining the conceptual pacing required in Real Analysis.

Student responses further substantiated the model's practicality. High scores on worksheet readability, clarity of instructions, and ease of role execution suggest that learners were able to engage with the tasks without experiencing extraneous cognitive load. Prior work has shown that Guided Discovery Learning succeeds only when task structures are explicit and coherent, particularly in multi-step reasoning contexts (S. Arifin et al., 2020). The high practicality scores across all evaluation phases indicate that this model successfully mitigated risks of unguided or unproductive exploration.

The field test provided the strongest evidence of operational feasibility. Lecturer observations reached the highest categories, revealing that the model not only functioned technically but also supported classroom interaction, collaboration, and sustained reasoning. The convergence of lecturer and student evaluations across stages demonstrates stable implement ability in authentic instructional conditions. By contrast, many prior development studies adopt minimalistic interpretations of practicality, often equating it with the absence of implementation barriers based on brief or single-class trials (Noor et al., 2023; Zhang et al., 2023). Such approaches neglect essential dimensions of practicality, particularly responsiveness, which concerns how instructional models adapt to diverse learners and dynamic classroom environments. By incorporating indicators of feasibility, usability, and responsiveness, the present study offers a more rigorous and operationally grounded definition of practicality than typically found in development research.

Implications for Advanced Mathematics Instruction

The convergence of high validity and practicality provides strong evidence that the GDL MRP model is theoretically sound and operationally feasible for use in Real Analysis. The model's design aligns with the cognitive demands of advanced mathematical reasoning, particularly the transition from exploratory engagement to formal proof construction. At the same time, practicality results confirm that the model can be implemented effectively without additional resources or specialized training, meeting two standards often emphasized yet rarely achieved simultaneously in instructional development research (Nieveen et al., 2013; Plomp, 2013).

The findings contribute meaningfully to the broader literature by advancing discussions beyond task-level opportunities for reasoning and proving (Stylianides et al., 2024; Weingarden et al., 2023). While many studies examine how individual tasks foster conjecturing or justifying, few address how sequences of such tasks can be integrated into full instructional models. This study fills that gap by demonstrating how MRP Tasks, when embedded within Guided Discovery Learning, function cohesively within structured instructional trajectories rather than as isolated elements. This approach offers a more holistic understanding of how reasoning and proving can be orchestrated across extended instructional episodes.

Methodologically, the study underscores the necessity of multi-stage validation in model development. Unlike studies that rely only on expert review or short-term pilots (Elfizon et al., 2021; Oluwadele et al., 2024), the present research integrates expert assessment, iterative refinement, and field implementation, establishing a more rigorous and defensible methodological standard. Pedagogically, the model provides a systematic mechanism for helping students bridge the gap between exploration and formal proof in Real Analysis. Because the GDL MRP model demonstrated both conceptual rigor and adaptability, it holds strong potential for implementation in other upper-division mathematics courses such as Abstract Algebra, Topology, and Number Theory, where guided inquiry and proof-based reasoning are equally essential.

Limitations and Directions for Future Research

Despite its strengths, the study has several limitations that offer opportunities for further investigation. First, the research focused exclusively on early-stage EDR, emphasizing validity and practicality but not instructional effectiveness. Future studies should examine the model's causal impact on students' proof construction skills and broader mathematical reasoning. Second, the field test was conducted within a single institutional context, which may limit generalizability. Cross-institutional trials could reveal how the model adapts to varying instructional cultures, student

backgrounds, and curricular structures (Noor et al., 2023; Zhang et al., 2023).

Further research is also needed to analyze how MRP-based scaffolding supports the long-term development of proving competencies. Longitudinal studies could clarify how students internalize conjecturing, justifying, and proving processes over time and across mathematical domains. Moreover, future work should explore the model's transferability to other advanced mathematics courses and whether its guided discovery structure requires adaptation for domains with different epistemic norms. Such investigations will deepen understanding of the cognitive mechanisms underlying students' transition from informal reasoning to formal proof and will position the GDL MRP model as both a validated instructional design and a methodological reference point for innovations in advanced mathematics pedagogy.

Conclusion

This study developed and systematically evaluated a Guided Discovery Learning model integrated with Mathematical Reasoning and Proving (MRP) Tasks for Real Analysis instruction, focusing on two foundational criteria in instructional design research: validity and practicality. Expert evaluations revealed that the model demonstrates strong conceptual and structural validity. Its instructional syntax, supporting materials, and overall design were consistently rated in the high category, as reflected in stable Aiken's V and ICC values. These results indicate that the model aligns well with the academic, cognitive, and pedagogical demands of advanced mathematical learning. The model also exhibited high practicality across sequential evaluation phases, including one-to-one evaluation, small group testing, and full classroom implementation. Both lecturers and students reported that the instructional steps were clear, executable, and compatible with typical classroom conditions. The MRP-based worksheets effectively structured students' reasoning processes, guiding them from exploratory engagement toward conjecture refinement and formal proof construction. These findings confirm that the model is not only theoretically coherent but also operationally feasible for routine use in Real Analysis courses. Collectively, the results contribute to the development of advanced mathematics pedagogy by presenting a coherent instructional model that unifies guided discovery with structured reasoning and proving. The convergence of validity and practicality provides a strong foundation for future investigations into the model's instructional effectiveness, its transferability to other upper-division mathematics courses, and the cognitive mechanisms that support students' transitions from informal reasoning to

formal proof. The GDL–MRP model therefore offers both a practical instructional resource for university educators and a conceptual contribution to ongoing efforts to strengthen mathematical reasoning and proving in higher education

Author Contributions
Conceptualization, Selvy Musdalifah, Sudarsana and Sukasyasa; methodology, Selvy Musdalifah, Dasa Ismaimuza; validation, Dasa Ismaimuza, Sutji Rochaminah. and Nurhayadi.; formal analysis, Selvy Musdalifah, Sudarsana.; investigation, Selvy Musdalifah, Nurhayadi; resources, Sukayasa; writing—original draft preparation, Selvy Musdalifah, Dasa Ismaimuza; writing—review and editing, Selvy Musdalifah; supervision, Dasa Ismaimuza; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest
The authors declare no conflict of interest.

References

Adleff, A., Ross, N., König, J., & Kaiser, G. (2023). Types of mathematical tasks in lower secondary classrooms in Germany. *Educational Studies in Mathematics*, *114*(3), 371–392. https://doi.org/10.1007/s10649-023-10254-9

Alanazi, A., Osman, K., & Halim, L. (2025). Enhancing physics problem-solving skills through guided discovery and scaffolding strategies. *LUMAT: International Journal on Math, Science and Technology Education*, *12*(4), 5. https://doi.org/10.31129/LUMAT.12.4.2329

Arifin, A. A., Ramdani, A., & Andayani, Y. (2024). Development of Learning Tools Based on the Culturally Responsive Transformative Teaching Model with a Socio-Scientific Issues Approach: Evaluation of Validity and Practicality. *Jurnal Penelitian Pendidikan IPA*, *10*(11), 9141–9155. https://doi.org/10.29303/jppipa.v10i11.9188

Arifin, S., Wahyudin, W., & Herman, T. (2020). The effects of contextual group guided discovery learning on students' mathematical understanding and reasoning. *Jurnal Prima Edukasia*, *8*(2), 106–114. https://doi.org/10.21831/jpe.v8i2.33059

Arnesen, K. K., & Skartsæterhagen, Ø. I. (2025). Mathematical induction in education research: a systematic review. *Educational Studies in Mathematics*, *119*(1), 79–100.

https://doi.org/10.1007/s10649-024-10373-x

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, *47*(1), 79–86. https://doi.org/10.1177/0748175613513808

Castillo-Martínez, I. M., Flores-Bueno, D., Gómez-Puente, S. M., & Vite-León, V. O. (2024). AI in higher education: a systematic literature review. *Frontiers in Education*, *9*(7), 1–24. https://doi.org/10.3389/feduc.2024.1391485

Chand, H. B. (2021). Difficulties Experienced by Undergraduate Students in Proving Theorems of Real Analysis. *Scholars' Journal*, *4*(December), 149–163. https://doi.org/10.3126/scholars.v4i1.42475

Chen, S. I, Yu, W., Dou, G., & Zhang, Q. (2025). A Review on Mechanical Proving and Formalization of Mathematical Theorems. *IEEE Access*, *13*(March), 50672–50686. https://doi.org/10.1109/ACCESS.2025.3552634

Demiray, E., Işıksal-Bostan, M., & Saygı, E. (2023). How argumentation relates to formal proof process in geometry. *International Electronic Journal of Mathematics Education*, *18*(3), em0741. https://doi.org/10.29333/iejme/13214

Elfizon, Dewi, M., Ganefri, & Sukardi. (2021). Vocational Education Project-Based Technopreneurship Learning Model. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(2), 3119–3126. https://doi.org/10.17762/turcomat.v12i2.2356

Esteban, R. F. C., Mamani-Benito, O., Huancahuire-Vega, S., Casildo-Bedón, N., Turpo-Chaparro, I., Turpo-Chaparro, J. E., & Josué. (2024). Design and validation of a scale of motivation for scientific publication in Peruvian university professors (MoSCPU-UP). *Frontie in Education*, *9*(September), 1–8. https://doi.org/10.3389/feduc.2024.1378626

Estrada Molina, O., Fuentes-Cancell, D. R., & García-Hernández, A. (2022). Evaluating usability in educational technology: A systematic review from the teaching of mathematics. *LUMAT: International Journal on Math, Science and Technology Education*, *10*(1), 65–88. https://doi.org/10.31129/LUMAT.10.1.1686

Hamdani, D., Sa'dijah, C., Subanji, S., & Subarinah, S. (2023). Development of ACERA Learning Model Based on Proof Construction Analysis. *JTAM (Jurnal Teori Dan Aplikasi Matematika)*, *7*(3), 560. https://doi.org/10.31764/jtam.v7i3.12354

Hein, K., & Prediger, S. (2024). Scaffolds for seeing, using, and articulating logical structures in proofs : Design research study with high school students. *Journal of Mathematical Behavior*, *74*(December 2023), 101123.

https://doi.org/10.1016/j.jmathb.2023.101123

Herizal, Marham, & Akmal, N. (2024). Students' Errors in Constructing Mathematical Proofs by Direct Method. *CMScE International Conference On Mathematics And Science Education*, *2024*, 651–659. https://doi.org/10.18502/kss.v9i13.15969

Hidayati, S., Susilawati, S., & Harjono, A. (2021). Validity and Practicality of Problem Based Learning (PBL) Model Learning Tools to Improve Students' Conceptual Understanding. *Prisma Sains : Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *9*(1), 82. https://doi.org/10.33394/j-ps.v9i1.3966

Jatisunda, M. G., Suciawati, V., & Nahdi, D. S. (2020). Discovery Learning with Scaffolding To Promote Mathematical Creative Thinking Ability And Self-Efficacy. *Al-Jabar : Jurnal Pendidikan Matematika*, *11*(2), 351–370. https://doi.org/10.24042/ajpm.v11i2.6903

Komatsu, K., & Jones, K. (2022). Generating mathematical knowledge in the classroom through proof, refutation, and abductive reasoning. *Educational Studies in Mathematics*, *109*(3), 567–591. https://doi.org/10.1007/s10649-021-10086-5

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Mata-Pereira, J., & da Ponte, J. P. (2017). Enhancing students' mathematical reasoning in the classroom: teacher actions facilitating generalization and justification. *Educational Studies in Mathematics*, *96*(2), 169–186. https://doi.org/10.1007/s10649-017-9773-4

McKenne, S., & Reeves, T. (2019). *Conducting Educational Design Research*. Routledge.

Miles, M ., Huberman, A ., & Saldana, J. (2014). *Qualitative Data Analysis A Methods Sourcebook*. SAGE Publications.

Murphy, C., Muir, T., & Thomas, D. (2025). Scaffolding Collaboration in Early Years Mathematics : A Practice - Based Case Study in Teaching Multiplicative Grouping. *Early Childhood Education Journal*, *53*(8), 3207–3217. https://doi.org/10.1007/s10643-025-01928-5

Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. In *Part A: An introduction* (pp. 152–169). SLO.

Noor, A. F., Yunus, R., Suyidno, S., & Fahmi, F. (2023). Development of Predict-Observe-Explain (Poe) Based Authentic Problems' Instructional Package To Improve Students' Critical Thinking Skills. *Jurnal Pendidikan Matematika Dan IPA*, *14*(1), 69. https://doi.org/10.26418/jpmipa.v14i1.53932

Oluwadele, D., Singh, Y., & Adeliyi, T. (2024). Operationalizing a Weighted Performance Scoring Model for Sustainable e-Learning in Medical Education: Insights from Expert Judgement. *Electronic Journal of E-Learning*, *22*(8), 24–40. https://doi.org/10.34190/ejel.22.8.3427

Pala, O., & Narlı, S. (2020). The Role of the Formal Knowledge in the Formation of the Proof Image : A Case Study in the Context of the Infinite Sets *. *Turkish Journal of Computer and Mathematics Education*, *11*(3), 584–618. https://doi.org/10.16949/turkbilmat.702540

Plomp, T. (2013). Educational Design Research: An Introduction. In *Educational design research* (pp. 10–51). SLO.

Risalah, D., & Hodiyanto, H. (2022). Mathematics communication as an alternative to overcome the obstacles of undergraduate students in mathematical proof. *International Journal of Trends in Mathematics Education Research*, *5*(2), 125–132. https://doi.org/10.33122/ijtmer.v5i2.141

Simsek, Z. Z. (2020). Constructing-Evaluating-Refining Mathematical Conjectures and Proofs. *International Journal for Mathematics Teaching and Learning*, *21*(2), 197–215. https://doi.org/10.4256/ijmtl.v21i2.263

Stylianides, G. J., Stylianides, A. J., & Moutsios, A. (2024). Proof and proving in school and university mathematics education research : a systematic review. *ZDM – Mathematics Education*, *56*(1), 47–59. https://doi.org/10.1007/s11858-023-01518-y

Subarinah, S., Sridana, N., & Prayitno, S. (2021). Development of Learning Model Based on Cognitive Processes in Mathematical Investigation. *Proceedings of the 2nd Annual Conference on Education and Social Science (ACCESS 2020)*, *556*(Access 2020), 531–533. https://doi.org/10.2991/assehr.k.210525.143

Sun, L., Kangas, M., Ruokamo, H., & Siklander, S. (2023). A systematic literature review of teacher scaffolding in game-based learning in primary education. *Educational Research Review*, *40*(July), 100546. https://doi.org/10.1016/j.edurev.2023.100546

Teledahl, A., Kilhamn, C., Ahl, L. M., & Helenius, O. (2025). Defining and measuring quality in students' mathematical writing: a systematic literature review. *Mathematics Education Research Journal*, *37*(3), 557–576. https://doi.org/10.1007/s13394-024-00501-4

Valenta, A., Rø, K., & Iversen, S. (2024). A framework for reasoning in school mathematics : analyzing the development of mathematical claims. *Educational Studies in Mathematics*, *116*, 91–111. https://doi.org/10.1007/s10649-024-10309-5

Vroom, K., & Ellis, B. (2024). Sociomathematical scaffolding as students engage in revising draft definitions, conjectures, and proofs. *Educational Studies in Mathematics*, *116*(1), 27–47. https://doi.org/10.1007/s10649-023-10294-1

Weingarden, M., & Buchbinder, O. (2023). Teacher learning to teach mathematics via reasoning and proving: a discursive analysis of lesson plans modifications. *Frontiers in Education*, *8*(August), 1–20. https://doi.org/10.3389/feduc.2023.1154531

Zhang, Y., Huang, T., Tang, M., Meng, L., Wu, X., & Chen, K. (2023). An effectiveness evaluation of a community - based course for medical students : a randomized controlled trial in the teaching of epidemiology. *BMC Medical Education*, *23*(807), 1–9. https://doi.org/10.1186/s12909-023-04787-z