EDUMATSAINS
JURNAL PENDIDIKAN MATEMATIKA DAN SAINS
E-ISSN:2527-7235
Volume 11 Issue. 2, October 2026, pp 179-194
http://ejournal.uki.ac.id/index.php/edumatsains

Prospective Mathematics Teachers' Concept Images of Group Structure: From Binary Operations to Axiomatic Coordination

I Made Suarsana [1*], Made Juniantari [2], I Made Sugiarta[3], Elah Nurlaelah[4], Tia Purniati[5], Bill Chairy Rizky Bustaren[6]

[1,2,3]Department of Mathematics Education, Faculty of Mathematics and Natural Science, Universitas Pendidikan Ganesha

[4,5,6]Department of Mathematics Education, Faculty of Mathematics and Natural Science Education, Universitas Pendidikan Indonesia

e-mail: made.suarsana@undiksha.ac.id

Article Info

Article history:
Received: August 2nd, 2026
Revised: October 23rd, 2026
Accepted: October 26th, 2026
Available online: October 31st, 2026

https://doi.org/10.33541/edumatsains.v11i2.8643

Abstract

Understanding group structure requires the coordination of a set, a binary operation, universal quantifiers, and the group axioms. Yet how these meanings are activated across connected tasks remains incompletely understood. This interpretive qualitative study examined prospective mathematics teachers' concept images as part of the prospective-analysis phase of Didactical Design Research (DDR). Thirty-six students enrolled in an abstract algebra course completed an eight-task diagnostic instrument, and eight purposively selected students participated in task-based interviews. Written responses, interview transcripts, and reflective notes were analyzed as meaning-making episodes and compared across tasks and representations. Three task-sensitive configurations were distinguished: operation-oriented reasoning, in which familiar calculations or visibly active operands dominated judgments; per-axiom reasoning, in which requirements were checked separately without consistent universal justification; and structurally coordinated reasoning, in which the set, operation, quantifiers, and relevant axioms jointly governed classification, proof, and refutation. Definition-governed reasoning sometimes weakened when the representation changed, particularly in operation-table tasks. These configurations describe evoked patterns of reasoning, not permanent learner types or developmental stages. The analysis connects binary-operation recognition to axiom coordination and suggests preliminary didactical needs involving non-prototypical operations, quantifiers, and representational flexibility. No didactical intervention was evaluated in this study.

Keywords: abstract algebra, concept image, binary operation, group axioms, didactical design research

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

1. Introduction

Group structure is a foundational idea in abstract algebra because it enables numbers, permutations, matrices, transformations, and functions to be examined through a common structural lens. For prospective mathematics teachers, learning groups also marks a transition from procedure-based activity to reasoning with definitions, quantifiers, and proof. This transition is demanding: students must coordinate prerequisite knowledge with unfamiliar representations and sustain conceptual reasoning when a task changes form (Agustyaningrum et al., 2021; Biehler et al., 2024; Veith et al., 2025).

A group is not merely a list of closure, associativity, identity, and inverse. A binary operation on a set S is a function $S \times S \rightarrow S$ that assigns exactly one element of S to every ordered pair. Associativity is a universal property of that operation; an identity must work from both sides for every element; and each inverse is defined relative to the specified operation and identity. Reasoning about a group therefore requires the coordinated use of the set, operation, domain and codomain, universal quantifiers, and relations among the axioms.

Knowing a formal definition does not guarantee that it governs reasoning. Tall and Vinner (1981) distinguish the concept definition from the concept image, which includes the examples, representations, procedures, properties, and experiences associated with a concept. Only part of this structure may be activated in a particular situation. This evoked concept image can be aligned with, partially aligned with, or in conflict with the formal definition. Thus, a student may state a definition accurately but revert to an arithmetic prototype when confronting an unfamiliar operation or representation.

Studies of abstract algebra show that students may interpret a binary operation as a calculation or recognizable symbol rather than as a function on a Cartesian product (Aksu et al., 2018; Melhuish, Ellis, et al., 2020). They may attend to the rule while overlooking its domain, codomain, everywhere-defined property, or closure (Uscanga et al., 2024; Uscanga & Cook, 2024). Further difficulties arise when identity, inverse, and associativity must be coordinated within one structure, particularly across symbolic rules, examples, proofs, and Cayley tables (Dubinsky et al., 1994; Veith et al., 2022b, 2022a). Familiar examples can support meaning, but surface similarity may also encourage inappropriate procedural imitation (Hicks, 2024). Concept-test research consequently emphasizes varied examples, non-examples, and representations to reveal whether decisions are definition-governed (Feudel & Unger, 2025; Veith et al., 2025).

Previous studies have documented difficulties with binary operations and individual group axioms, but they have less often examined how definitional reasoning is evoked across a connected set of tasks involving operation recognition, axiom checking, classification, proof, refutation, and Cayley tables. Examining these demands together can reveal recurring interpretive criteria and exceptions without treating students' local responses as evidence of stable learner types or a sequential developmental trajectory.

Accordingly, this study asks: What characterizes prospective mathematics teachers' evoked concept images of group structure across tasks concerning binary operations, group axioms, and structural justification? We distinguish three analytic configurations: operation-oriented, per-axiom, and structurally coordinated. Building on Tall and Vinner's (1981) distinction between

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

concept image and concept definition, the analysis examines how strongly the formal definition governs judgments as tasks and representations change. These configurations are context-sensitive descriptions of reasoning, not proposed developmental stages

2. Methods

2.1. Research Design and Participants

This interpretive qualitative study was situated within the prospective-analysis phase of Didactical Design Research (DDR). DDR provided the broader methodological framework for examining students' mathematical meanings and anticipating didactical needs before designing learning situations (Suryadi, 2013). The present article analyzes written and interview-based reasoning about binary operations and group structure. An interpretive orientation guided repeated readings of participants' explanations and the task contexts in which they were produced; hermeneutic phenomenology was not employed as a separate methodological design. Although DDR encompasses prospective, metapedadidactic, and retrospective analyses, this paper reports only the prospective component. It does not report the implementation or evaluation of a didactical design. Rather, the analysis provides an empirical basis for provisional didactical-pedagogical anticipations to be examined in subsequent DDR stages.

Participants were 36 prospective mathematics teachers from one cohort of an undergraduate abstract algebra course in Indonesia. During the even semester of 2025/2026, they had received formal instruction on binary operations, group axioms, examples and non-examples, and operation tables. All participated voluntarily. Eight students (P03, P07, P11, P16, P21, P27, P31, and P35) were selected for task-based interviews through maximum-variation purposive sampling. Selection considered differences in personal definitions, response patterns, consistency across representations, and negative or deviant cases, rather than total scores.

2.2. Instruments and Data Collection

Data comprised an eight-task diagnostic concept-image test, task-based semi-structured interviews, and researcher reflective notes. The tasks elicited personal definitions, examples and non-examples, classification, changes of representation, axiom verification, proof, and refutation. The test was completed in 150 minutes. Interviews additionally explored participants' reasoning with an operation table. Using varied forms was important because the aspects of a concept image evoked can differ across tasks and representations (Tall & Vinner, 1981). Table 1 summarizes the three analytic dimensions covered by the diagnostic instrument.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Table 1.
Analytic Coverage of the Diagnostic Tasks

Dimension	Concept examined	Form of evidence
Binary operation	Two inputs, function, domain, codomain, closure	Definition, example production, classification, and representation translation
Group axioms	Associativity, identity, and inverse	Verification, justification, error analysis, and counterexamples
Structural coordination	Relations among set, operation, quantifier, and all axioms	Structural classification, boundary cases, proof, and refutation

A blueprint mapped the content, reasoning processes, representations, and expected evidence. One abstract-algebra expert and one mathematics-education expert reviewed the draft for conceptual accuracy, domain coverage, task relevance, and response potential. Readability testing and cognitive interviews with eight students outside the research cohort informed revisions to wording, task sequence, boundary cases, representations, and interview prompts.

Interviews probed the reasoning behind written responses and the stability of that reasoning when participants encountered comparison cases. Neutral prompts included “What do you mean by this statement?”, “How do you know that?”, and “Does that reasoning hold for every member of the set?” Instructional feedback was withheld during probing. With participants' consent, interviews were recorded, transcribed verbatim, anonymized, and linked to written responses. Reflective notes documented contextual information, emerging conjectures, interview decisions, and follow-up questions.

2.3. Data Analysis

Written responses, interview transcripts, and reflective notes were read repeatedly. The primary unit of coding was a meaning-making episode, defined as a task-specific mathematical decision together with its supporting claim, representation, justification, conflict, or shift in reasoning. Episodes, rather than participants, were assigned codes. Each participant's response set was reviewed before segmentation to preserve links between local statements and reasoning across tasks.

Analysis proceeded in five connected steps. First, written responses and interview transcripts were segmented into meaning-making episodes. Second, the mathematical criteria used to justify each decision were coded, including reliance on arithmetic prototypes, the active-operand criterion, tests of only one side of an identity, example-based verification, explicit universal quantification, and coordination of axioms. Third, these codes were compared across tasks and representations within participants and across participants. Fourth, recurring forms of definition use were synthesized into three configurations while contradictory or representation-sensitive episodes were retained. Finally, each interpretation was checked against the associated response and task context. Broader descriptive codes, such as example-bound, procedure-dominated, representation-dependent, and definition-aligned reasoning, informed this comparison.

Operational criteria were applied to episodes, not persons. An episode was coded as operation-oriented when its justification privileged arithmetic actions, familiar symbols, or the apparent influence of both operands over the definition of a mapping $S \times S \rightarrow S$; as per-axiom when

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

individual properties were checked but their universal or joint applicability was incompletely justified; and as structurally coordinated when the specified set, operation, quantifiers, and axioms were used together in the justification. For example, P21's statement in Task 2d that 'both elements have to change or be combined' was treated as raw evidence, coded as an active-operand criterion, grouped within prototype-dependent operation recognition, and interpreted as operation-oriented reasoning for that episode. P21's different explanation of the same mapping as a pair-input machine was retained as representation-sensitive evidence. This analytic trace does not license a permanent label for P21.

Frequencies and percentages describe task-level response patterns in this cohort, using 36 as the denominator for diagnostic responses and eight for interviews when relevant. They do not indicate the prevalence of fixed concept-image types, constitute population estimates, or support causal claims. Configuration assignments were made at the episode level and interpreted through cross-task comparisons, not through a numerical cutoff or a single isolated answer.

2.4. Validity of Interpretation

Credibility was strengthened by triangulating written answers, interview explanations, cross-task responses, and reflective notes; examining negative cases; and discussing competing interpretations within the research team. A purposive subset representing response variation was independently reviewed by another researcher, and discrepancies were resolved through discussion of the response context and supporting evidence. An audit trail documented code development, analytic decisions, negative cases, and the formation of the three configurations (Nowell et al., 2017).

3. Results and Discussion

The analysis distinguished three configurations of evoked reasoning: operation-oriented, per-axiom, and structurally coordinated. In operation-oriented episodes, familiar arithmetic actions, symbols, or the visible participation of both operands guided judgments more than the mapping-based definition. In per-axiom episodes, particular requirements were identified or checked, but universal quantification or their joint applicability to one set and operation was not consistently justified. In structurally coordinated episodes, the specified set, operation, quantifiers, and relevant axioms collectively governed classification, proof, or refutation. These are analytic configurations of task-level reasoning, not stable student categories or successive developmental stages.

3.1. Operation-Oriented Concept-Image Configuration

Task 1 of the diagnostic instrument asked students to explain the meaning of a binary operation in their own words, to produce examples of a binary operation, and to evaluate a given rule presented in symbolic and permutation-composition form, as shown in Figure 1.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Figure 1
Diagnostic Task 1: Interpreting the Binary Operation

a. *Explain, in your own words, what is meant by a binary operation.*
b. *Give two examples of a binary operation. Explain why each example can be called a binary operation.*
c. *Given the operation $*$ on the set of real numbers $\mathbb{R}$, defined by the rule*
$$a * b = a, for\ every\ a, b \in R$$
Is $*$ a binary operation on R? Explain your reasoning.
d. *Can the composition of two permutations be called a binary operation? Explain your answer with reference to the domain and codomain of the operation.*

On Task 1a, 10 of the 36 students (27.8%) were able to coordinate all three attributes of a binary operation: two inputs from the same set, a single output, and closure of the result. Seventeen students (47.2%) mentioned combining two elements and closure but had not yet asserted the functional property and uniqueness of the output. Meanwhile, nine students (25.0%) construed a binary operation as an operation involving two numbers or the use of the symbols +, −, ×, and ÷. P07's response, "A binary operation is an operation involving two numbers, such as addition, subtraction, multiplication, and division" (P07, Task 1a), suggests that, in this response, a familiar computational prototype took precedence over understanding a binary operation as a function $S \times S \rightarrow S$. This was also evident when P07 accepted division without first specifying the set on which the operation was to hold. On Task 1b, the dominance of the computational image of operation became even more apparent. Of the 72 examples given, 61 examples (84.7%) were arithmetic operations on sets of numbers, while only 11 examples (15.3%) used non-numeric objects, such as functions, permutations, sets, and logic. Some students also concluded closure solely on the basis of success on a few examples or substitution cases. In those responses, closure was treated as success on selected cases rather than as the universal statement $\forall a, b \in S, a * b \in S$. This finding aligns with Uscanga et al. (2024), who showed that in advanced mathematical contexts, determining the status of a function requires coordinated attention to domain, codomain, the everywhere-defined property, and well-defined. A technique that succeeds on one task form is also not automatically connected to the definition of function when the mathematical context and objects change.

On Task 1c, 22 students rejected the rule $a * b = a$ as a binary operation on R, even though the rule is defined for every pair of real numbers, yields exactly one output, and always returns a result in R. The rejection occurred because b did not appear to affect the outcome. These responses indicate reliance on the visible form of the rule rather than on the defining properties of the mapping. In fact, a binary operation does not require that both inputs appear to be processed or that both equally influence the output. For these episodes, we use the descriptive code active-operand criterion: the assumption that both inputs must be seen to be used, processed, or to exert an observable influence on the output. This criterion is not part of the definition of a binary operation, but it likely arises from an arithmetic prototype in which both operands are typically transformed or combined.

On Task 1d, P07 initially rejected permutation composition as a binary operation because it lacked familiar arithmetic symbols. After a follow-up question directing attention to two permutation inputs and the resulting permutation, P07 accepted composition as a binary operation. This

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

interview response shows that a mapping-based interpretation became accessible under probing in this episode; it does not establish that the interpretation was stable or spontaneously available across other tasks. The contrast suggests that the presentation of the mathematical objects influenced P07's immediate judgment. This interpretation is consistent with Hicks (2024), who described difficulties in recognizing structural relations across familiar and unfamiliar forms.
Task 2 asked students to classify five rules as binary operations or non-examples on a specified set, as shown in Figure 2.

Figure 2
Diagnostic Task 2: Classifying Binary Operation Rules

For each of the following rules, determine whether it defines a binary operation on the given set. Explain your reasoning.
a. $a * b = a - b, on\ Z$
b. $a * b = a/b, on\ Z$
c. Ordinary addition, on {0, 1}.
d. $a * b =$*a (projection), on R.*
e. $a * b = \sqrt{a^2 + b^2}$*, on R*

Task 2 showed that classification success shifted when a rule's structural properties and surface features were in conflict. Thirty-one students (86.1%) recognized $a * b = a - b$ as a binary operation on Z, and 28 students (77.8%) rejected $a * b = a/b$ on Z because the rule is not always defined or its result is not always in Z. Thirty students (83.3%) likewise rejected ordinary addition on {0, 1} because 1 + 1 = 2 does not lie in that set. However, only 19 students (52.8%) accepted $a * b = \sqrt{a^2 + b^2}$*, on* R, and only 14 students (38.9%) accepted the projection operation $a * b =$*a* on R. Twenty-two students rejected the projection operation because b did not appear to be used. P21 explained, "If b isn't included in the calculation, this isn't an operation on two numbers. In my view, both elements have to change or be combined" (P21, Task 2d). In the interview, P21 recognized the same rule as a function when it was presented as a machine that takes a pair (a, b) and outputs the first component, yet still hesitated to call it a binary operation. This episode suggests that recognizing a function $S \times S \rightarrow S$ was not sufficient, in this context, for P21 to apply the definition of a binary operation consistently across representations. This finding is consistent with research on the importance of coordinating domain, codomain, rule, everywhere-definedness, and well-definedness when evaluating functions in advanced mathematical contexts (Uscanga et al., 2024; Uscanga & Cook, 2024).

3.2. Per-Axiom Concept-Image Configuration

Task 3 asked students to examine how closure changes when the underlying set changes, as shown in Figure 3.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Figure 3
Diagnostic Task 3: Closure and Change of Set

Consider the rule

$$a + b = a + b + 1$$

a. Is $*$ closed on Z? Explain in general.
b. Is $*$ closed on the set of even numbers? Explain.
c. Is $*$ closed on the set of odd numbers? Explain.
d. What do your answers show about the relation between a rule and the set on which it is defined?

On Task 3, 34 of the 36 students (94.4%) correctly stated that $a * b = a + b + 1$ is closed on Z. However, 12 students based this conclusion on only a few examples; they had not yet explained in general that for every $a, b \in Z$, the result $a + b + 1$ is also a member of Z. In these responses, checking selected examples replaced a general justification of closure for all pairs in the given set. In other words, students were aware of the "for every" requirement but had not used it consistently in constructing their justification. This finding relates to Uscanga et al. (2024), who showed that students may possess relevant knowledge of a function's definitional properties yet not always use it when facing a different task form or mathematical object. In this study, students were able to show that the operation's result lay in Z for several pairs of numbers, but they did not always connect the operation rule, domain, codomain, and its validity for every pair of elements when determining closure in general.
Task 4 asked students to test and justify associativity, as shown in Figure 4.

Figure 4
Diagnostic Task 4: Associativity

On Z, define

$$a * b = a + b + ab.$$

a. Test the rule on two triples of numbers of your own choosing.
b. Are the results of these two tests sufficient to conclude that $*$ is associative? Why or why not?

On Task 4, 32 students (88.9%) obtained the same result for two selected triples, but only 15 (41.7%) explained why such tests were insufficient and constructed a general argument. Six students interchanged the order of elements, suggesting that, in those episodes, associativity was conflated with commutativity. The contrast between particular checks and universal justification illustrates per-axiom reasoning rather than a general inability to state the definition. After being asked to write the definition, P27 altered the response; this shows activation of relevant definitional knowledge under probing, not evidence of a lasting conceptual change. Hicks (2024) similarly documents difficulties in activating structural relations when familiar procedures dominate mathematical judgments.

Task 5 asked students to determine and justify the identity element of a given operation, as shown in Figure 5.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Figure 5
Diagnostic Task 5: Identity Element

On R, define the operation
$$a * b = a + b - 3.$$
a. Determine the identity element for $*$, if one exists.
b. Explain why that element must work from both the left and the right.

On Task 5, 27 students (75.0%) correctly determined 3 as the identity for $a * b = a + b - 3$, but only 11 students checked both conditions $a * e = a$ and $e * a = a$. A one-sided check could justify the other side only if commutativity of the specified operation had been established and explicitly invoked. Without that justification, the use of 0 or 1 together with a one-sided check indicates a concept image of identity as a fixed arithmetic element. Formally, an identity element e must satisfy $a * e = e * a = a$ for every a in the set; the value of the identity is thus determined by the operation, not intrinsic to the symbol 0 or 1. The dominance of prototypical examples from addition and multiplication reveals a conflict between concept image and concept definition (Tall & Vinner, 1981). This pattern aligns with the finding that the meaning attributed to a binary operation can be influenced by the characteristics of familiar operations (Melhuish, Ellis, et al., 2020) and reinforces the value of conceptual tasks for revealing understanding that is not always visible through procedural answers (Feudel & Unger, 2025).
Task 6 asked students to determine inverses and to relate them to the operation and identity from Task 5, as shown in Figure 6.

Figure 6
Diagnostic Task 6: Inverse Element

Use the operation $*$ from Task 5.
a. Determine the inverse of an element a with respect to $*$.
b. Determine the inverse of 7 and check your answer.
c. Explain the relation among an element a, its inverse, the operation $*$, and the identity element.

On Task 6, 21 students (58.3%) successfully determined the general form of the inverse, $a^{-1} = 6 - a$, while 25 students (69.4%) correctly determined the inverse of 7 as −1. However, some students still used $-a$ or $1/a$. These responses show that the concept image of inverse remained dominated by the opposite under ordinary addition or the reciprocal under ordinary multiplication, rather than by the relation among operation, identity, and inverse.
"The inverse of 7 is −7, because a number plus its opposite equals zero." (P09, Task 6b)
P09's written answer applied ordinary addition and the identity 0, rather than the operation $a * b = a + b - 3$ and the identity 3. In fact, the inverse of 7 must satisfy $7 * 7^{-1} = 3$, so that $7 + 7^{-1} - 3 = 3$, giving $7^{-1} = -1$. In this particular response, the relation $a * a^{-1} = a^{-1} * a = e$ was not spontaneously applied. A similar variation appeared with P31, who could state the formal definitions of identity and inverse but reverted to 0, 1, −a, or 1/a on an application task. Element, operation, identity, and inverse were not always viewed as a single, unified relation. This pattern reveals a conflict between concept image and concept definition, in which arithmetic experience dominates over the relational definition of inverse (Melhuish, Ellis, et al., 2020; Tall

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

& Vinner, 1981). This finding also reinforces the importance of misconception-based conceptual tasks for revealing reasoning that may not be visible through procedural problems (Feudel & Unger, 2025).

3.3. Structurally Coordinated Concept-Image Configuration

Task 7 asked students to define a group, state and relate its required conditions, and give a supporting example, as shown in Figure 7.

Figure 7
Diagnostic Task 7: The Meaning of a Group

a. *In your own words, explain what is meant by a group.*
b. *Write down all the conditions you believe a pair $(G,*)$ must satisfy to be a group.*
c. *Explain the relations among those conditions. Is it sufficient to check them separately, without regard to the set and the operation?*
d. *Give one example of a group and explain why it satisfies your definition.*

On Task 7a, 14 students (38.9%) defined a group as a set equipped with a binary operation satisfying the required axioms. Sixteen students (44.4%) supplied near-formal definitions but omitted a component, most often closure or the universal requirement that each element have an inverse. Six students (16.7%) described a group as a set with certain properties or as a collection of numbers that can be operated on. The responses reveal different degrees of structural coordination in this task. They do not establish that any student possessed a permanently complete or incomplete concept image. Related difficulties in coordinating the components of the group definition were reported by Veith et al. (2022a) and may be interpreted through the distinction between concept image and concept definition (Tall & Vinner, 1981).

On Tasks 7b and 7c, 30 students named at least three group axioms, but only 12 explained that all requirements must hold jointly for the same set and operation. Eighteen presented the conditions as an independent checklist, whereas six included commutativity as mandatory. In these task episodes, the abelian-group prototype sometimes displaced the general group definition. The distinction between listing properties and coordinating them is central to the per-axiom configuration, but it should not be read as a fixed attribute of the respondents. The pattern is compatible with research on familiar-operation prototypes (Melhuish, Ellis, et al., 2020) and supports contrasting groups with non-groups and abelian with non-abelian groups in conceptual tasks (Feudel & Unger, 2025).

On Task 7d, classification of $(Z,\times)$, $(N,+)$, and $(2Z,+)$ suggested that the definition did not govern every decision in this task. For $(Z,\times)$, some students stopped after finding closure, associativity, identity, and inverse for a few elements without activating the "every element" quantifier. For $(N,+)$, the presence of 0 was regarded as sufficient to accept the structure. Meanwhile, 14 students rejected $(2Z,+)$ because it does not contain every integer, or because the inverse of 2 was assumed to be 1/2. The latter response applied the multiplicative inverse, whereas the operation in use was addition, so the inverse of 2 is -2. These responses illustrate that definition-governed classification was not consistently evoked in this task and could still be overridden by an arithmetic prototype. This finding aligns with the difficulties in coordinating

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

operation, identity, inverse, and quantifiers reported by Veith et al. (2022a), and it supports the use of non-examples and misconception-based distractors (Feudel & Unger, 2025).
Task 8 asked students to prove or refute two candidate group structures and to compare their strategies, as shown in Figure 8.

Figure 8
Diagnostic Task 8: Proving and Refuting

a. *On $\mathbb{Z}$, define $a * b = a + b + 1$. Prove or refute that $(Z,*)$ is a group.*
b. *On $\mathbb{Z}$, define $a \diamond b = a - b$. Prove or refute that $(Z,\diamond)$ is a group.*
c. *Compare how you solved parts (a) and (b). When is the definition used to prove, and when is a single counterexample sufficient to refute?*

On Task 8a, 18 students (50.0%) produced a complete or nearly complete proof that $(Z,*)$, with $a * b = a + b + 1$, is a group. Ten students correctly obtained some components but did not prove associativity in general or did not check identity and inverse from both sides. Eight students denied the structure because its identity was not 0 or because its operation was not ordinary addition. In successful proof episodes, all relevant axioms were coordinated; in other responses, familiar operation and identity prototypes influenced the justification. This is consistent with Melhuish et al. (2020), who showed that students' understanding of binary operations can be influenced by the characteristics of familiar operations, and with Veith et al. (2022a), who identified students' difficulty in coordinating operation, identity, inverse, and other axioms when evaluating a group structure.
On Task 8b, 31 students (86.1%) correctly stated that $(Z,\diamond)$, with $a \diamond b = a - b$, is not a group. However, only 17 students used a valid counterexample for associativity or demonstrated the absence of a two-sided identity. Nine students rejected it on the grounds that it is not commutative, even though commutativity is not a group axiom. These episodes included valid axiom-based refutations as well as invalid rejections based on a property that is not required of groups. The contrast concerns the quality of task-specific justifications and does not establish an individual developmental progression. The tendency to conflate commutativity with the group axioms likewise reflects an understanding of axioms as a list of operation properties that has not yet been conceptually coordinated (Veith et al., 2022a). Tasks that ask students to compare valid and invalid justifications are therefore important for revealing whether their decisions are genuinely governed by the definition or merely by resemblance to familiar group examples (Feudel & Unger, 2025).
The operation-table representation provided further evidence of task-sensitive reasoning. All eight interviewees identified closure in the table, six found a two-sided identity, and five identified an inverse for each element; only three adequately justified associativity. Four participants inferred that the table defined a group because each symbol occurred exactly once in every row and column. Such Latin-square regularity does not, by itself, establish associativity. P03 used the defining properties consistently across formula-based and tabular tasks, whereas P21's justification was less consistent in the table context than with a symbolic rule. These comparisons illustrate changes in the reasoning evoked by a representation rather than differences in permanent conceptual ability. Related variation across abstract-algebra concept items has been documented by Veith et al. (2025).

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Across tasks, the analytic configurations differed in how strongly the formal definition governed mathematical judgments. Operation-oriented episodes privileged arithmetic or visible-operand features; per-axiom episodes activated selected axioms without consistently coordinating their quantifiers and shared structure; and structurally coordinated episodes used the definition to justify classification, proof, or refutation. The same participant could evoke more than one configuration, as the contrasting responses of P03, P21, and P31 illustrate. The subsection sequence is therefore an organizational device, not a developmental pathway. The findings concern situated concept-image activation (Tall & Vinner, 1981), and no individual developmental trajectory was investigated.

Table 2 synthesizes the task-specific evidence supporting the three configurations. The configuration assigned to a row is an interpretation of reasoning in that context, not a classification of the students who gave the responses.

Table 2.

Cross-task Summary of Responses, Difficulties, and Concept-Image Interpretations

Tasks	Student responses	Types of difficulties	Interpretation	Concept-image configuration
Task 1. Binary-operation definition and examples	10/36 stated the key defining attributes; 22/36 rejected the projection rule $a * b = a$ on $\mathbb{R}$.	Arithmetic prototype; active-operand criterion.	A familiar computational image overrode the functional definition in specific responses.	Operation-oriented
Task 2. Classification of rules	22/36 rejected projection; 31/36 accepted subtraction on $\mathbb{Z}$; 30/36 rejected addition on $\{0,1\}$.	Attention to visible operands varied despite successful closure checking on familiar examples.	Definitional control differed across tasks and representations.	Operation-oriented; task-sensitive
Task 3. Closure with a change of set	34/36 judged the rule $a * b = a + b + 1$ closed on $\mathbb{Z}$; 12 based justifications on a few examples.	Examples used in place of a universal argument.	The role of "for every pair" was not consistently expressed.	Per-axiom
Task 4. Associativity	32/36 obtained equal results on selected triples; only 15 constructed a general argument.	Example-based verification conflated with proof; some mixed associativity and commutativity.	Correct checks on particular triples were not always generalized.	Per-axiom
Task 5. Identity	27/36 identified 3 for $a * b = a + b - 3$ on $\mathbb{R}$; 11 checked both sides.	Identity treated as 0 or 1 by prototype; unilateral verification.	The identity was sometimes detached from the defined operation.	Per-axiom
Task 6. Inverse	21/36 derived $6 - a$; 25/36 identified -1 as inverse of 7; P09 used -7.	Additive opposite or reciprocal substituted for operation-relative inverse.	Identity and inverse were not always coordinated with the specified operation.	Per-axiom

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Tasks	Student responses	Types of difficulties	Interpretation	Concept-image configuration
Task 7. Group definition and classification	14/36 stated the complete structural definition; 12/36 coordinated axioms in Tasks 7b–c.	Axiom checklist; missing quantifiers; unnecessary commutativity.	A formally stated definition did not always govern a new classification.	Per-axiom; structurally coordinated where justified
Task 8. Proof and refutation	18/36 gave a complete or near-complete group proof; 17/36 offered valid grounds to refute the second structure.	Incomplete universal proof; invalid rejection based on noncommutativity.	Successful judgments were not always supported by relevant axioms.	Per-axiom; structurally coordinated where justified
Interview: operation table	8/8 recognized closure; 6/8 found two-sided identity; 5/8 found all inverses; 3/8 justified associativity.	Latin-square regularity mistaken for sufficient evidence of group structure.	Definition-based reasoning could weaken in table representation.	Representation-sensitive across configurations

Note. Diagnostic-task frequencies refer to n = 36; the operation-table evidence is from n = 8 interviewees. Configuration labels are episode-specific

Table 2 summarizes the observable responses, associated difficulties, bounded interpretations, and configurations across tasks. Building on the distinction between concept image and concept definition (Tall & Vinner, 1981), this analysis extends prior accounts of difficulties with binary operations and group axioms (Melhuish, Lew, et al., 2020; Veith et al., 2022b, 2022a) by connecting operation recognition, separate axiom checking, and coordinated structural justification within one interpretive account. Its contribution is not a universal learner taxonomy. Rather, it identifies how the control exerted by a formal definition may vary across mathematically related demands and representations. A correct final answer alone does not establish such control when reasoning relies on a few examples, familiar prototypes, or incomplete justification.

Within DDR prospective analysis, the three configurations suggest specific but provisional didactical needs. Operation-oriented episodes point to contrasting arithmetic with non-arithmetic and projection operations, explicitly considering domain, codomain, totality, and uniqueness. Per-axiom episodes motivate comparisons of particular examples with universal arguments, two-sided identity and inverse checks, and critique of incomplete proofs. Structurally coordinated episodes suggest activities connecting formulas, boundary cases, counterexamples, and Cayley tables while requiring explicit coordination of all relevant axioms. These implications are consistent with guided reinvention of the group concept (Larsen, 2013), but they are didactical hypotheses, not findings about the effectiveness of an implemented design. As synthesized in Table 3, subsequent DDR work would need to translate them into a hypothetical didactical design, analyze implementation metapedadidactically, and revisit initial conjectures retrospectively.

Table 3 integrates the three configurations with their supporting evidence, interpretive limits, and provisional didactical implications.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Table 3.

Synthesis of Concept-Image Configurations and Prospective Didactical Needs

Configuration	Reported evidence	Bounded interpretation	Prospective didactical need
Operation-oriented	Projection a * b = a was rejected by 22/36 in Task 1c; P21 required both operands to appear active in Task 2d.	In these episodes, an arithmetic prototype or visible-operand criterion governed recognition. This does not mean each participant reasons this way on every task.	Contrast numerical and nonnumerical operations; foreground domain, codomain, totality, and uniqueness of a binary operation.
Per-axiom	Only 11/36 explicitly checked both sides of the identity; 12/36 coordinated the group conditions in Tasks 7b–c.	Individual axioms were sometimes identified but not joined through common set, operation, quantifiers, and mutual dependence.	Use universal-versus-example tasks, two-sided tests, proof critiques, and axiom coordination activities.
Structurally coordinated	18/36 produced complete or nearly complete proof on Task 8a; only 3/8 interviewees adequately justified associativity from the operation-table context.	Definition-governed reasoning was observable but need not transfer reliably to another representation. The figures describe different tasks and samples.	Compare formulas, examples, counterexamples, and Cayley tables; require explicit explanation of why all group requirements hold.

Note. These implications are prospective didactical hypotheses; the study did not test an instructional intervention.

4. Conclusion

The analysis distinguished three context-sensitive configurations of prospective mathematics teachers' concept images of group structure: operation-oriented, per-axiom, and structurally coordinated reasoning. The first describes episodes governed by arithmetic or notational prototypes; the second, episodes in which axioms are recognized but insufficiently connected through common quantifiers, sets, and operations; and the third, episodes in which formal conditions jointly support classification, proof, or refutation. Each configuration describes reasoning evoked in particular tasks, not a stable type of student or a developmental stage.

The findings suggest that recalling a group definition does not guarantee its activation in unfamiliar situations. Apparent success in one representation may coexist with incomplete justification or prototype-based reasoning in another. Accordingly, responses should be interpreted through their justifications and cross-task contrasts rather than through final-answer correctness alone. No longitudinal evidence of movement between configurations was collected, so no individual developmental progression is claimed.

As summarized in Table 3, the analysis contributes a task-sensitive account linking binary-operation recognition, the use of individual axioms, and structural coordination. For DDR prospective analysis, the findings support preliminary didactical anticipations involving varied objects and representations, careful attention to universal quantifiers, and comparisons of valid proofs with counterexamples. The evidence comes from one cohort and eight interviews; transferability therefore depends on the similarity of other educational contexts, and the configurations should not be generalized as population categories. Future work should design and

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

test didactical situations informed by these findings, conduct metapedadidactic analysis, and undertake retrospective analysis before drawing conclusions about instructional effectiveness.

5. Acknowledgments

The authors thank the Institute for Research and Community Service (LPPM) of Universitas Pendidikan Ganesha and FORUM LPPM ALPTKNI for supporting this study through the 2026 RLKI scheme (Contract No. 251/UN48.16/PT.01.01/2026). We also thank the collaborating researchers at Universitas Pendidikan Indonesia and all study participants.

6. References

Agustyaningrum, N., Sari, R. N., Abadi, A. M., & Mahmudi, A. (2021). Dominant Factors That Cause Students' Difficulties in Learning Abstract Algebra: A Case Study at a University in Indonesia. *International Journal of Instruction*, *14*(1), 847–866. https://doi.org/10.29333/iji.2021.14151a

Aksu, Z., Konyalıoğlu, A. C., Kul, Ü., & Kul, Ü. (2018). Pre-Service Mathematics Teachers' Conceptual Knowledge of Binary Operation. *Khazar Journal of Humanities and Social Sciences*, *21*(4), 55–70. https://doi.org/10.5782/2223-2621.2018.21.4.55

Biehler, R., Durand-Guerrier, V., & Trigueros, M. (2024). New trends in didactic research in university mathematics education. *ZDM - Mathematics Education*, *56*(7), 1345–1360. https://doi.org/10.1007/s11858-024-01643-2

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory. *Educational Studies in Mathematics*, *27*(3), 267–305. https://doi.org/10.1007/BF01273732

Feudel, F., & Unger, A. (2025). Development and implementation of concept-test questions in abstract algebra. *International Journal of Mathematical Education in Science and Technology*, *56*(5), 869–898. https://doi.org/10.1080/0020739X.2023.2297002

Hicks, M. D. (2024). "I'll just try to mimic that": an exploration of students' analogical structure creation in abstract algebra. *Educational Studies in Mathematics*, *117*(2), 303–327. https://doi.org/10.1007/s10649-024-10345-1

Larsen, S. P. (2013). A local instructional theory for the guided reinvention of the group and isomorphism concepts. *The Journal of Mathematical Behavior*, *32*(4), 712–725. https://doi.org/10.1016/j.jmathb.2013.04.006

Melhuish, K., Ellis, B., & Hicks, M. D. (2020). Group theory students' perceptions of binary operation. *Educational Studies in Mathematics*, *103*(1), 63–81. https://doi.org/10.1007/s10649-019-09925-3

Melhuish, K., Lew, K., Hicks, M. D., & Kandasamy, S. S. (2020). Abstract algebra students' evoked concept images for functions and homomorphisms. *The Journal of Mathematical Behavior*, *60*, 100806. https://doi.org/10.1016/j.jmathb.2020.100806

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*(1), 1609406917733847.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Suryadi, D. (2013). Didactical design research (DDR) to improve the teaching of mathematics. *Far East Journal of Mathematical Education*, *10*(1), 91–107. https://doi.org/10.17654/0973563113006

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

Uscanga, R., & Cook, J. P. (2024). Analyzing the Structure of the Non-examples in the Instructional Example Space for Function in Abstract Algebra. *International Journal of Research in Undergraduate Mathematics Education*, *10*(1), 7–33. https://doi.org/10.1007/s40753-022-00166-z

Uscanga, R., Melhuish, K., & Cook, J. P. (2024). Students' techniques for approaching defining properties of functions. *Educational Studies in Mathematics*, *117*(3), 457–484. https://doi.org/10.1007/s10649-024-10344-2

Veith, J. M., Bitzenbauer, P., & Girnat, B. (2022a). Assessing Learners' Conceptual Understanding of Introductory Group Theory Using the CI2GT: Development and Analysis of a Concept Inventory. In *Education Sciences* (Vol. 12, Issue 6, p. 376). https://doi.org/10.3390/educsci12060376

Veith, J. M., Bitzenbauer, P., & Girnat, B. (2022b). Exploring Learning Difficulties in Abstract Algebra: The Case of Group Theory. In *Education Sciences* (Vol. 12, Issue 8, p. 516). https://doi.org/10.3390/educsci12080516

Veith, J. M., Girnat, B., Krause, M., Straulino, M., Becker, E., Becher, R., & Bitzenbauer, P. (2025). Consistent Responses on Concept Items Reflect Confidence, and Vice Versa. *International Journal of Science and Mathematics Education*, *23*(8), 3903–3928. https://doi.org/10.1007/s10763-025-10605-3

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia