International Journal of
Humanities, Education, and Social Sciences

e-ISSN : 3026-0892
p-ISSN : 3026-1422

Index: Harvard, Boston, Sydney, Dimensions, Lens, Scilit, Semantic, Google, etc

https://doi.org/10.58578/IJHESS.v4i3.12670

An Analysis of Students' Mathematical Reasoning in Solving Indefinite Integral Problems Based on Lithner's Framework

Khadijah
Universitas Negeri Makassar, Indonesia
khadijah@unm.ac.id

Submitted:	Revised:	Accepted:	Published:
Oct 4, 2026	Nov 2, 2026	Nov 13, 2026	Nov 18, 2026

Abstract

Correct solutions to indefinite integral problems can result from applying routine procedures and therefore provide limited evidence of the reasoning students used. Drawing on Lithner's framework, this study examined the types and components of mathematical reasoning evident in students' written solutions and explored how errors related to reasoning type. A descriptive–interpretive qualitative document analysis was conducted on 20 scanned handwritten scripts, comprising 19 scripts covering 13 odd-numbered antiderivative items and one definite-integral script used for comparison. Qualitative content analysis involved deductive coding of strategy choice, implementation, argumentation, and mathematical foundation, followed by classification into four levels (A–D) according to the dominant reasoning pattern and categorization of errors. All 19 antiderivative scripts exhibited algorithmic reasoning. Eleven showed procedural steps alone, five stated a general rule, and three provided reasons or specified conditions for a rule's validity. No script met all three criteria for creative reasoning or showed verification by differentiation. Errors were associated with the reuse of procedural templates across items, copying slips, and uncritical interpretations of notation. The findings show that the routine tasks examined elicited imitative reasoning and underscore that creative reasoning requires strategic novelty alongside mathematical grounding. The study provides coding indicators and reasoning levels for analyzing written solutions and identifies ways to design calculus assessments that require students

IJHESS Journal is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

to verify results, justify strategies, and test the conditions under which rules apply. The analysis was limited to written evidence and relied on a single coder.

Keywords: Mathematical Reasoning; Imitative Reasoning; Creative Reasoning; Indefinite Integrals; Calculus Assessment

INTRODUCTION

Calculus is a gateway course for the sciences, engineering, and many other study programmes, and integration is among the topics students meet most often in it (Jones, 2013; Tatira, 2025). In mathematics learning more generally, learning often means carrying out demonstrated procedures: recognising a task type, applying a standard routine, and obtaining a single answer. This pattern is natural at the beginning, but it raises a more basic question: do students also *reason*, choosing and justifying their steps on the basis of mathematical properties, or do they merely follow procedures? The topic that provides the context for this study is the indefinite integral, that is, finding the general antiderivative F(x) + C. For power functions and their sums, antiderivatives follow from the power rule and linearity, and this skill underlies the evaluation of definite integrals. The antiderivative is also one of the ways in which students make sense of the integral, alongside area and adding up pieces (Jones, 2013, 2015b). Because the basic tasks can all be solved algorithmically, this topic is a sharp place to tell carrying out a procedure from reasoning.

Several lines of evidence support this concern. Bergqvist (2007) found that about 70% of the tasks in university mathematics exams can be solved with imitative reasoning, and that 15 exams could be passed without any creative reasoning. Jäder et al. (Jäder, Lithner and Sidenvall, 2020) analysed 5,738 tasks from textbooks in twelve countries and found that most could be solved by following a template, while few required a solution to be constructed. For integration specifically, Mahir (2009) reported that students who had completed a year of calculus did not yet have a satisfactory conceptual understanding of integration. Since Orton (1983) examined students' errors in integration, research has moved towards conceptions such as symbolic forms (Jones, 2013), Riemann sums (Sealey, 2014), and applications (Tatira, 2025), and still warns that students can be fluent in symbolic techniques without matching conceptual understanding (Tatira, 2025).

The actual condition behind this study can be seen in the written work of 20 students collected for this purpose. They solved routine antiderivative items and, on first inspection,

produced very uniform sequences of steps and mostly correct answers. Uniformity and accuracy show the *product* of the work, but not the reasons behind the choice of steps. The main problem, therefore, is that procedural accuracy does not reveal the kind of reasoning that underlies it. A correct and confident answer can rest on remembered procedures and surface cues (Altindis, 2025). As a result, it is unclear how students decide on a strategy and justify a result when solving indefinite integrals.

The consequences of this problem are long-term. According to Lithner (2008), reasoning based on given algorithms tends to lead to rote learning, with negative consequences for conceptual understanding (Jonsson, Granberg and Lithner, 2020). Experiments show that practising with tasks that demand creative reasoning gives better outcomes than practising with given algorithms (Jonsson, Norqvist, Liljekvist and Lithner, 2014; Jonsson, Granberg and Lithner, 2020), and university students' learning difficulties have been linked to imitative reasoning (Lithner, 2011). When students' reasoning never becomes visible, errors that stem from shallow reasoning are also hard to catch early. The main cause lies in tasks and assessment. The type of task influences the type of reasoning students use (Boesen, Lithner and Palm, 2010), and calculus tasks, exam questions, and textbooks typically supply solution templates (Lithner, 2004; Bergqvist, 2007; Jäder, Lithner and Sidenvall, 2020). When the method is already given, students' reasoning tends to become imitative (Jonsson, Granberg and Lithner, 2020). The opportunities for creative reasoning in undergraduate calculus assessment have been questioned (Mac an Bhaird, Nolan, O'Shea and Pfeiffer, 2017), and Calculus I final exams have been characterised systematically (Tallman, Carlson, Bressoud and Pearson, 2016).

Previous research can be grouped into four strands. The first analyses tasks, exams, and textbooks, or students working on them, with Lithner's framework (Lithner, 2000, 2003, 2004; Bergqvist, 2007; Jäder, Lithner and Sidenvall, 2020). The second consists of experiments and eye-tracking studies on the effects of creative and algorithmic reasoning (Jonsson, Norqvist, Liljekvist and Lithner, 2014; Norqvist, Jonsson, Lithner, Qwillbard, *et al.*, 2019). The third studies the understanding of integrals (Orton, 1983; Mahir, 2009; Jones, 2013, 2015a, 2015b; Sealey, 2014; Tatira, 2025). The fourth applies reasoning frameworks to students' work, namely on functions (Altindis, 2025) and on visual reasoning about integrals (Sholihah and Asyhar, 2023). Other work examines reasoning in students' textbook task solving and non-routine tasks (Bergqvist, Lithner and Sumpter, 2008; Sidenvall, Lithner and

Jäder, 2015; Jäder, Sidenvall and Sumpter, 2017) and in teachers' presentations (Bergqvist and Lithner, 2012).

These findings do not all point in the same direction. Mahir (2009) found that students who understood the concept of integration also succeeded procedurally, whereas Tatira (2025) cautions that fluency with techniques can be separated from understanding. Jones (2015a) showed that area and antiderivative conceptions differ in productivity between mathematical and applied contexts. Such differences indicate that procedural performance alone is not enough to infer the quality of understanding, and that the type of task and its context help determine what becomes visible.

From the literature we reviewed, the following research gap appears. Studies on integrals concentrate on conceptions of the definite integral (area, Riemann sums, applications), whereas symbolic antidifferentiation is rarely examined in terms of reasoning type. Studies based on Lithner's framework mostly analyse tasks, exams, and textbooks rather than real solutions written by students. Three gaps follow: a theoretical gap (Lithner's reasoning framework has not been related to the indefinite integral), a methodological gap (there are no explicit coding criteria for reasoning levels in the written traces of real solutions), and an empirical gap (there is no evidence from students' work on routine integral tasks). This judgement rests on the literature we reviewed, not on a systematic review. The limitations of earlier studies reinforce the gap. Task analyses show what is *demanded*, not how students *solve* the tasks. Altindis (2025) analysed three of eight students in depth on function tasks, and Sholihah and Asyhar (2023) involved three students, so conclusions about gender differences cannot be generalised.

The theoretical foundation is the framework of Lithner (2008). Lithner views task solving as a sequence of four steps: a problematic task is met, a strategy is chosen, the strategy is implemented, and a conclusion is reached (Lithner, 2008). Altindis (2025) networked this framework with covariational reasoning (Carlson, Jacobs, Coe, Larsen, *et al.*, 2002). Argumentation may be *predictive*, supporting the choice of strategy, or *verificative*, explaining why the implementation and conclusion are true or plausible (Jonsson, Granberg and Lithner, 2020). *Imitative* reasoning comprises memorised reasoning (recalling a full answer) and algorithmic reasoning (recalling a procedure so that only its execution remains), which may be familiar, delimiting, or guided by a text or another person (Lithner, 2008). *Creative* reasoning must meet three criteria: the strategy or reasoning sequence is new or re-created

by the reasoner, the arguments are plausible, and the arguments are anchored in intrinsic mathematical properties (Lithner, 2008; Jonsson, Granberg and Lithner, 2020). Creativity here means new to the reasoner, not extraordinary (Jonsson, Norqvist, Liljekvist and Lithner, 2014), and intrinsic properties are essential to the concept, whereas surface properties are not (Lithner, 2008; Altindis, 2025). Other frameworks view reasoning as conjecturing, generalising, and justifying (Jeannotte and Kieran, 2017; Mata-Pereira and da Ponte, 2017), and rubrics that assess explanation and justification can form progression scales (Callingham and Siemon, 2021). The relation between procedural and conceptual knowledge has been reconsidered conceptually (Baroody, Feil and Johnson, 2007) and empirically for university students (Engelbrecht, Harding and Potgieter, 2005), and principles for designing tasks that elicit creative reasoning have been formulated (Lithner, 2017).

The novelty of this study lies in three aspects. Theoretically, it relates Lithner's framework to symbolic antidifferentiation, a topic seldom examined in terms of reasoning. Methodologically, it proposes explicit coding indicators and four levels for the written traces of solutions. Empirically, it presents evidence from the real work of 20 students rather than from tasks or textbooks. Accordingly, the focus of this study is mathematical reasoning according to Lithner's framework in the symbolic solution of indefinite integrals. The objects of the study are 20 handwritten student scripts: 19 antiderivative scripts as the main data and one definite-integral script as a comparison. The aims are to describe the types of reasoning visible in the written traces, to describe which reasoning components are visible and which are not, and to explain how errors and differences in interpreting the items relate to reasoning type. These aims are expressed in three research questions: (RQ1) Which types of Lithner's reasoning are visible in the written traces of indefinite-integral solutions? (RQ2) Which reasoning components (strategy choice, implementation, predictive and verificative argumentation) are visible and which are not, and how do they vary between students? (RQ3) How do errors and differences in interpreting the items relate to reasoning type?

METHODS

1. Type of Research

This study is a descriptive-interpretive qualitative document study. The documents analysed are students' written solutions. They were interpreted deductively with the reasoning framework of Lithner (2008), leaving room for categories that emerged from the data.

2. Research Design

The design consisted of four stages: (1) collecting the scanned scripts and anonymising them; (2) transcribing each item as a sequence of written steps; (3) coding each sequence with the indicators in Table 2 and assigning every script a dominant level; and (4) grouping errors and interpretive variants and relating them to the research questions. All coding was done by one researcher.

3. Participants and Sampling Technique

The data came from 20 students taking a basic mathematics course, who are coded M01 to M20. Nineteen scripts contain antiderivatives (indefinite integrals) and one contains three definite-integral items. The latter was analysed separately as a comparison and is not part of the level classification. The scripts are those that were available to the researcher; the sample is not random and is not intended for generalisation. No demographic information was collected, and the completeness of the work varies: not every script contains all 13 items.

All indefinite-integral scripts answer the same instruction, namely to find the general antiderivative $F(x) + C$ for the odd-numbered items 1 to 25 (Table 1). The task is routine: every item can be solved with the power rule (Equation 1), the linearity property (Equation 2), and two basic trigonometric antiderivatives, and a result can be checked by differentiation (Equation 3). In Lithner's framework, a task of this kind invites familiar algorithmic reasoning.

Table 1. Items Analysed

No.	Item	Salient demand
1	$f(x) = 5$	Antiderivative of a constant
3	$f(x) = x^2 + \pi$	The constant π as a term, not a variable
5	$f(x) = x^{5/4}$	Fractional exponent, $r + 1 = 9/4$
7	$f(x) = 1/\sqrt[3]{(x^2)}$	Converting a root to a negative fractional exponent
9	$f(x) = x^2 - x$	Linearity
11	$f(x) = 4x^5 - x^3$	Linearity and coefficients
13	$f(x) = 27x^7 + 3x^5 - 45x^3 + \sqrt{2}x$	Many terms, irrational coefficient, meaning of $\sqrt{2}x$
15	$f(x) = 3/x^2 - 2/x^3$	Negative exponents
17	$f(x) = (4x^6 + 3x^4)/x^3$	Simplifying before integrating
19	$\int(x^2 + x)\,dx$	Linearity
21	$\int(x + 1)^2\,dx$	Expanding the square before integrating
23	$\int(z^2 + 1)^2/\sqrt{z}\,dz$	Expanding, splitting terms, fractional exponents
25	$\int(\sin\theta - \cos\theta)\,d\theta$	Basic trigonometric antiderivatives

$$\int x^r \, dx = \frac{x^{r+1}}{r+1} + C, \; r \neq -1 \qquad (1)$$

$$\int [a\, f(x) + b\, g(x)] \, dx = a\int f(x) \, dx + b\int g(x) \, dx \qquad (2)$$

$$F'(x) = f(x) \implies \int f(x) \, dx = F(x) + C \qquad (3)$$

4. Instruments and Data Collection

The data collection technique was documentation: the students' handwritten work was scanned, and the scans were the only data source. The research instrument was a coding sheet derived from the structure and criteria of Lithner (2008) (Table 2). Student identities were replaced by codes M01 to M20, and the figure excerpts were chosen so that no names or student numbers appear.

5. Data Analysis Technique

The data were analysed with qualitative content analysis and deductive coding. Reasoning types follow Lithner (2008): memorised, algorithmic, and local or global creative; for this topic we expect algorithmic reasoning to appear as applying the power rule and linearity in sequence, and creative reasoning as rewriting an item for a stated algebraic reason or verifying by differentiation (our operationalisation, not Lithner's own classification).

Each sequence of steps was coded with the indicators in Table 2, which derive from the structure and criteria of Lithner (2008). Each script was then given a **dominant level** based on its written traces, and errors and interpretive variants were grouped into procedural errors, notation errors, and item interpretation.

Table 2. Coding Indicators for Written Traces

Component	Indicator in the written trace	Example
Strategy choice	The strategy used and whether it is stated	Power rule term by term; converting a root to a fractional power
Implementation	Order and correctness of steps and intermediate forms	$x^{r+1}/(r+1)$ computed term by term
Predictive argument	A reason or condition accompanying a step	"because π is a constant"; "as long as $n \neq -1$"
Verificative argument	Checking the result or commenting on its plausibility	Differentiating F to check F' = f
Mathematical foundation	Reference to intrinsic rather than surface properties	Conditions of the rule vs. appearance patterns
Novelty	A strategy that is new or re-created by the reasoner	Deriving the rule oneself

The dominant levels are defined as follows. **Level A (plain algorithmic)**: procedural steps with no statement of a rule or a reason. **Level B (rule-stating algorithmic)**: a general rule (for example the power-rule formula) is written without a reason for applying it. **Level C (reasoned algorithmic)**: steps are accompanied by a reason or by the conditions under which the rule holds. **Level D (creative, local or global)**: a strategy that is new to the reasoner, supported by plausible arguments anchored in intrinsic properties, following the three criteria of Lithner (2008).

Levels A to D are our operationalisation and not standard Lithner categories. Coding was carried out by one researcher without a second coder, so inter-coder reliability has not been tested. To give readers the chance to check the interpretation, evidence is presented as excerpts of the original work with detailed description. Using excerpts of student work in publication requires the consent of the students and the instructor.

RESULTS

1. Main Findings

Distribution of reasoning levels. No script reached the creative level. All 19 antiderivative scripts showed algorithmic reasoning, from plain to reasoned (Table 3). Purely memorised reasoning was not found, because every script showed steps. The classification is holistic and based mainly on the first pages of each script, so the distribution is indicative.

Table 3. Dominant Reasoning Level in the 19 Antiderivative Scripts

Level	n (%)	Scripts	Trace
A, plain algorithmic	11 (58%)	M02, M03, M04, M06, M09, M10, M13, M14, M15, M16, M18	Routine steps, no rule or reason
B, rule-stating algorithmic	5 (26%)	M01, M05, M08, M12, M19	General formula written
C, reasoned algorithmic	3 (16%)	M07, M11, M17	Reason, condition, or rule parameter given
D, creative	0	–	Not found

Familiar algorithmic reasoning and similarity of steps. All scripts chose the same strategy: the power rule term by term with linearity. According to Lithner (Lithner, 2008), in algorithmic reasoning the strategy choice is only recalling the algorithm and the implementation only running its steps, so no argument is needed. This is visible in item 9: three scripts wrote almost identical lines, $f(x) \rightarrow \int(x^2 - x)dx = \int x^2 dx - \int x\,dx$, then $x^{2+1}/(2+1) - x^{1+1}/(1+1) + C$ (Figure 1).

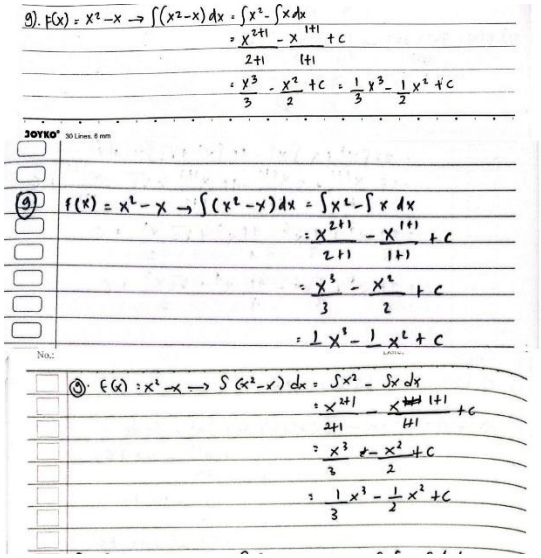


Figure 1. Item 9 in Three Different Scripts (Top to Bottom: M02, M04, M06)

Similar patterns appear in item 23, the longest item. The intermediate form $z^{4-1/2} + 2z^{2-1/2} + z^{-1/2}$ and the final form $(2/9)z^4\sqrt{z} + (4/5)z^2\sqrt{z} + 2\sqrt{z}$ appear in at least four scripts (M02, M06, M15, M18; Figure 2). This pattern is consistent with guided algorithmic reasoning, in which the algorithm is supplied by an instructor's example, a textbook, or a peer (Lithner, 2008). However, the written traces cannot distinguish this from the coincidence of a standard procedure, so we do not conclude that copying occurred.

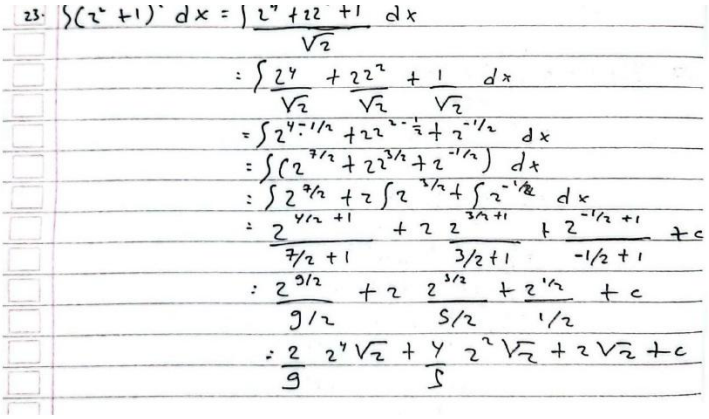


Figure 2. Chain of Steps for Item 23 in M18: Expansion, Splitting Terms, Power Rule, and Writing the Final Form

Changing the form of the item. Items 7, 17, 21, and 23 require a change of form before the power rule can be used: a root becomes a fractional power, the numerator is divided by the denominator, or a square is expanded. The scripts that attempted these items did so, except for item 23, which M07 crossed out and abandoned. The decision is a strategy choice in Lithner's sense, but the techniques are commonly taught and therefore familiar, and they were seldom accompanied by a written reason. Item 7 requires the following change of form before the power rule is used (Equation 4):

$$\frac{1}{\sqrt[3]{x^2}} = x^{-2/3} \implies \int x^{-2/3}\,dx = \frac{x^{1/3}}{\frac{1}{3}} + C = 3x^{1/3} + C \qquad (4)$$

M07 stands out for writing out the structure of the computation: the general rule, then "for n = −2/3, n + 1 = −2/3 + 1 = 1/3" (Figure 3). Computing r + 1 openly anchors the result in the parameters of the rule, but the strategy choice itself remains a recalled one.

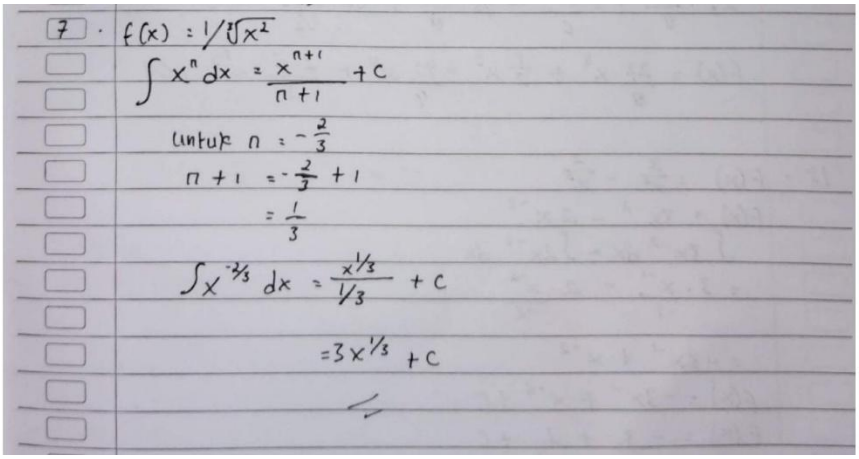


Figure 3. M07, Item 7: General Rule and Explicit Computation of n + 1

Level B: stating the general rule. Five scripts (M01, M05, M08, M12, M19) wrote a general rule, for example $\int a\,dx = ax + C$ or $\int x^n dx = x^{n+1}/(n+1) + C$. After solving item 5, M12 added in coloured ink the formula $\int x^r dx = (1/(r+1))x^{r+1}$ next to the case 5/4 (Figure 4). This move from a case to a rule approaches generalisation (Jeannotte and Kieran, 2017; Mata-Pereira and da Ponte, 2017). However, there was no reason why the rule is true and no condition of validity, so the formula remains a memorised one.

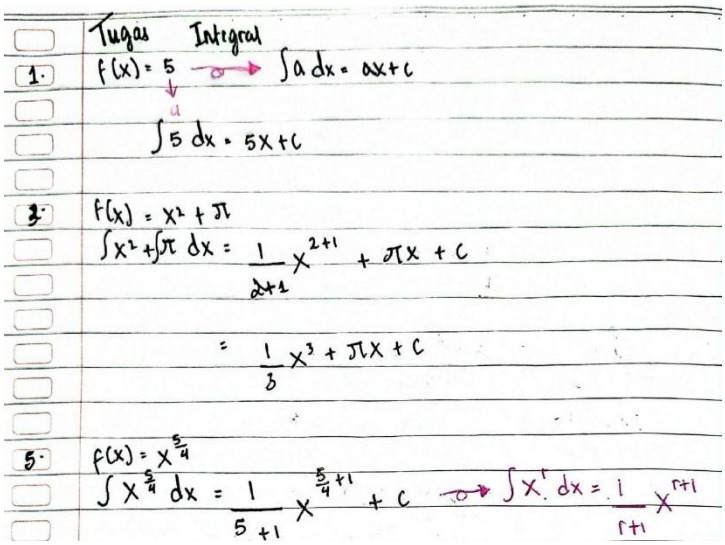


Figure 4. M12: Statement of the General Rule for a Constant (Item 1) and the Power Rule Added after Item 5

Level C: reasons, conditions, and rule parameters. Three scripts accompanied their steps with reasons. M11 wrote the basic formula with the condition "as long as $n \neq -1$", then "substitute $n = 5/4$" (Figure 5), and used separate constants C_1 and C_2 for each term before combining them. M17 wrote "$\int \pi \, dx = \pi x$ (because π is a constant)" and labelled "for the first term $c = 4$, $n = 5$" and "for the second term $c = 1$, $n = 3$" (Figure 6). M07 also belongs at this level (Figure 3).

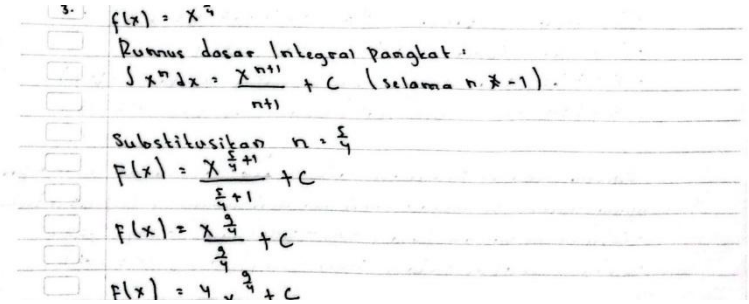


Figure 5. M11, Item 5: Basic Formula with the Condition $n \neq -1$ and Substitution $n = 5/4$

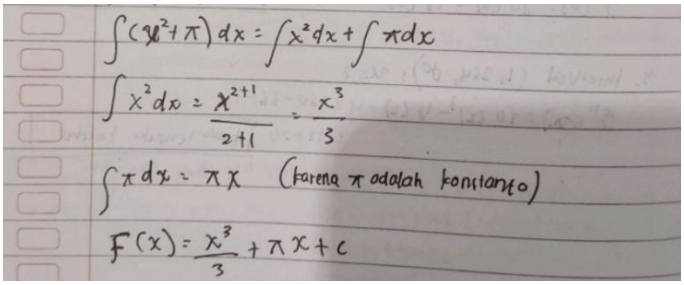


Figure 6. M17, Item 3: The Reason "Because π Is a Constant"

These reasons are predictive arguments anchored in intrinsic properties (the meaning of a constant, the domain of validity of the rule, the parameters of the rule). The criteria of plausibility and mathematical foundation in Lithner (2008) are therefore partly met. The creativity criterion is not met because the strategy was recalled rather than constructed, so this reasoning remains imitative even though it is the closest to local creative reasoning. All three criteria must be met together (Lithner, 2008; Jonsson, Granberg and Lithner, 2020), so mathematical foundation alone is not enough. In addition, the item set contains no case $r = -1$, so the condition $n \neq -1$ was written but never tested, and stating it cannot be distinguished from quoting a formula.

2. Supporting Findings

Errors and interpretation. Table 4 summarises the errors and interpretive variants that are visible. All of them centre on surface properties of notation, in line with the finding of Altindis (2025) that imitative reasoning rests on surface cues.

Table 4. Errors and Interpretive Variants in the Written Traces

Type	Form	Script	Link to reasoning
Notation	∫5 dx written as 5 + c (variable lost)	M10	Surface pattern "a constant stays a constant"
Carried-over template	The function is changed to f(x) + 5 so the result contains +5x	M08	Procedure run without verification
Item interpretation	√2x read as √2·x by most scripts; M07 read it as √(2x)	Many; M07	Notational ambiguity not questioned
Notation	A × written in place of +	M08	Surface slip
Copying slip	$(x^3 + x)$ written for $(x^2 + x)$ and then solved as x^2	M18	No check back against the item
Item abandoned	Item 23 crossed out and not continued	M07	The long algorithm was not available

First, M10 wrote ∫5 dx = 5 + c (Figure 7): the variable x is missing from the result. This can be read as the surface pattern "a constant stays a constant", and it is inconsistent with the intrinsic property ∫a dx = ax + C, which other scripts wrote explicitly (M01, M12).

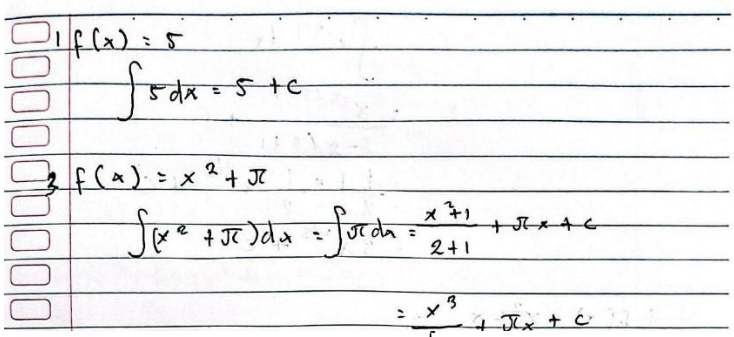


Figure 7. M10, Items 1 and 3: ∫5 dx Written as 5 + c

Second, M08 wrote g(x) = f(x) + 5 in item 11 and the following item, so the result contains the term "+5x" (Figure 8). The constant 5 from item 1 appears to have been carried over as a template into another item. Differentiating the result would give f(x) + 5 rather than f(x), so one verification step would be enough to catch this error. A third error is the copying slip of M18, who wrote $(x^3 + x)$ for $(x^2 + x)$, a mismatch that rereading the item would also have caught.

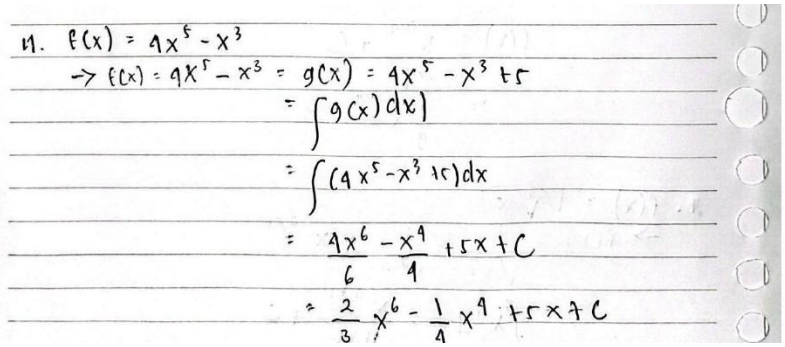


Figure 8. M08, Item 11: The Function Changed to f(x) + 5 before Integration

Absence of verificative argumentation. On the pages we examined, no script differentiated F(x) to check that F'(x) = f(x), and none commented on the plausibility of a result. Yet verification is cheap for antiderivatives and rests on the intrinsic relation between the integral and the derivative. Verificative argumentation is a component of creative reasoning in Lithner's framework (Jonsson, Granberg and Lithner, 2020). Its absence can be explained by a task instruction that did not ask for it, and by the fact that written traces record only what was written. The consequence is nevertheless real: the errors of M08 and M18 were not caught by their authors.

Comparison case: the definite integral. M20 worked on a definite integral. For the integral from −1 to 2 of $(3x^2 - 2x + 3)dx$, the script wrote the antiderivative $x^3 - x^2 + 3x$ and then lines labelled "Value at 2: 10", "Value at −1: −5", and "Difference: 10 − (−5) = 15" (Figure 9). The structure is well organised, but no reason is given for why the difference of antiderivative values gives the result, so it remains algorithmic. This is consistent with the finding that students rely more on antiderivative and area conceptions than on Riemann-sum conceptions (Jones, 2015b).

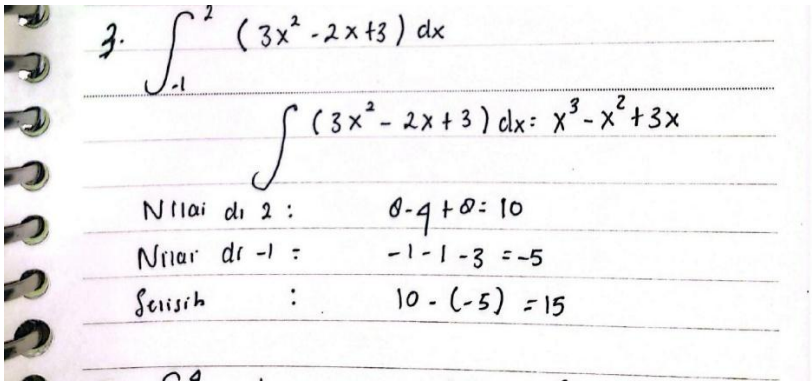


Figure 9. M20: Definite Integral with the Stages "Value at the Upper Limit", "Value at the Lower Limit", and "Difference"

3. Negative Data, Divergent Findings, or Anomalies

Four findings did not follow the general pattern or deserve caution. First, the most divergent interpretation came from M07, who read √2x in item 13 as √(2x), integrated it as $\sqrt{2}\cdot x^{1/2}$, and obtained $(2\sqrt{2}/3)x^{3/2}$, whereas most scripts read it as $\sqrt{2}\cdot x$ and obtained $(\sqrt{2}/2)x^2$. The notation of the item allows both readings, but no script mentioned the ambiguity or the reading it chose. Second, the same M07 crossed out and abandoned item 23, although it was one of the scripts that wrote the most reasons; reasoning at Level C did not therefore guarantee persistence with a long algorithm. Third, no script reached Level D. This is a negative result about the written traces and the task, and it should not be read as evidence that these students cannot reason creatively. Fourth, the condition $n \neq -1$ was written but never tested, because the items did not include $r = -1$.

DISCUSSION

1. Analysis and Interpretation of Findings

The main finding is that all 19 antiderivative scripts showed algorithmic reasoning (RQ1). The strategy was the same everywhere and only the amount of rule and reason that was written down varied. This does not show that the students cannot reason creatively; it shows that the task did not demand it and that the written traces did not display it. The routine items were designed to be solved by remembered rules, and the students solved them in exactly that way.

For RQ2, strategy choice and implementation were visible in every script, predictive argumentation only in three (Level C), and verificative argumentation in none. The variation between students therefore lies in how far rules and reasons are stated, not in the strategy. For RQ3, the errors came from surface properties: templates carried across items, notation, copying slips, and item readings that were not questioned. They are the kind of error that one verification step would catch, which links the absence of verification to the persistence of the errors.

2. Comparison with Theories and Previous Literature

The dominance of algorithmic reasoning agrees with research showing that routine tasks elicit imitative reasoning (Lithner, 2004; Bergqvist, 2007; Boesen, Lithner and Palm, 2010; Jäder, Lithner and Sidenvall, 2020) and that students facing given methods tend to

imitate them (Jonsson, Granberg and Lithner, 2020). It also fits the report that creative reasoning has few opportunities in undergraduate calculus assessment (Mac an Bhaird, Nolan, O'Shea and Pfeiffer, 2017). Our result adds to these task analyses the view from students' own written solutions: the imitative character predicted from the tasks is indeed what appears in the work.

Like Altindis (2025), we found dependence on surface cues and little justification; the difference is that our items had no quantitative context, so covariational reasoning (Carlson, Jacobs, Coe, Larsen, *et al.*, 2002) was not invited. Sholihah and Asyhar (2023) examined visual reasoning about integrals; the present symbolic work complements it. Mahir (2009) linked conceptual understanding to good procedural performance. Our data do not measure conceptual understanding, but the Level C scripts contained more conceptual cues than the Level A scripts. This is consistent with the claim that procedural and conceptual knowledge are related rather than separate (Engelbrecht, Harding and Potgieter, 2005; Baroody, Feil and Johnson, 2007), and with the caution of Tatira (2025) that technical fluency is not enough. The finding that a script can state the correct condition without ever testing it shows why mathematical foundation alone does not make reasoning creative, in line with the three-criteria definition of Lithner (2008). Levels A to D could be validated further with a progression-scale approach as in Callingham and Siemon (2021).

3. Research Implications

Theoretically, the study shows how Lithner's framework can be operationalised for the written traces of symbolic antidifferentiation, with indicators (Table 2) and four levels that separate imitative from creative reasoning by the three criteria. Practically, following the design principles of Lithner (2017), tasks could ask students to verify a result by differentiating it and to say why it is correct; include an item with $r = -1$ so that the condition $n \neq -1$ is tested rather than quoted; present an incorrect solution (as in Figure 8) for students to correct; ask students to reconstruct the power rule from the derivative rule; link antiderivatives to the graphs of f and F (Jones, 2013); and use a levelled rubric such as Levels A to D for formative assessment.

4. Research Limitations and Directions for Future Research

This study has several limitations. Coding was done by one researcher without a test of inter-coder reliability. The data are written traces only, without interviews or think-aloud protocols, so the absence of an argument does not necessarily mean the absence of reasoning,

and we do not know whether students were asked to explain their steps. The level classification rests mainly on the first pages of each script and is holistic; the sample is not random; the similarity of steps cannot be explained from these data; and the completeness of work differs between scripts. Future research can add think-aloud interviews, a second coder, and tasks designed to demand creative reasoning.

CONCLUSION

In answer to the research questions, the reasoning visible in the 19 antiderivative scripts was entirely algorithmic (11 plain, 5 rule-stating, 3 reasoned), and none met the three criteria of creative reasoning. Strategy choice and implementation were clearly visible, predictive argumentation appeared in only three scripts, and verificative argumentation did not appear at all. The errors came from templates carried across items, copying slips, and notation read without question, and such errors are easily caught when verification is required. Routine indefinite-integral tasks therefore invite only imitative reasoning.

The study contributes to knowledge by offering coding indicators and four reasoning levels for the written traces of indefinite-integral solutions, empirical evidence from authentic student scripts, and concrete suggestions for tasks and assessment. For creative reasoning to be practised and assessed, tasks need to demand verification, justification, and the testing of the conditions of a rule.

REFERENCES

Altindis, N. (2025). Networking theories of quantitative reasoning and mathematical reasoning to explore students' understanding of functions. *Journal of Mathematical Behavior, 80*, Article 101276.

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education, 38*(2), 115–131.

Bergqvist, E. (2007). Types of reasoning required in university exams in mathematics. *Journal of Mathematical Behavior, 26*, 348–370.

Bergqvist, T., & Lithner, J. (2012). Mathematical reasoning in teachers' presentations. *Journal of Mathematical Behavior, 31*(2), 252–269.

Bergqvist, T., Lithner, J., & Sumpter, L. (2008). Upper secondary students' task reasoning. *International Journal of Mathematical Education in Science and Technology, 39*(1), 1–12.

Boesen, J., Lithner, J., & Palm, T. (2010). The relation between types of assessment tasks and the mathematical reasoning students use. *Educational Studies in Mathematics, 75*(1), 89–105. https://link.springer.com/article/10.1007/s10649-010-9242-9

Callingham, R., & Siemon, D. (2021). Connecting multiplicative thinking and mathematical reasoning in the middle years. *Journal of Mathematical Behavior, 61*, Article 100837.

Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education, 33*(5), 352–378.

Engelbrecht, J., Harding, A., & Potgieter, M. (2005). Undergraduate students' performance and confidence in procedural and conceptual mathematics. *International Journal of Mathematical Education in Science and Technology, 36*(7), 701–712.

Jäder, J., Lithner, J., & Sidenvall, J. (2020). Mathematical problem solving in textbooks from twelve countries. *International Journal of Mathematical Education in Science and Technology, 51*(7), 1120–1136.

Jäder, J., Sidenvall, J., & Sumpter, L. (2017). Students' mathematical reasoning and beliefs in non-routine task solving. *International Journal of Science and Mathematics Education, 15*(4), 759–776. https://link.springer.com/article/10.1007/s10763-016-9712-3

Jeannotte, D., & Kieran, C. (2017). A conceptual model of mathematical reasoning for school mathematics. *Educational Studies in Mathematics, 96*(1), 1–16. https://doi.org/10.1007/s10649-017-9761-8

Jones, S. R. (2013). Understanding the integral: Students' symbolic forms. *Journal of Mathematical Behavior, 32*(2), 122–141.

Jones, S. R. (2015a). Areas, anti-derivatives, and adding up pieces: Definite integrals in pure mathematics and applied science contexts. *Journal of Mathematical Behavior, 38*, 9–28.

Jones, S. R. (2015b). The prevalence of area-under-a-curve and anti-derivative conceptions over Riemann sum-based conceptions in students' explanations of definite integrals. *International Journal of Mathematical Education in Science and Technology, 46*(5), 721–736.

Jonsson, B., Granberg, C., & Lithner, J. (2020). Gaining mathematical understanding: The effects of creative mathematical reasoning and cognitive proficiency. *Frontiers in Psychology, 11*, Article 574366. https://doi.org/10.3389/fpsyg.2020.574366

Jonsson, B., Norqvist, M., Liljekvist, Y., & Lithner, J. (2014). Learning mathematics through algorithmic and creative reasoning. *Journal of Mathematical Behavior, 36*, 20–32.

Lithner, J. (2000). Mathematical reasoning in school tasks. *Educational Studies in Mathematics, 41*(2), 165–190. https://link.springer.com/article/10.1023/A:1003956417456

Lithner, J. (2003). Students' mathematical reasoning in university textbook exercises. *Educational Studies in Mathematics, 52*(1), 29–55. https://link.springer.com/article/10.1023/A:1023683716659

Lithner, J. (2004). Mathematical reasoning in calculus textbook exercises. *Journal of Mathematical Behavior, 23*(4), 405–427.

Lithner, J. (2008). A research framework for creative and imitative reasoning. *Educational Studies in Mathematics, 67*(3), 255–276. https://link.springer.com/article/10.1007/s10649-007-9104-2

Lithner, J. (2011). University mathematics students' learning difficulties. *Education Inquiry, 2*(2), 289–303.

Lithner, J. (2017). Principles for designing mathematical tasks that enhance imitative and creative reasoning. *ZDM Mathematics Education, 49*(6), 937–949. https://link.springer.com/article/10.1007/s11858-017-0867-3

Mac an Bhaird, C., Nolan, B. C., O'Shea, A., & Pfeiffer, K. (2017). A study of creative reasoning opportunities in assessments in undergraduate calculus courses. *Research in Mathematics Education, 19*(2), 147–162.

Mahir, N. (2009). Conceptual and procedural performance of undergraduate students in integration. *International Journal of Mathematical Education in Science and Technology, 40*(2), 201–211.

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://link.springer.com/article/10.1007/s10649-017-9773-4

Norqvist, M., Jonsson, B., Lithner, J., Qwillbard, T., & Holm, L. (2019). Investigating algorithmic and creative reasoning strategies by eye tracking. *Journal of Mathematical Behavior, 55*, Article 100701.

Orton, A. (1983). Students' understanding of integration. *Educational Studies in Mathematics, 14*(1), 1–18. https://link.springer.com/article/10.1007/BF00704699

Sealey, V. (2014). A framework for characterizing student understanding of Riemann sums and definite integrals. *Journal of Mathematical Behavior, 33*, 230–245.

Sholihah, U., & Asyhar, B. (2023). The students' visual reasoning in solving integral problems. *Journal Focus Action of Research Mathematic (Factor M), 6*(2), 15–26. https://jurnalfaktarbiyah.iainkediri.ac.id/index.php/factorm/article/view/1988

Sidenvall, J., Lithner, J., & Jäder, J. (2015). Students' reasoning in mathematics textbook task-solving. *International Journal of Mathematical Education in Science and Technology, 46*(4), 533–552.

Tallman, M. A., Carlson, M. P., Bressoud, D. M., & Pearson, M. (2016). A characterization of Calculus I final exams in U.S. colleges and universities. *International Journal of Research in Undergraduate Mathematics Education, 2*(1), 105–133. https://link.springer.com/article/10.1007/s40753-015-0023-9

Tatira, B. (2025). Undergraduate students' understanding of the application of integral calculus in kinematics. *Eurasia Journal of Mathematics, Science and Technology Education, 21*(3), Article em2601. https://doi.org/10.29333/ejmste/16049