Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

How Junior Secondary Mathematics Teachers Reason About Financial Literacy Tasks Beyond Calculation

Endah Budi Rahaju[1]

endahrahaju@unesa.ac.id

Universitas Negeri Surabaya

Yurizka Melia Sari

Universitas Negeri Surabaya

Yulia Izza El Milla

Universitas Negeri Malang

Nia Wahyu Damayanti

Universitas Negeri Surabaya

Research Article

Keywords: Problem posing, Financial literacy, Mathematical reasoning, Junior secondary mathematics teachers, Context-based tasks

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11617759/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Designing context-based mathematics tasks requires teachers to interpret a situation, decide which information is relevant and sufficient, and judge whether claims and decisions fit the context. This collective case study examined how six junior secondary mathematics teachers reasoned while working on two financial literacy tasks and discussing their reasoning in task-based interviews. Transcripts were analysed using nine operational indicators organised into understanding, indirect evidence relevant to problem formulation, and evaluation. The analysis showed that all teachers could reconstruct the financial situations, but they differed in how they selected information, identified missing or ambiguous conditions, and justified decisions. Four teachers questioned the sufficiency of the task data. Evaluation combined calculations with contextual judgements, personal experience, rules of thumb, internet-based information, and normative principles. Several conceptual slips involved gross and net income, emergency-fund calculations, and interest periods. Because participants were not asked to generate new problems, evidence concerning problem formulation is indirect and is interpreted cautiously. The findings suggest that preparation for financial task design should integrate mathematical reasoning, financial concepts, explicit treatment of assumptions, and critical evaluation of contextual information. The study also identifies data sufficiency and the distinction between calculation and contextual judgement as important considerations for future research on teachers' financial problem posing.

1. Introduction

Problem posing, the generation of new problems or the reformulation of given ones (Silver, 1994), has drawn sustained research attention for more than two decades (Singer et al., 2013). Compared with solving, posing requires assembling given elements into a coherent problem that can actually be solved (Cai & Hwang, 2002; Silver & Cai, 1996). Work with prospective and practising teachers shows that posing can be learned and that it shapes what teachers ask students to think about (Cai & Hwang, 2021; Crespo, 2003; Ellerton, 2013). Learning to pose is also a route to learning to teach (Cai & Hwang, 2020), and posing has been linked to affect as well as cognition (Cai & Leikin, 2020).

Curricula increasingly ask teachers to place mathematics in real-life contexts. One context of growing importance is financial literacy, which the OECD describes as knowledge and understanding of financial concepts and risks together with the skills, motivation, and confidence to apply that knowledge in effective financial decisions (OECD, 2019). Financial literacy has been assessed in PISA since 2012 (OECD, 2014), and it is associated with household decisions and economic outcomes (Lusardi & Mitchell, 2014). In Indonesia the national survey reported a financial literacy index of 49.68 per cent (Otoritas Jasa Keuangan, 2022), and Indonesian students performed in the lower range in the PISA 2018 financial literacy assessment (OECD, 2020). Mathematics supplies much of the machinery for financial decisions, including percentages, ratios, proportions, algebra, and statistics (Aprea et al., 2016; Blue et al., 2021), which makes mathematics teachers natural authors of financial literacy tasks, yet teachers' own perceptions of financial literacy suggest a need for professional learning (Sawatzki & Sullivan, 2017).

A good context-based task does more than wrap a mathematical concept in a story. It should invite students to reason meaningfully about a realistic situation (Van Den Heuvel-Panhuizen, 2005). Yet tasks produced for classrooms tend to treat context superficially, using ambiguous wording, unfamiliar terms, or contexts that are not authentic (Palm, 2011; Wijaya et al., 2015), and what counts as authentic is itself debated (Vos, 2018). Writing a better task starts with reading the context accurately, noticing what is missing, and judging whether the numbers make sense. These are the first and last steps of a posing process.

Research on problem posing has largely examined the problems produced rather than the reasoning that precedes and accompanies their production (Cai et al., 2013). Process-oriented studies have subsequently described activities such as situation analysis, variation, generation, problem solving, and evaluation (Baumanns & Rott, 2022), as well as orientation, connection, generation, and reflection (Cai & Rott, 2024). Most of this work has involved students or prospective teachers (for example, Ponte & Henriques, 2013; Xie & Masingila, 2017) and mathematical rather than financial contexts. Less is known about how practising junior secondary mathematics teachers interpret and evaluate financial contexts before they are asked to design tasks. The present study addresses that gap by examining teachers' reasoning with ready-made financial literacy tasks. Importantly, the study does not treat interview evidence as direct evidence of problem posing; instead, it examines reasoning that may be relevant to the design of financially and contextually sound tasks.

RQ1: How do junior secondary mathematics teachers reason when they interpret a financial literacy task?

RQ2: How do they evaluate the sufficiency, realism, and feasibility of the data, claims, and decisions presented in the task?

RQ3: What implications do these reasoning patterns have for teachers' preparation to design financial literacy problems?

2. Conceptual Framework

2.1 Reasoning in problem posing

Process models of problem posing agree that posing is not a single act. Baumanns & Rott (2022) observed prospective teachers in task-based interviews and described five recurring activities: situation analysis, variation, generation, problem solving, and evaluation, which are linked cyclically rather than in a fixed order. Cai & Rott (2024) reviewed 75 papers and proposed four phases: orientation, connection, generation, and reflection. Earlier, Christou et al. (2005) classified posing processes as editing, selecting, comprehending and organising, and translating, and Kontorovich et al. (2012) proposed a framework for the complexity of posing in small groups (see also Koichu & Kontorovich, 2013). For teachers specifically, Cai & Hwang (2020) set out how learning to teach through posing can be conceptualised and studied.

The larger project uses three analytic stages (understanding the task, formulating the problem, and evaluating the problem) as a working synthesis. This three-stage grouping is the authors' interpretation rather than a framework proposed by any one cited source. Figure 1 maps the three stages onto elements of published process models. In the present study, the first and third stages are examined directly through participants' work and explanations. The second stage is examined only indirectly because participants were not asked to formulate a new problem. This distinction is central to the interpretation of the findings.

2.2 Indicators used in this study

We derived nine operational indicators from the study's working framework. The indicators are the authors' operationalisation rather than a pre-existing validated scale. U1–U3 capture understanding; F1–F2 capture indirect evidence relevant to formulation; and E1–E4 capture evaluation. Because no participant was asked to pose a problem, F1 and F2 are not treated as measures of problem-posing performance. They identify talk and reasoning moves that may provide material for later task design.

Table 1
Indicators used to code reasoning

Code	Stage	Indicator
U1	Understanding the task	States the information given (income, spending, savings, debt) accurately
U2	Understanding the task	Separates relevant, irrelevant, and unknown or ambiguous information
U3	Understanding the task	Identifies the mathematical concepts and financial literacy elements involved
F1	Formulating the problem (proxy)	Proposes new scenarios, questions, or options arising from the situation
F2	Formulating the problem (proxy)	Sets explicit assumptions or conditions so the problem can be solved
E1	Evaluating the problem	Checks whether the data in the task are sufficient and complete
E2	Evaluating the problem	Checks whether figures and claims are realistic in the context
E3	Evaluating the problem	Judges the fit and feasibility of options in the financial context, with justification
E4	Evaluating the problem	Reflects on and revises the work or strategy

Note. Codes U, F, and E stand for understanding, indirect evidence relevant to formulation, and evaluating, respectively. F1 and F2 are proxy indicators and are not treated as direct measures of problem-posing performance.

2.3 Financial contexts and task quality

The four content areas of the financial literacy tasks used here (earning income, managing budgets and debt, saving, and managing risk and emergencies) are standard in financial literacy frameworks (OECD, 2019). Their mathematics is often simple, but the reasoning is not. A poser must decide which figures are reliable, whether a promised return is plausible, and whether a stated expense is realistic. These decisions concern authenticity, which depends on whether a task reflects real situations and real criteria (Palm, 2011; Vos, 2018), and which influences how students make sense of word problems (Palm, 2008). They also depend as much on financial knowledge as on mathematics. Real-life artifacts such as price lists and bank documents can serve as sources for posing (Bonotto, 2013), but using them well requires the poser to read them critically first.

3. Methods

3.1 Research design

We used a qualitative collective case study design (Stake, 1995; Yin, 2018). Each teacher constituted an individual case, and the six cases were subsequently compared to identify recurring and contrasting patterns in teachers' reasoning. The data comprised teachers' responses to two written financial literacy tasks and one follow-up task-based interview with each participant. Analysis combined deductive matrix coding using the nine operational indicators with thematic synthesis of recurring patterns and contrasts across cases. The design was appropriate for examining how teachers interpreted and evaluated financial contexts in depth rather than estimating population-level effects.

3.2 Participants

Six junior secondary mathematics teachers who were members of the Mathematics MGMP (*Musyawarah Guru Mata Pelajaran*) for junior secondary schools in one city in East Java, Indonesia, participated in the study. Their teaching experience ranged from 3 to 10 years. The participants comprised three teachers recorded as high and three as medium in the project classification. However, the classifications of T-H3 and T-M3 were provisional and were based on the present analysis. Because this creates potential circularity when the classification is compared with the same reasoning data, financial-literacy level is treated only as contextual participant information and is not used as an inferential grouping variable. Participant identities are represented by codes.

Table 2
Participants and level assignment

Code	Project level	Basis of classification	Total coding score (0 to 27)
T-H1	High	Recorded before analysis	16
T-H2	High	Recorded before analysis	15
T-H3	High	Provisional; based on present analysis	22
T-M1	Medium	Recorded before analysis	20
T-M2	Medium	Recorded before analysis	13
T-M3	Medium	Provisional; based on present analysis	13

Note. Total coding score is reported to describe the coding matrix only. Financial-literacy level is not used as an inferential grouping variable because two classifications were provisional and were based on the present analysis.

3.3 Tasks

Participants completed two written financial literacy tasks, each presenting a short financial scenario followed by several items. Context 1 described Dito, a recent graduate with a monthly salary of IDR 4,000,000 and an online business generating about IDR 1,500,000 per month on average. His monthly expenses were IDR 1,200,000 for rent, IDR 1,250,000 for food, IDR 800,000 for transport, IDR 450,000 for other needs, and IDR 800,000 for a motorbike instalment. He planned to buy a camera for live selling, and his online income could fall by half in some months.

Context 2 described Bu Dian, a junior secondary teacher who had run a home catering business for two years, with gross business income of about IDR 5,000,000 and net business income of about IDR 2,800,000 per month. Her spouse was also a teacher earning IDR 6,000,000. Household expenses were IDR 7,500,000 per month, savings were IDR 8,000,000, and the family had no investments and no emergency fund. She wanted to prepare IDR 100,000,000 for her children's education in ten years. Items concerned an industrial oven costing IDR 7,500,000, three financing options, a product-distribution opportunity requiring IDR 3,000,000 and promising IDR 1,500,000 per month for three hours of work per day, and a scenario in which income fell by 60% for six months.

3.4 Interviews

Following completion of the financial literacy tasks, each participant took part in one individual semi-structured, task-based interview lasting approximately 40–60 minutes. The interview explored how teachers interpreted the task information, identified missing or ambiguous information, justified their responses, and evaluated the sufficiency, realism, and feasibility of the information provided. The interviews were conducted in Indonesian, audio-recorded with participants' consent, and transcribed. Relevant excerpts were translated into English by the authors and lightly edited for readability while retaining their intended meaning.

3.5 Data analysis

The analysis proceeded in two complementary stages. First, interview transcripts and relevant task responses were examined against the nine operational indicators: U1–U3 for understanding, F1–F2 as indirect evidence relevant to formulation, and E1–E4 for evaluation. For each teacher and indicator, the evidence was coded on a four-point scale: 0 (not observed), 1 (partly present, general, or containing an error affecting accuracy), 2 (present and mostly accurate), and 3 (strong, accurate, and justified or self-checked). Supporting transcript passages were retained alongside the coding matrix so that interpretations could be traced to the underlying data. Second, coded evidence was synthesised thematically to identify recurring patterns, contrasts, and illustrative cases across participants (Braun & Clarke, 2006; Nowell et al., 2017; Tracy, 2010). The resulting scores were used as descriptive case summaries rather than as a validated measurement scale, and no inferential comparison was made between the high- and medium-level classifications.

To strengthen coding credibility, a second researcher independently coded a subset of the data. Inter-rater reliability was assessed using the intraclass correlation coefficient (ICC), yielding an ICC of 0.73 (for guidance on intercoder reliability in qualitative research, see O'Connor & Joffe, 2020). Disagreements were discussed by the researchers and resolved through consensus before the final coding matrix was used for interpretation.

4. Results

4.1 Overview of coding

Table 3 provides the participant-level coding matrix. The matrix is used as a descriptive aid rather than as a psychometric scale. The most consistent pattern was relatively strong evidence of understanding the presented information, while evidence relevant to formulation was limited because participants were not asked to generate problems. Evaluation varied more substantially across cases, particularly in how participants tested realism, justified decisions, and reflected on errors or alternative strategies.

Table 3
Coding matrix of nine indicators by participant

Teacher	U1	U2	U3	F1	F2	E1	E2	E3	E4	Total
Project high-level code										
T-H1	**2**	**3**	**2**	**1**	**0**	**0**	**3**	**3**	**2**	**16**
T-H2	**1**	**3**	**2**	**1**	**0**	**0**	**3**	**2**	**3**	**15**
T-H3	**3**	**3**	**2**	**2**	**2**	**3**	**2**	**3**	**2**	**22**
Project medium-level code										
T-M1	**2**	**3**	**2**	**2**	**2**	**1**	**3**	**3**	**2**	**20**
T-M2	**2**	**2**	**1**	**1**	**0**	**2**	**1**	**2**	**2**	**13**
T-M3	**2**	**2**	**2**	**1**	**0**	**2**	**1**	**2**	**1**	**13**

Note. Scores run from 0 (not observed) to 3 (strong). U = understanding; F = indirect evidence relevant to formulation; E = evaluating. Scores are descriptive case summaries, not validated measurement scores.

4.2 Understanding the task

Rebuilding the financial picture. Every teacher reconstructed the situation from the figures and calculated what remained after expenses. T-H3 worked through Context 1 under the 50 per cent scenario, adding IDR 750,000 of reduced online income to the IDR 4,000,000 salary and subtracting IDR 4,500,000 of expenses to leave IDR 250,000. For Context 2 under the 60 per cent scenario, the same teacher computed 40 per cent of IDR 2,800,000 as IDR 1,120,000 and added the spouse's IDR 6,000,000 to reach IDR 7,220,000. T-M3 computed IDR 5,500,000 of income against IDR 4,500,000 of spending for Context 1, leaving IDR 1,000,000. T-H1 organised the information by type, separating fixed from unstable income, compulsory expenses, savings, installments, and running debts.

The way teachers selected information differed. T-M1 marked the text as a working method:

"What I underline is what I use as the basis for analysis. What is not underlined, I consider not important to analyse." (T-M1)

T-M1 also separated the certain salary of IDR 4,000,000 from the average online income of IDR 1,500,000 and treated only the salary as reliable. On that basis the teacher concluded that nothing remained each month. Strictly, a salary of IDR 4,000,000 against expenses of IDR 4,500,000 leaves a deficit of IDR 500,000, which the teacher acknowledged elsewhere in the interview as running "minus". The example shows a sound conceptual move, discounting uncertain income, carried out with imprecise arithmetic.

Distinguishing kinds of income and cost. Concepts of gross and net were handled unevenly. T-M1 treated the IDR 1,500,000 online income as gross profit that still had to bear operating costs. T-M3 combined IDR 5,000,000, which is the gross income of the catering business, with another figure to reach

IDR 7,800,000 for Bu Dian's own income and then IDR 13,800,000 for the household, which suggests that gross and net figures were mixed. Household income for Context 2 ended up between IDR 7.8 million (T-M1) and IDR 13.8 million (T-M3), with T-H1 reporting IDR 11 million, whereas the net business income and the spouse's salary alone add up to IDR 8.8 million.

Gaps and ambiguity in the task. Four teachers noticed that something was missing. T-H3 pointed out that Bu Dian's own salary was not stated, assumed IDR 6,000,000 at first, and then set it aside:

"Here she is a teacher, but her salary is not mentioned. So yesterday I took 6 million, because it is not stated here. This is what still bothers me, I think the problem is still incomplete." (T-H3)

T-M3 asked whether the household expense of IDR 7,500,000 already included the children's school costs, and said the price of the camera was unknown. T-M2 observed that the kind of small business was not given although it would affect income. T-M1, asked for a break-even point, answered that there was none in the task. The other two teachers at the high level (T-H1 and T-H2) did not raise such concerns.

4.3 Formulating the problem: proxy evidence

No teacher was asked to pose a problem, so what follows concerns proxy moves only. T-M1 set explicit limits and introduced new options: the motorbike installment should not exceed about IDR 600,000 for Dito to break even on a IDR 4,000,000 income, a drop of up to 50 per cent in online sales should be treated as normal, a government-backed small business loan could replace a consumer installment, and Bu Dian's IDR 8,000,000 savings could be split into emergency and investment funds. T-H3 stated an assumption (Bu Dian's salary) and offered scenarios such as renting out the camera if it was rarely used and expanding catering through new contacts without buying an oven. T-H1 and T-H2 produced conditional statements, for instance that a business offer was realistic "if the business really runs and produces IDR 1,500,000 a month", but did not state assumptions or extend the situation. T-M2 proposed financing half from savings and half from credit. T-M3 suggested ways for Dito to earn more, such as cheaper rent, night work, and additional online accounts, which are solutions to the given problem rather than new problems.

4.4 Evaluating the problem and the decisions

Checking completeness. Four of the six teachers (T-H3, T-M1, T-M2, T-M3) questioned whether the data were sufficient, as described above. Neither T-H1 nor T-H2 did so.

Testing claims against realism. Four teachers tested claims in the tasks against what seemed realistic. T-H1 and T-H2 both doubted the promised payback of two to three months:

"The logic has to be checked first. Maybe it is just a promotional lure." (T-H1)

"Not every business can be stable within two to three months." (T-H2)

T-M1 argued that a return of 50 per cent of the investment every month was impossible “unless there are no operating costs”, and challenged the figures in Context 1, noting that IDR 1,250,000 for food was under IDR 50,000 a day and “below average”, while rent was “above average”. T-H3 computed that three hours of work for about IDR 50,000 was “not worth it”. The tests were weaker in two other cases. T-M3 judged Bu Dian in Context 2 as already “more than enough” and “not very risky” because of the spouse’s income, even though the task included a 60 per cent fall in business income. T-M2 called Dito’s decision risky because the salary was “not that big” but gave no figures.

Justifying choices. Teachers justified their choices through calculation, personal heuristics, or a mixture. Calculations were most visible in T-M1, who rejected the bank loan because its IDR 1,200,000 monthly installment (a total of IDR 8,400,000 over seven months) exceeded the IDR 300,000 that remained each month, and in T-M3, who compared the total of installment options with the IDR 7,500,000 price of the oven. T-H1 chose the supplier option because it finished in four months and had a lower total cost (IDR 7,693,904 in the written work), accepting a larger monthly payment.

Heuristics supplied many of the other justifications. T-H1 and T-H2 favoured gold on the basis of their own buying experience. T-H1 said:

“I usually invest in gold, because I see for myself that the price always goes up and it is more stable than other investments I do not understand yet.” (T-H1)

T-M1 drew on a bank rule that installments should not exceed one third of salary to argue for an emergency fund of 20 per cent. T-H3 took the advice that an emergency fund should cover nine to twelve months of expenses from an internet search. T-M3 said openly that decisions involve more than calculation:

“In matters of financial literacy, sometimes we also use feelings, and also principles, like in religion, if possible, do not borrow.” (T-M3)

Conceptual and unit slips. Three teachers made slips that would matter if they were writing tasks. T-H3 recommended an emergency fund of IDR 9 million to IDR 12 million when the advice they cited meant nine to twelve times monthly expenses of IDR 7,500,000, which is about IDR 67.5 million to IDR 90 million. T-M1 described bond returns of 5 to 10 per cent “per month”. T-M3 mixed gross and net figures, as noted above.

Reflection and revision. All teachers offered some reflection, but its depth varied. T-H2 separated three possible sources of error:

“I check whether the mistake lies in the calculation, in the decision, or in an outside factor such as the economy.” (T-H2)

T-H1 gave a similar but more general answer. T-M2 corrected an earlier misreading during the interview (“I think I was wrong from the start”) and arrived at a remaining IDR 1,000,000 after rereading the

numbers. T-H3 noted that business forecasts cannot be exact and proposed reusing assets. T-M1 proposed cheaper rent and splitting savings. T-M3, asked about a failed investment, said that investment was not well understood, and offered little reflection.

4.5 Cross-case patterns

Across the six cases, understanding was the most consistent feature (Table 3). Every teacher scored 2 or 3 on separating relevant from uncertain or missing information (U2), and five of the six scored 2 or 3 on stating the given information accurately (U1). Evidence relevant to formulation was thinner. Explicit assumptions or conditions (F2) were coded at level 2 for two teachers (T-H3 and T-M1) and were not observed in the other four cases, which is consistent with the fact that no teacher was asked to pose a problem.

Evaluation showed the widest variation, and the profiles differed in shape more than in overall height. T-H3 and T-M1 had the fullest profiles, with at least partial evidence on every indicator. T-H1 and T-H2 showed strong realism testing (E2) and justification (E3) but no checking of data sufficiency (E1), whereas T-M2 and T-M3 questioned the task data (E1 = 2) but tested claims against realism less convincingly (E2 = 1). Reflection and revision (E4) ranged from 1 to 3 and was most developed in T-H2. Total scores ranged from 13 to 22. Because the scores are descriptive summaries and two of the project-level classifications were provisional, they are not used to rank or group the teachers. What recurs across the cases is a separation between reading a financial situation, which every teacher did, and examining it critically, which only some teachers did consistently.

4.6 Reasoning pathway across the cases

Figure 2 brings the three stages together as a flow of reasoning that emerged from the six cases, with the number of teachers at each project level (three high, three medium) who showed each move. It is a descriptive synthesis of the evidence in Sections 4.2 to 4.5, not a model tested against the data, and the dashed border marks the formulating stage, which rests on indirect evidence only.

Reading the figure by level, the two groups were close in reconstructing and sorting information, in the limited evidence for formulation, and in their use of heuristics. The levels were furthest apart on two moves. All three high-level teachers tested claims against realism, compared with one of the three medium-level teachers, and the high-level teachers showed somewhat deeper reflection (mean E4 of 2.33 against 1.67). In contrast, all three medium-level teachers questioned the sufficiency of the task data, compared with one high-level teacher. Conceptual slips appeared in one high-level and two medium-level teachers. With three teachers per level and two provisional classifications, these contrasts are treated as tendencies to be examined in later work, not as differences between groups.

5. Discussion

The findings indicate that interpreting a financial context involves more than extracting and calculating the numbers. All six teachers could reconstruct the basic financial situation, but their treatment of

information differed. Some distinguished stable from uncertain income or organised information by type, whereas others relied more heavily on arithmetic. This pattern is consistent with the selecting and comprehending-and-organising processes described in earlier problem-posing research (Christou et al., 2005) and with situation analysis in Baumanns & Rott (2022) phase model. In a financial task, such selection is consequential because the task designer must decide which figures are reliable, which conditions are essential, and which uncertainties should remain visible.

A second pattern concerns data sufficiency. Four teachers questioned whether the task contained enough information, and some introduced their own assumptions when they encountered gaps. The spread of household-income interpretations in Context 2 illustrates why this matters: the same financial scenario can support different calculations when a key condition is unclear. At the same time, the study cannot determine whether this variation originated in the task wording or in participants' interpretations until the administered version of the task is verified. The contribution is therefore not that the task was definitively flawed, but that noticing and articulating missing conditions emerged as an important form of contextual reasoning relevant to future task design.

A third pattern is the coexistence of calculation and heuristic judgement. Participants sometimes justified decisions through explicit calculations, but they also drew on personal experience, rules of thumb, internet searches, and normative principles. This suggests that financial task evaluation is not reducible to mathematical correctness. For teachers who design context-based tasks, the issue is not whether personal experience or values should disappear, but whether assumptions and sources of judgement are made explicit and whether the resulting task remains open to defensible reasoning rather than embedding an unexamined personal rule as the expected answer. This observation extends the discussion of authenticity by drawing attention not only to whether a context looks realistic, but also to the criteria used to judge what is realistic.

The conceptual slips identified in the interviews are also important for task design. Confusions involving gross and net income, emergency-fund magnitudes, and interest periods could be transferred directly into a classroom task or answer key. The finding therefore supports a distinction between mathematical calculation and financial-conceptual knowledge: accurate arithmetic does not by itself guarantee a financially coherent interpretation. Because the present study did not include actual problem posing, these implications should be treated as propositions for subsequent research rather than evidence that the observed reasoning predicts the quality of problems teachers would produce.

5.1 Implications

Three implications follow for research and teacher professional development. First, financial problem-design activities should require teachers to state assumptions and identify the information a solver would need, making data sufficiency visible rather than implicit. Second, professional learning should address financial concepts such as gross and net income, emergency funds, financing costs, and interest periods alongside problem-posing strategies. Third, task-design discussions should make the

basis of contextual judgements explicit, including the sources and values used to decide whether a financial claim is realistic. The next empirical step is to examine teachers' actual posed problems, together with their explanations and revisions, so that the formulation stage can be studied directly.

6. Limitations

This study offers an in-depth examination of the reasoning underlying teachers' engagement with mathematical tasks, with several considerations defining its analytical scope and providing opportunities for further investigation. The analysis draws on teachers' work with ready-made tasks and subsequent interviews, offering insights into the considerations that inform task design while establishing a foundation for examining teachers' problem-posing practices directly in future research. The collective case study involved six teachers from one MGMP in an urban setting in East Java, with the findings intended to support analytical rather than statistical generalisation and provide a basis for comparison across different educational contexts. Project-level classifications were included to contextualise participants' professional backgrounds; however, as two classifications were provisionally assigned during the analysis, comparisons across levels are interpreted descriptively. The nine analytical indicators were developed by the authors to capture dimensions of teachers' reasoning relevant to the research objectives, with a second researcher independently coding a subset of the data and yielding an intraclass correlation coefficient (ICC) of 0.73, providing a basis for further refinement and application across broader datasets. In addition, a small number of passages in one transcript required contextual interpretation, while the interpretation of Context 2 figures remained situated within the specific conditions of task administration. Although interview accounts provide valuable access to teachers' articulated reasoning, complementary approaches, such as observations of design processes, think-aloud protocols, and direct analyses of teacher-posed problems, could offer further insights into how pedagogical considerations translate into mathematical task design. Taken together, the findings provide a foundation for extending research on teachers' mathematical task-design reasoning, particularly by examining the relationship between their articulated pedagogical considerations and the mathematical problems they formulate.

7. Conclusions

Six junior secondary mathematics teachers were able to reconstruct the financial situations presented in two tasks, but they differed in how they selected information, identified missing conditions, evaluated realism, and justified decisions. Four teachers explicitly questioned data sufficiency, while evaluation frequently combined calculation with personal experience, rules of thumb, external information, and normative principles. Conceptual slips in financial ideas were also evident. These findings suggest that preparation for designing financial literacy tasks should integrate mathematical reasoning with financial-conceptual knowledge, explicit assumptions, and critical evaluation of contextual information. Because the study did not observe actual problem posing, the proposed implications require confirmation through research that analyses the problems teachers themselves generate and revise.

Declarations

Funding: This research was funded by the 2026 FMIPA UNESA Policy Research Program (Penelitian Kebijakan FMIPA UNESA Tahun 2026).

Ethics approval and consent to participate: Ethical approval for this study was granted by the Research Ethics Committee of Universitas Negeri Surabaya (UNESA), under Ethical Exemption Certificate No. 232/UN38.16/KEP/EC/2026. All procedures were performed in accordance with the relevant guidelines and regulations, and the study was conducted in accordance with the principles of the Declaration of Helsinki. Informed consent was obtained from all participants before data collection. They were informed about the purpose of the study and agreed to participate and to allow their data to be used for research purposes. All participants were adult teachers, and no minors were involved.

Use of AI tools: ChatGPT was used during manuscript preparation for English-language proofreading and improving sentence structure. The authors retained responsibility for the content, interpretation, analysis, and final wording of the manuscript.

Author contributions: EBR: Validation; Investigation and data collection. YMS: Conceptualization; Methodology; Writing – original draft; Writing – review and editing. YIE: Investigation and data collection; Formal analysis. NWD: Investigation and data collection; Formal analysis.

Data availability statement

The datasets generated during and/or analysed during the current study are not publicly available due to the interview transcripts containing information that could identify the participants, but are available from the corresponding author on reasonable request.

References

1. Aprea C, Wuttke E, Breuer K, Koh NK, Davies P, Greimel-Fuhrmann B, Lopus JS. (2016). *International Handbook of Financial Literacy*(C. Aprea, E. Wuttke, K. Breuer, N. K. Koh, P. Davies, B. Greimel-Fuhrmann, & J. S. Lopus, Eds.). Springer Singapore. https://doi.org/10.1007/978-981-10-0360-8
2. Baumanns L, Rott B. The process of problem posing: Development of a descriptive phase model of problem posing. Educational Stud Math. 2022;110(2):251–69. https://doi.org/10.1007/s10649-021-10136-y.
3. Blue L, O'Brien M, Makar K. Financial literacy in the mathematics curriculum: Developing critical citizenship. In: Shimizu Y, Vithal R, editors. School Mathematics Curriculum Reforms: Challenges, Changes and Opportunities. Springer; 2021. pp. 261–78.
4. Bonotto C. Artifacts as sources for problem-posing activities. Educational Stud Math. 2013;83(1):37–55. https://doi.org/10.1007/s10649-012-9441-7.

5. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Res Psychol. 2006;3(2):77–101. https://doi.org/10.1191/1478088706qp063oa.
6. Cai J, Hwang S. Generalized and generative thinking in U.S. and Chinese students' mathematical problem solving and problem posing. J Math Behav. 2002;21(4):401–21. https://doi.org/10.1016/S0732-3123(02)00142-6.
7. Cai J, Hwang S. Learning to teach through mathematical problem posing: Theoretical considerations, methodology, and directions for future research. Int J Educational Res. 2020;102:101391. https://doi.org/10.1016/j.ijer.2019.01.001.
8. Cai J, Hwang S. Teachers as redesigners of curriculum to teach mathematics through problem posing: Conceptualization and initial findings of a problem-posing project. ZDM. 2021;53(6):1403–16. https://doi.org/10.1007/s11858-021-01252-3.
9. Cai J, Leikin R. Affect in mathematical problem posing: Conceptualization, advances, and future directions for research. Educational Stud Math. 2020;105(3):287–301. https://doi.org/10.1007/s10649-020-10008-x.
10. Cai J, Moyer JC, Wang N, Hwang S, Nie B, Garber T. Mathematical problem posing as a measure of curricular effect on students' learning. Educational Stud Math. 2013;83(1):57–69. https://doi.org/10.1007/s10649-012-9429-3.
11. Cai J, Rott B. On understanding mathematical problem-posing processes. ZDM. 2024;56(1):61–71. https://doi.org/10.1007/s11858-023-01536-w.
12. Christou C, Mousoulides N, Pittalis M, Pitta-Pantazi D, Sriraman B. An empirical taxonomy of problem posing processes. ZDM Mathematics Education. 2005;37(3):149–58. https://doi.org/10.1007/s11858-005-0004-6.
13. Crespo S. Learning to pose mathematical problems: Exploring changes in preservice teachers' practices. Educational Stud Math. 2003;52(3):243–70. https://doi.org/10.1023/A:1024364304664.
14. Ellerton NF. Engaging pre-service middle-school teacher-education students in mathematical problem posing: Development of an active learning framework. Educational Stud Math. 2013;83(1):87–101. https://doi.org/10.1007/s10649-012-9449-z.
15. Koichu B, Kontorovich I. Dissecting success stories on mathematical problem posing: A case of the Billiard Task. Educational Stud Math. 2013;83(1):71–86. https://doi.org/10.1007/s10649-012-9431-9.
16. Kontorovich I, Koichu B, Leikin R, Berman A. An exploratory framework for handling the complexity of mathematical problem posing in small groups. J Math Behav. 2012;31(1):149–61. https://doi.org/10.1016/j.jmathb.2011.11.002.
17. Lusardi A, Mitchell OS. The economic importance of financial literacy: Theory and evidence. J Econ Lit. 2014;52(1):5–44. https://doi.org/10.1257/jel.52.1.5.
18. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: Striving to meet the trustworthiness criteria. Int J Qualitative Methods. 2017;16(1):1–13. https://doi.org/10.1177/1609406917733847.

19. O'Connor C, Joffe H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods*, *19*. https://doi.org/10.1177/1609406919899220
20. OECD. PISA 2012 Results: Students and Money: Financial Literacy Skills for the 21st Century: VI. OECD Publishing; 2014.
21. (2019). *PISA 2018 Assessment and Analytical Framework*. OECD OECD, Publishing. https://doi.org/10.1787/b25efab8-en
22. OECD. PISA 2018 Results (Volume IV): Are Students Smart about Money? IV. OECD Publishing; 2020.
23. Otoritas Jasa Keuangan. (2022). *Survei Nasional Literasi dan Inklusi Keuangan Tahun 2022*. Otoritas Jasa Keuangan. https://ojk.go.id/id/Fungsi-Utama/Perilaku-Pelaku-Usaha-Jasa-Keuangan/SNLIK/Pages/Booklet-Survei-Nasional-Literasi-dan-Inklusi-Keuangan-2022.aspx
24. Palm T. Impact of authenticity on sense making in word problem solving. Educational Stud Math. 2008;67(1):37–58. https://doi.org/10.1007/s10649-007-9083-3.
25. Palm T. (2011). Features and impact of the authenticity of applied mathematical school tasks. In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. Stillman, editors, *Trends in Teaching and Learning of Mathematical Modelling*(pp. 201–208). Springer. https://doi.org/10.1007/978-94-007-0910-2_20
26. Ponte JP da, Henriques A. (2013). Problem posing based on investigation activities in university settings. *Educational Studies in Mathematics*, *83*(1), 145–156. https://doi.org/10.1007/s10649-012-9443-5
27. Sawatzki CM, Sullivan PA. Teachers' perceptions of financial literacy and the implications for professional learning. Australian J Teacher Educ. 2017;42(5):51–65. https://doi.org/10.14221/ajte.2017v42n5.4.
28. Silver EA. On Mathematical Problem Posing. Learn Math. 1994;14(1):19–28.
29. Silver EA, Cai J. An analysis of arithmetic problem posing by middle school students. J Res Math Educ. 1996;27(5):521–39. https://doi.org/10.2307/749846.
30. Singer FM, Ellerton NF, Cai J. Problem-posing research in mathematics education: New questions and directions. Educational Stud Math. 2013;83(1):1–7. https://doi.org/10.1007/s10649-012-9434-6.
31. Stake RE. The Art of Case Study Research. Sage; 1995.
32. Tracy SJ. Qualitative quality: Eight big-tent" criteria for excellent qualitative research. Qualitative Inq. 2010;16(10):837–51. https://doi.org/10.1177/1077800410383121.
33. Van Den Heuvel-Panhuizen M. The Role of Contexts in Assessment Problems in Mathematics. Learn Math. 2005;25(2):2–9.
34. Vos P. How real people really need mathematics in the real world": Authenticity in mathematics education. Educ Sci. 2018;8(4):195. https://doi.org/10.3390/educsci8040195.

35. Wijaya A, van den Heuvel-Panhuizen M, Doorman M. Opportunity-to-learn context-based tasks provided by mathematics textbooks. Educational Stud Math. 2015;89(1):41–65. https://doi.org/10.1007/s10649-015-9595-1.
36. Xie J, Masingila JO. Examining interactions between problem posing and problem solving with prospective primary teachers: A case of using fractions. Educational Stud Math. 2017;96(1):101–18. https://doi.org/10.1007/s10649-017-9760-9.
37. Yin RK. Case Study Research and Applications: Design and Methods. 6th ed. Sage; 2018.

Figures

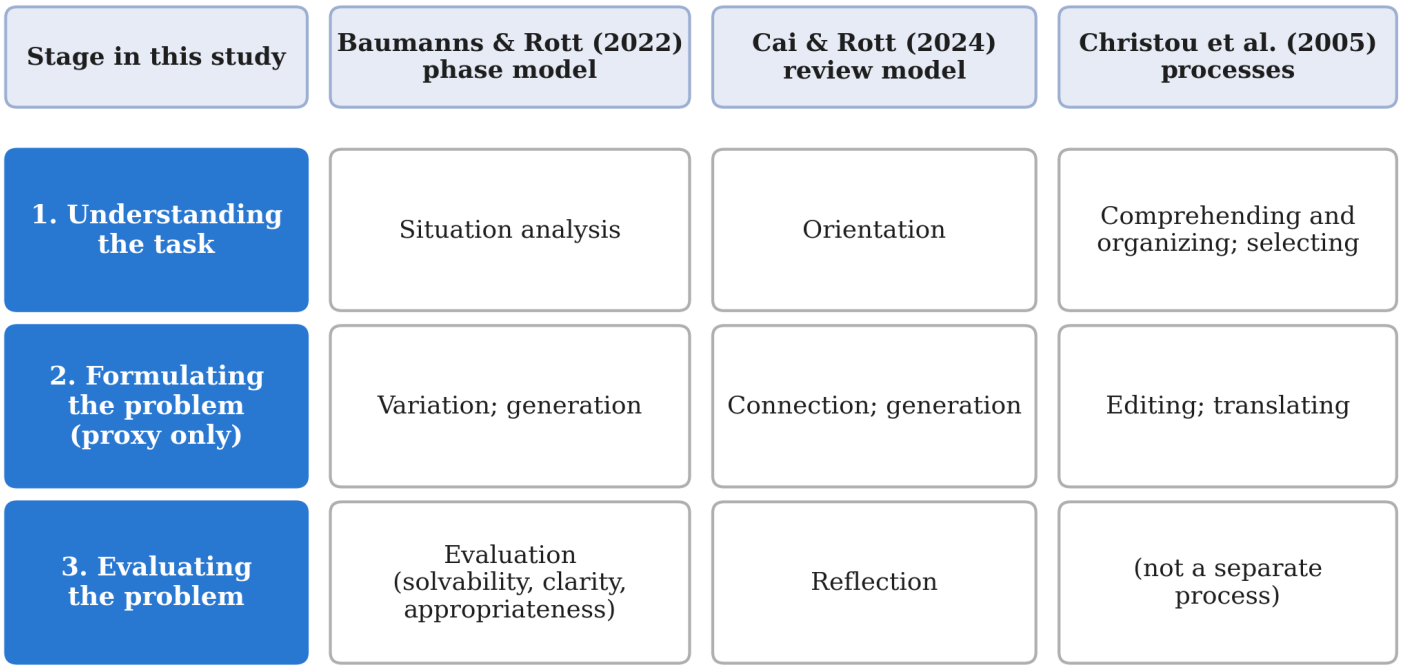


Figure 1

Alignment of the three stages used in this study with published models of problem posing processes

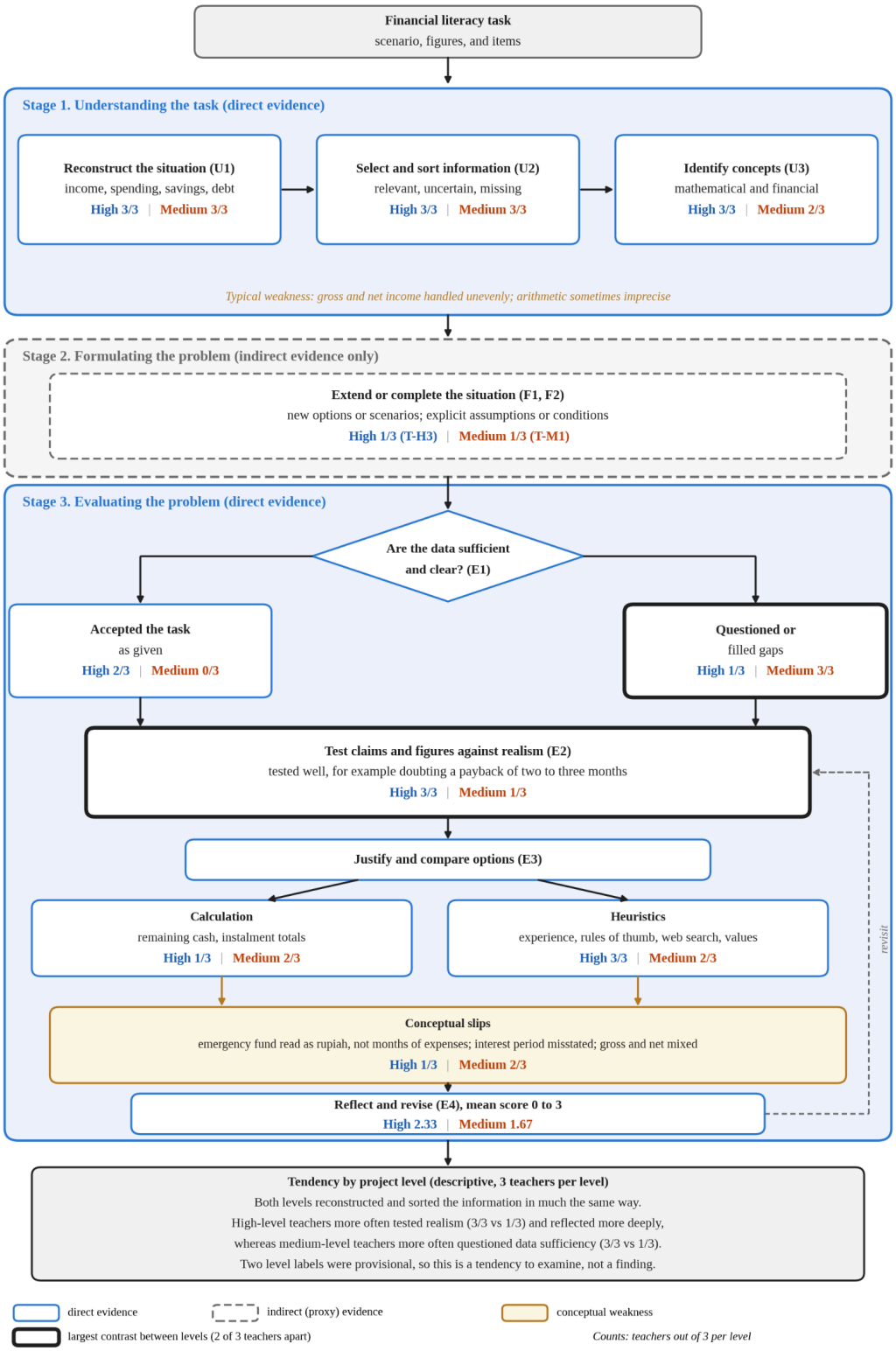


Figure 2

Reasoning pathway when working with the financial literacy tasks, by project level