

Real Contexts in Mathematical Problem Posing. An Exploratory Replicated Study of Students' Creativity

Diana Silva-Pereira^a, Marta Pla-Castells^a, Carmen Melchor-Borja^a

Corresponding author: Carmen Melchor Borja (carmen.melchor-borja@uv.es)

^aDepartamento de Didáctica de la Matemática, Universitat de València, 46022, España

Keywords: Mathematical Problem Posing (MPP), real contexts, creativity, replicated study, Primary Education

Abstract

Mathematical Problem Posing (MPP) is key to developing creativity, but its study in relation to contexts can still be expanded. This work analyses how the context influences the creativity of sixth-grade primary school students when posing problems involving fractions. Thirty-seven students from a public Spanish school in the sixth grade participated in two 40-minute sessions where they formulated problems based on a real mathematical context and a real-life context. Firstly, the productions were classified as problems or statements and were assessed based on their mathematical nature, solvability, and relationship with the context. Then, a creativity index was calculated by comparing the problems posed by each student. The results indicate that the definition used to distinguish between a problem and a statement notably affects the number of productions analysed. Most were mathematical and solvable, but the connection with the context was stronger in the real-life context than in the mathematical one. Finally, no significant differences were found in the levels of creativity between contexts, suggesting that the interpretation of the stimulus may have more relevance.

Introduction

Mathematical Problem Posing (MPP) is a fundamental tool in Mathematics Education, promoting conceptual understanding and higher-order skills (Kilpatrick, 1987; Kilpatrick et al., 2001; Silver, 1994; English, 1997). In contrast to instruction centred on solving routine exercises, MPP is an active process in which students create or reshape problems using given contexts or stimuli, drawing on prior knowledge, making connections, and developing their own representations (Stoyanova and Ellerton, 1996; Brown and Walter, 2005), developing creativity through dimensions such as fluency, flexibility, and originality (Guilford, 1959; Torrance, 1966).

The context in which problems are posed is essential. Authors such as Freudenthal (1991) and Palm (2006) have emphasised the importance of the situations presented to students being recognisable, functional, and plausible, so that they activate mathematical knowledge from prior experience. In this sense, these studies have made it possible to differentiate various types of contexts that provide distinct opportunities for creativity, such as school mathematical contexts—connected to explicit references addressed in class—and real-life contexts, which relate more closely to students' everyday experiences and activities outside school.

Passarella's (2022) study examined how real contexts shape sixth-grade students' creativity in fraction problem posing. Using Stoyanova and Ellerton's (1996) typology, the tasks were

semi-structured: students posed at least three fraction problems from two visual stimuli—a number line (real-mathematical context) and a phone-discount leaflet (real-life context). The results showed that, although no significant differences were found in creativity levels between the two contexts, relevant variations occurred in how students linked their productions to the stimuli and how they redefined the presented situations.

The purpose of this study is twofold. First, we conduct a critical conceptual replication of Passarella's (2022) classroom investigation, examining how different real contexts influence sixth-grade students' creativity in mathematical problem posing with fractions. Second, we refine and problematise the definition of a "mathematical problem" adopted in that study, analysing how alternative criteria for distinguishing problems from statements affect the set of productions included in the creativity analysis.

Rather than a strict statistical replication, this a conceptual replication conducted in a different national and institutional setting. We closely mirror Passarella's (2022) design and procedures, but treat cross-sample contrasts as exploratory, given the small, non-random samples.

With this in mind, the research objectives are:

O1: To examine how the nature of the context (real mathematical vs real-life) influences sixth-grade students' mathematical creativity in problem-posing tasks with fractions, by carrying out a critical conceptual replication of Passarella's (2022) classroom study in a different educational setting.

O2: To refine the definition of a mathematical problem used in Passarella's (2022) work in light of the broader problem-solving and problem-posing literature, and to investigate how this definitional shift impacts the classification of students' productions and the resulting creativity indices in both studies.

This study offers a dual contribution: providing additional empirical evidence on the role of real contexts in primary students' creativity, and showing how alternative, theory-driven definitions of "problem" can alter the empirical picture of creative performance, highlighting the need for conceptual clarity in MPP research.

Theoretical Framework

Mathematical Problem Posing (MPP) and its context

MPP is a process in which students generate or reformulate problems based on diverse stimuli, mobilising prior knowledge and establishing connections (Stoyanova and Ellerton, 1996; Brown and Walter, 2005). This approach is essential for developing mathematical competence (Kilpatrick, 1987; Kilpatrick et al., 2001), promoting conceptual understanding and the ability to think creatively and flexibly (Silver, 1994; English, 1997; Bevan et al., 2019).

In Mathematical Problem Posing (MPP), context shapes how students interpret a task and construct productions. From Realistic Mathematics Education, experiential contexts can bridge formal knowledge and students' reality (Freudenthal, 1991). Palm (2006) proposed that verbal problems should meet the criteria of familiarity, transparency, and consistency with students' prior experience. Cultural artefacts activate connections between school

mathematics and everyday life (Bonotto, 2013), and contexts must be meaningful and recognisable to students (Van den Heuvel-Panhuizen and Drijvers, 2014; Albarracín and Gorgorió, 2019). Similarly, studies on problem-based learning in rich, authentic contexts have shown that such environments can elicit unexpected mathematical responses from young students, even when the focus is on problem solving rather than problem posing (Russo et al., 2020).

In this study, a real context is defined as any situation presented to students that satisfies the specified criteria and is perceived as both plausible and meaningful. Two categories are distinguished: real mathematical contexts, with an explicit school anchorage that students perceive as functional and connected to what is worked on in class, and real-life contexts, linked to extracurricular situations close to students' everyday lives.

Creativity in Mathematical Problem Posing

Mathematical creativity is multidimensional. Guilford (1959) and Torrance (1966) established three dimensions: fluency (produce numerous ideas from the same stimulus), flexibility (variety of approaches, structures, or representations), and originality (novelty or rarity of the productions). Applied to MPP, these dimensions identify students' capacity to generate diverse statements, moving away from routine patterns, and adding mathematical value (Leung and Silver, 1997).

In the present study, creativity is understood as the capacity to produce problems that demonstrate fluency and flexibility within the framework of closed problems. For its evaluation, the rubric developed by Xie and Masingila (2017) is adopted, which allows for a comparative assessment of creativity levels based on the generated productions.

Mathematical Problem

The concept of a mathematical problem has been defined from multiple perspectives in the literature. Authors such as Polya (1981) and Schoenfeld (1992) emphasise the centrality of cognitive challenge: a problem situation for which no immediate solutions method is available, requiring exploration, strategy formulation, and decision-making. Neither author requires that a problem be closed—what matters is the absence of a routine algorithmic path and the need for genuine mathematical reasoning.

Other authors add further structural requirements. Silver (1994) and Rodríguez (2012) argue that problems should also be meaningful, contextualised, and conducive to new learning. Additionally, some frameworks include closed structure (containing all necessary data and conditions for a verifiable solution) as a characteristic of well-formed problems, particularly in classroom assessment contexts (Carrillo, 1996; Espinoza et al., 2014).

In this study, we adopt a more restrictive operational definition of "mathematical problem" for analytical purposes, requiring three conditions: (1) a real cognitive challenge, without an immediate or routine path to the solution; (2) a closed structure, meaning it contains the data and conditions necessary to reach a reasonable solution; and (3) minimum contextualisation that gives meaning to the statement within a school or everyday framework. This choice is a methodological decision aimed at ensuring objective solvability and enabling systematic comparison of structural diversity through the creativity index c (Xie & Masingila, 2017). We recognise that this restricts our analysis to closed problem-posing contexts and does not capture other valid forms of creativity in MPP, such as semantic richness or open-ended exploration (Silver, 1994; Cai et al., 2015).

Productions that do not meet these criteria—such as routine exercises, unsolvable tasks, or statements lacking genuine challenge—are classified as "statements" rather than problems. It is important to highlight that Passarella's (2022) study adopted a more inclusive criterion, encompassing all productions that contained a context and a clearly identified unknown, regardless of cognitive demand. Consequently, some formulations classified as statements in the present study were considered problems in the original research. This definitional contrast is central to our second objective and allows us to examine how analytical framing shapes empirical conclusions about students' creativity in MPP.

This study uses a focused operational framework based on four constructs that shape the design, coding, and analysis: creativity, defined as fluency and flexibility in closed problems and measured through a structural c-index; a context typology distinguishing real-mathematical from real-life situations; a definition of mathematical problem that requires cognitive challenge, closed structure, and meaningful contextualisation; and MPP tasks understood as semi-structured problem reformulations anchored in visual stimuli.

Methodology

Sample

The research involved 37 sixth-grade primary students (aged 11-12 years) from a public urban school in Spain. The group was selected because they had recently worked on fractions, ensuring minimum prior knowledge for the problem-posing task.

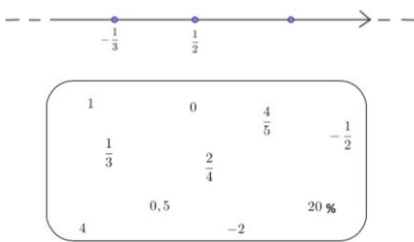

This enables a conceptual replication of Passarella's (2022) study with 22 Italian sixth-grade students. Key similarities include the same grade level, recent fraction instruction, identical semi-structured MPP tasks with visual stimuli, and equivalent timing (two 40-minute sessions). However, differences are inevitable: national curricula vary, no directly comparable pretest exists, and prior exposure to representational tools may differ. These factors preclude assuming statistical equivalence between samples and position cross-study comparisons as exploratory rather than confirmatory.

Participation was voluntary and anonymous, and informed consent was obtained from families, always ensuring respect for the ethical principles of confidentiality and data protection.

Data Collection

The intervention was designed to replicate the conditions of Passarella's (2022) study and ensure comparability between investigations. Two distinct sessions were organised, using semi-structured tasks (Stoyanova and Ellerton, 1996) focused on posing fraction problems (Table 1). Each session lasted 40 minutes, during which participants were asked to formulate at least three mathematical problems based on the provided stimulus.

Table 1. Contexts for the MPP sessions in both samples

Contexts for Fraction MPP	
Context 1	Context 2
 <p>A number line with arrows at both ends, showing points for $-\frac{1}{3}$, $\frac{1}{2}$, and 0. Below the number line is a rounded rectangle containing the following numerical representations: 1, 0, $\frac{4}{5}$, $-\frac{1}{2}$, $\frac{1}{3}$, $\frac{2}{4}$, 0.5, 20%, 4, and -2.</p>	 <p>An advertising leaflet for mobile phones. It features several phone models with prices and discounts. Visible prices include 249, 399, 199, 299, and 549. There are also smaller text elements and icons related to mobile phone features.</p>

Context 1 consisted of a number line with fractions. This was considered a real mathematical context due to its functional value: it was recently worked on in the classroom hence was recognisable and meaningful to the students. Context 2, on the other hand, was presented through an advertising leaflet with telephone discounts, designed according to Palm's (2006) criteria of contextual authenticity. Its selection is justified by students' familiarity with this type of material, allowing connection with their social and cultural environment and offering a favourable setting for creative problem posing.

This design maintains the core conditions of Passarella's (2022) study (semi-structured tasks, visual stimuli, and timing of the sessions) to enable a conceptual replication, while recognising that cross-sample comparisons are exploratory and theoretically informative rather than confirmatory, given the small, context-specific samples.

Data Analysis

Given the exploratory nature of this replication study and the small sample sizes ($n = 37$; $n = 22$), inferential statistics are used in a supportive way. With these sample sizes, statistical power is limited for detecting small effects, so the analyses focus on identifying patterns that are theoretically informative rather than drawing strong generalisations or causal claims from the observed differences.

The data analysis was framed within the positivist paradigm, which seeks to establish systematic relationships between variables (Cohen et al., 2011). A mixed methodology combined quantitative analysis, aimed at identifying patterns and significant differences, with qualitative analysis focused on interpreting representative examples of the diversity of productions. For the statistical treatment, PSPP (2022) software was used, which allowed for contrast tests and obtaining descriptive measures of the variables of interest.

First Coding Phase

In the first phase, students' productions in both contexts were analysed according to the criterion used by Passarella (2022), which considered any production that included a context and a defined unknown as a problem. Subsequently, the productions underwent a second classification based on three dimensions, following the model of Leung and Silver (1997), also used in Passarella's (2022) study. First, it was analysed whether the formulation could be considered of a mathematical nature, meaning it explicitly included content specific to

the area. Second, its solvability was assessed, understood in this study solely in terms of closed structure: that is, it implicitly contained all the data and conditions necessary to reach a reasonable solution. Finally, the relationship with the context was analysed, ensuring that only those productions preserving a meaningful connection to the initial situation, whether related to school or everyday life, were acknowledged as such.

Second Coding Phase

This phase aimed to select productions to estimate the influence of the type of context on creativity by measuring the flexibility and fluency of the problems each student can generate. Unlike the first phase, in this second phase, productions were not required to be explicitly related to the initial context. As in the original study, imposing a narrative anchor to the stimulus could introduce bias into that variability; therefore, contextual relation was retained as a descriptor for contrast purposes but was not used as an inclusion criterion in the calculation of creativity. Next, individual creativity for each student in each context was calculated. The procedure described by Passarella (2022) based on Xie and Masingila (2017) was used, which consists of comparing the valid problems produced by the same individual and assigning them a value according to their degree of structural difference. The sum of these comparisons constitutes the creativity index (c), calculated independently for the real mathematical context and the real-life context.

Finally, the c value was translated into three levels of creativity according to Xie and Masingila (2017). A student was considered to have low creativity when $c \leq 0$, medium creativity when it was 1 or 2, and high creativity when $c \geq 3$. In this way, each individual was classified at a different level in each of the contexts, allowing for the comparison of the distribution of creativity levels and the analysis of the influence exerted by the nature of the context on the production of mathematical problems.

Modifications

Although the design generally followed that of Passarella (2022), several modifications were introduced. First, the materials were translated and linguistically adapted into Spanish to fit the local educational context. Second, while Passarella's study did not include an initial diagnostic test, we ensured that the sixth-grade students had recently worked on fractions content, thus establishing a comparable conceptual foundation for the task. Third, we implemented a double-coding procedure to differentiate the classification of problems and statements according to both the theoretical references adopted in this study and Passarella's (2022) inclusive criterion, directly addressing the second research objective.

This constitutes a conceptual replication that preserves key design elements (tasks, stimuli, procedures, and timing) while explicitly acknowledging that the two samples are not statistically equivalent. Consequently, contrasts are interpreted as exploratory and theoretically informative, rather than as evidence of causal effects or generalisable differences.

Results

First Coding Phase

Three key dimensions were analysed: whether the problem could be considered mathematical, whether it was solvable, and whether it maintained an explicit relationship with the initial context. The results of this classification are shown in Figure 1, which compares the percentages obtained in this research with those of the replicated study.

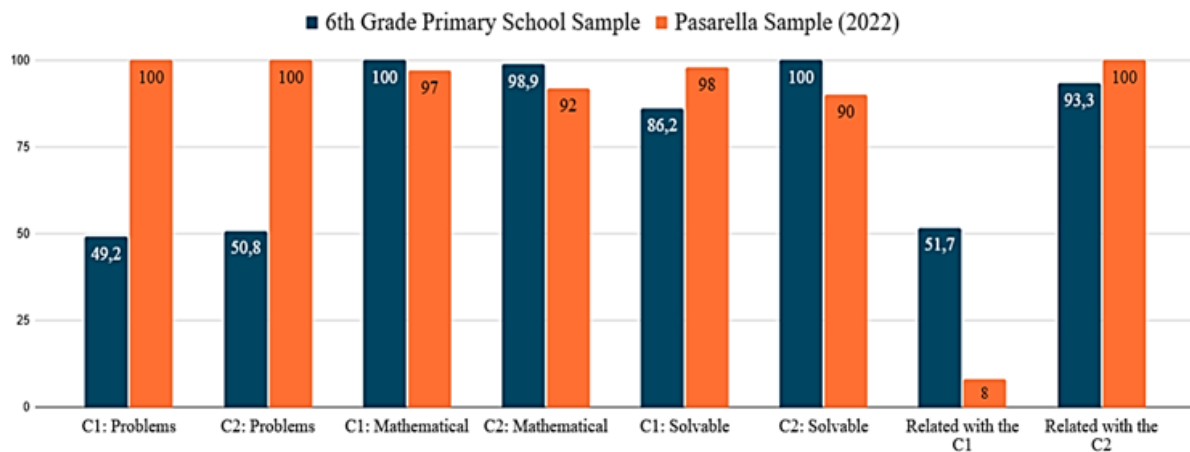


Figure 1. Results of the first coding phase comparing problem classification, mathematical nature, solvability, and context-stimulus connection between the Spanish sample (this study) and Passarella's (2022) Italian sample across both contexts (C1: real-mathematical, C2: real-life)

In the Spanish sample, 49.2% of productions in Context 1 and 50.8% in Context 2 were classified as problems, compared to 100% in both contexts in Passarella's (2022) study. This highlights that students in this research showed greater tendency to generate statements not meeting minimum problem criteria, anticipating interpretative divergence from the original sample.

Regarding the mathematical nature, no significant differences were found between the two samples. In the current study, 100% of problems from Context 1 and 98.9% from Context 2 were considered mathematical, compared to 97% and 92%, respectively, in Passarella's (2022) sample. Furthermore, only one production in this study was not classified as mathematical: "Which phone cost 249 euros?" Such formulations are not recognised as mathematical tasks since their resolution does not require calculations or application of domain knowledge but simply locating a specific datum within the stimulus.

Many problems in both samples were classified as solvable, with percentages approaching 100%, suggesting that participants tend to formulate mathematical problems that can be solved with the available information. Table 2 illustrates productions that were not considered solvable. The first example lacks the specific question, or all required data, deviating from the closed problem definition. The second omits the essential datum of total students, preventing complete resolution. Although these might be solved in open problem formulations, they were classified as unsolvable due to lacking a clear and verifiable solution.

Table 2. Examples of Unsolvable Problems

In a family, 500 euros are distributed among its members. Knowing that there are three members, and that the eldest receives one-third more than the middle child, who in turn receives twice as much as the youngest

In class 6°B and 6°A, they need to form 6 groups for the mini-games prepared by the teachers. How many students will be in each group?

Regarding the relationship with the context, the results show pronounced differences between the two samples. In the Spanish sample, 51.7% of the problems in Context 1 and 93.3% in Context 2 maintained a direct link with the stimulus, whereas in Passarella's (2022) study, only 8% of the productions in Context 1 were related to the number line, compared to 100% in Context 2. Both studies point to the real-life context as a favourable stimulus for generating contextualised problems, but the mathematical context behaved differently. To illustrate examples of these differences, Table 3 is presented.

Table 3. Examples of Problems Related and Not Related to the Context

Related Problems	C1	The fractions have gone off track and become unordered. Find the order they should be placed in on the number line shown above.
	C2	Lorena is going to buy a mobile phone, 25% is paid for her and she pays the rest. If she is going to buy the HTC, how much would she pay?
Unrelated Problems	C1	If there are 40 of us in class, how many children make up 20%?
	C2	In a library, there are 20 books. If this month they have sold $\frac{5}{20}$, how many books remain in the library?

As observed in Table 3, formulations considered related to the context make explicit use of the proposed material, whereas unrelated productions employ contexts disconnected from the initial stimulus, even when they are mathematically valid and solvable.

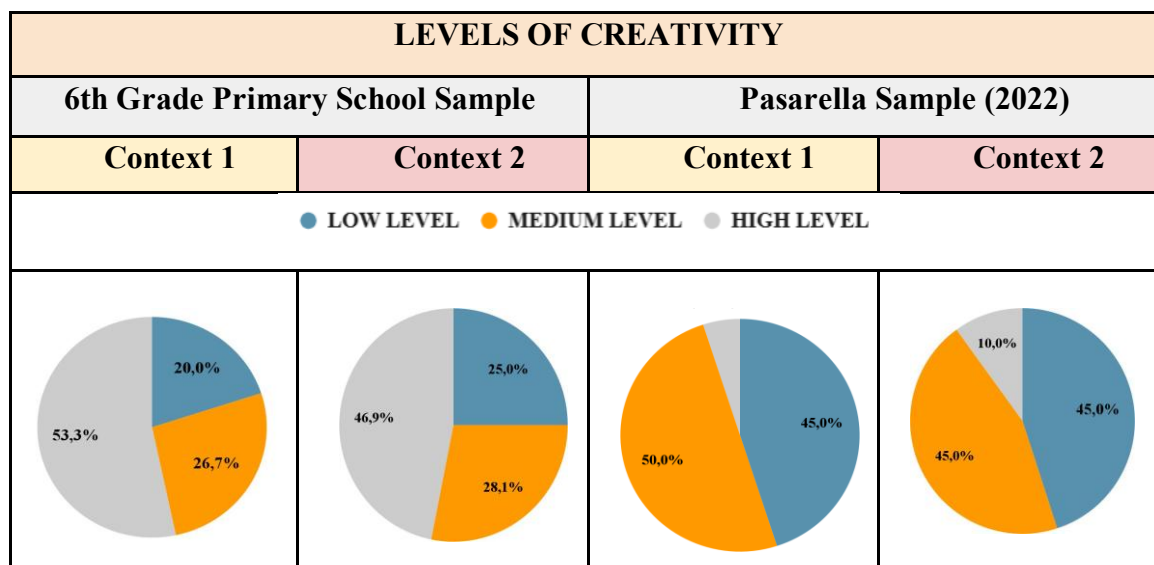
Both studies show high rates of mathematical nature and solvability but diverge in stimulus connection and problem classification. In our sample, around half of the productions in each context met Passarella's (2022) inclusive problem criterion, whereas in the original study all formulations were considered problems. These patterns suggest that real-life contexts favour stronger contextual anchoring in both studies, while differences in the mathematical context and in the adopted problem criteria lead to substantial variation in how students' productions are categorised.

Second Coding Phase

In this phase, only productions that met the three requirements were considered: being classified as problems, including mathematical content, and being solvable. Unlike the first phase, both problems related and unrelated to the initial context were incorporated. The creativity assessment was based exclusively on the structural diversity of the problems posed, without the contextual relationship influencing the calculation.

Not all students reached the minimum of three problems per context indicated in the initial instructions, as several statements were discarded for not meeting the criteria. In total, out of the 164 valid productions analysed, 75 corresponded to the real mathematical context and 89 to the real-life context. From these productions, each student's c value was calculated, and three levels were established. The distribution of these levels is shown in Table 4, where the results of this study's sample are compared with those obtained in the replicated one.

Table 4. Distribution of Creativity Level



The distribution of creativity levels varies noticeably between the two samples. In Context 1, 53.3% of the students in this study were classified at the high creativity level, 26.7% at the medium and 20% at the low level, whereas in Passarella's (2022) sample only 5% attained the high level, with 50% at the medium and 45% at the low level. In Context 2, this study's sample showed 46.9% at the high level, 28.1% at the medium, and 25% at the low level compared with 10%, 45% and 45% in the replicated study.

When analysing the differences between contexts within each sample, the test for related samples did not detect significant differences between the number line and the leaflet in this study ($z = -0.90$; $p = 0.371$) or in Passarella's (2022) research ($p = 0.9$). Therefore, although the percentages suggest contrasts between the samples, neither study can affirm that the type of context has a statistically verifiable effect on the creativity levels achieved.

To illustrate how these differences materialise and the characteristics of each level of creativity, Table 5 presents examples of students' productions in Context 1, at high, medium, and low creativity levels.

Table 5. Examples of students' productions in Context 1 at different levels

High level	P1	Juan and his brother bought a pizza. Juan ate $\frac{1}{3}$ and his brother ate $\frac{2}{4}$. What fraction of the pizza did they eat?
	P2	Manuel is going to buy a mobile phone that costs 300 euros, and he has a 20% discount. What is the final price of the mobile phone?
	P3	There are 2 computers, one costing 1300€ but with a 40% discount and another costing 1000€ with a 25% discount. Which computer is cheaper?
Medium level	P1	A mobile phone costs 300€. It has a 20% discount. What is its final price?
	P2	A tablet costs 200€. A quarter of its price is discounted. How much do you have to pay?
	P3	A mobile phone costs 400€. One-third of its price is discounted, and the buyer has already paid 100€. How much does the buyer still need to pay?
Low level	P1	The teacher draws a number line and asks her students to place the fraction $\frac{2}{4}$ on it. Help her.
	P2	The teacher draws a number line and asks her students to place the fraction $\frac{1}{2}$ on it. Help her.

At the high level, one individual formulated three valid problems that clearly differed in context, structure, and type of unknown: dividing fractions in a pizza, applying a single discount, and comparing prices with a double discount. The comparisons (P1, P2), (P2, P3), and (P1, P3) received values of 2, 1, and 2, totalling $c = 5$. This score reflects high creativity, as it shows that the individual does not repeat patterns but introduces substantial variations in operations, contexts, and reasoning.

At the medium level, the three problems revolved around calculating prices with discounts with slight modifications in objects, fractions, or inclusion of an advance payment. The comparisons yielded scores of 1, 1, and -1 , for a total of $c = 1$ indicating some flexibility but not enough structural variation to be classified at the high level.

At the low level, the problems asked to place fractions on the number line, changing only the value of the fraction. The comparison between both resulted in $c = 0$, as they were considered comparable. This lack of changes in procedure and reasoning confirms what Guilford (1959) and Silver (1994) stated: creativity requires introducing transformations in the structure of the problem, and when these do not appear, flexibility is null, and the classification is at the lowest level.

Overall, the creativity index revealed higher levels of structural diversity in the Spanish sample than in Passarella's (2022) study, with a larger proportion of students reaching the high creativity level in both contexts. However, in neither case did the type of context (real-mathematical vs real-life) produce statistically significant differences in creativity

distributions, reinforcing the idea that the nature of the stimulus alone does not determine students' creative output. Instead, these results point to the importance of how students interpret and transform the given situations, as well as to the influence of the analytical framework used to classify their productions.

Variation in the definition of problem from the original work

As previously mentioned, the original work used a non-restrictive definition, considering any production with a context and defined unknown as a problem. For example, asking a fictional character to place numbers on a number line was classified as a problem due to narrative context, though it could be understood as a simple school exercise. The two low-level problems in Table 5, for example, would not be classified as problems according to the theoretical references of this study.

Therefore, the previous analysis was completed by obtaining a subset of the productions considered problems according to Passarella (2022), using the definition from the theoretical references adopted in this research.

The comparison between the two reference frameworks used revealed notable differences in the categorization of productions. While only 44.1% of the formulations were classified as problems under the theoretical references, this percentage rose to 80.5% when applying Passarella's (2022) approach. The difference was confirmed by a Chi-square test, which showed a statistically significant difference between both classifications ($p < 0.001$). These results indicate that the chosen definition of a mathematical problem directly influences the number of productions categorized accordingly, which subsequently affects the analysis of creativity. To illustrate these differences, Table 6 is presented.

Table 6. Problems according to the adopted approach

Problems according to theoretical references	C1	In a school, 20% of all the litres of water were used in the first trimester. In the second trimester, $\frac{2}{4}$ of the remaining water was used. How many litres are left for the third trimester if there were a total of 100L at the beginning of the school year?
	C2	Jaime has 20% of the money for a tablet he wants to buy for 363 euros. If he had to spend 9% of the money, how much does he have left to buy the tablet?
Problems according to Passarella (2022)	C1	In a supermarket, Marta buys a bottle of water that costs 2 €. Calculate how much 3 bottles would cost her.
	C2	Help Jose find out how much money he saves on the mobile phone that used to cost 200 € and now costs 150 €.
Statements in both approaches	C1	-Order the following fractions from greatest to least: $\frac{1}{3}$; $\frac{1}{2}$; $\frac{4}{5}$; $\frac{2}{4}$. -Write how these fractions are read: $\frac{4}{5}$; $\frac{1}{3}$; $\frac{2}{4}$.
	C2	- Discount all phones by 50%. -Calculate: $199+299+399$.

In the case of problems according to the theoretical references, Table 6 productions pose a true cognitive challenge, requiring students to interpret data, establish relationships, and perform chained calculations to reach the solution. This group includes percentage and fraction problems in which information is presented sequentially, demanding progressive reasoning and planning of intermediate steps.

Conversely, problems according to Passarella (2022) considered statements by theoretical references show a much lower demand. They present a context and a clear unknown but are solved through mechanical procedures such as a simple multiplication or direct discount calculation, without requiring decision-making or complex reasoning. However, they meet Passarella's (2022) minimum requirements, which explains their inclusion in the problem category.

Finally, statements in both approaches include productions without a contextualized situation or a real unknown such as ordering fractions, performing additions, or representing quantities on the number line. These routine tasks do not involve a cognitive challenge or require interpretation, so they are not considered problems in either framework.

These findings show that the proportion and type of productions classified as problems depend heavily on the adopted definition. With theory-driven criteria emphasising cognitive challenge, closed structure, and contextualisation, less than half of the formulations qualify as problems, whereas under Passarella's (2022) inclusive criterion more than four out of five productions are counted as such. This discrepancy, confirmed statistically, indicates that differences in creativity outcomes may reflect analytical framing and categorisation choices at least as much as students' actual problem-posing behaviour.

The comparative overview in Table 7 illustrates methodological alignments and divergences between both studies, particularly regarding problem definition criteria, production classification rates, and creativity distributions.

Table 7. Key similarities and differences between this study and Passarella (2022)

Aspect	This study (Spanish sample)	Passarella (2022, Italian sample)
Grade level and content	6th grade; recent work on fractions	6th grade; recent work on fractions, proportionality, percentages
Task type and stimuli	Semi-structured number line (real-mathematical), phone leaflet (real-life)	Same semi-structured MPP; same number line and phone leaflet stimuli
Operational definition of "problem"	Theory-driven: cognitive challenge, closed structure, contextualisation	Inclusive: any production with context and defined unknown
% of productions classified as problems	44.1% under theory-driven criteria vs 80.5% under Passarella's criterion	100% under Passarella's inclusive criterion in both contexts
Mathematical nature	≈100% (C1) and 98.9%	97% (C1) and 92% (C2)

	(C2) of problems classified as mathematical	classified as mathematical
Solvability of problems	Percentages approaching 100% in both contexts	Percentages approaching 100% in both contexts
Context–stimulus connection	C1: 51.7% related; C2: 93.3% related	C1: 8% related; C2: 100% related
Creativity levels – Context 1	High 53.3%, Medium 26.7%, Low 20%	High 5%, Medium 50%, Low 45%
Creativity levels – Context 2	High 46.9%, Medium 28.1%, Low 25%	High 10%, Medium 45%, Low 45%
Context effect on creativity (stats)	No significant differences between contexts ($z = -0.90$; $p = 0.371$)	No significant differences between contexts ($p = 0.9$)

Conclusions

This study addressed two objectives that were examined systematically through the results.

Objective O1 (Context Effects): We confirm Passarella's (2022) core finding of no significant context effect on creativity levels. Neither the Spanish sample ($z = -0.90$, $p = .371$) nor the Italian sample ($p = .9$) showed statistically significant differences in creativity distributions between real-mathematical and real-life contexts. However, both studies show stronger stimulus connection in real-life contexts (Spanish: 93% vs 52%; Italian: 100% vs 8%), suggesting that contextual authenticity facilitates problem-context linkage regardless of creativity outcomes.

Objective O2 (Problem Definition): This study reformulates the concept of "mathematical problem" by adopting three theory-grounded criteria absent from Passarella's (2022) inclusive approach (context + defined unknown): cognitive challenge requiring non-routine reasoning, closed structure with all data for a verifiable solution, and contextualisation providing meaningful school or everyday anchoring. This reformulation reduced "problems" from 80.5% (Passarella criterion) to 44.1% (theory-driven criteria), a statistically significant difference (χ^2 , $p < .001$), showing that analytical framing fundamentally shapes MPP creativity research outcomes.

Theoretical implications: creativity indices in MPP research are highly sensitive to definitional choices, underscoring explicit construct operationalisation needs. Contextual authenticity aids stimulus engagement but does not guarantee creative diversity, as students' capacity to reinterpret given situations appears more determinative.

Practical implications: real-life artefacts strengthen problem–context connections, supporting their use in MPP tasks, and structured MPP activities that require transforming given stimuli can foster the reinterpretation skills associated with higher creativity levels. Overall, these results advance construct clarity in MPP research while confirming the robustness of Passarella's (2022) null context effect and positioning reinterpretation capacity as a promising direction for future creativity studies.

Acknowledgements

This work has been carried out within the framework of regional project CIGE/2024/64 of the Regional Ministry of Education, Culture, Universities and Employment of the Valencian Regional Government.

References

- Albarracín, L., & Gorgorió, N. (2019). Using large number estimation problems in primary education classrooms to introduce mathematical modelling. *International Journal of Innovation in Science and Mathematics Education*, 27(2), 45-57. <https://doi.org/10.30722/IJISME.27.02.004>
- Bevan, D., Williams, A. M., & Capraro, M. M. (2019, August). Strike a pose: The impact of problem-posing on elementary students' mathematical attitudes and achievement. In *International symposium elementary mathematics teaching* (pp. 80-88).
- Bonotto, C. (2013). Artifacts as sources for problem-posing activities. *Educational studies in Mathematics*, 83, 37-55. <https://doi.org/10.1007/s10649-012-9441-7>
- Brown, S. I. & Walter, M. I. (2005). *The art of problem posing*. (3rd ed.). Lawrence Erlbaum Associates.
- Cai, J., Hwang, S., Jiang, C., y Silver, S. (2015). Problem-posing research in mathematics education: Some answered and unanswered questions. *Mathematical problem posing: From research to effective practice*, 3-34. https://doi.org/10.1007/978-1-4614-6258-3_1
- Carrillo, Y. (1996). *La resolución de problemas en la didáctica de las matemáticas*. Universidad de Granada.
- Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). Routledge.
- English, L. D. (1997). Promoting a problem-posing classroom. *Teaching Children Mathematics*, 4(3), 172–179.
- Espinoza, J., Lupiáñez, J. L., & Segovia, I. (2014). La invención de problemas y sus ámbitos de investigación en educación matemática. *Revista Digital: Matemática, Educación e Internet*, 14(2), 1-12. <https://doi.org/10.18845/rdmei.v14i2.1664>
- Freudenthal, H. (1991). *Revising mathematics education: Thoughts and realities*. Kluwer Academic Publishers.
- Guilford, J.P. (1959) Traits of Creativity. In: Anderson, H.H., Ed., *Creativity and Its Cultivation*, Harper and Row, New York, 142-161.
- Kilpatrick, J. (1987). Problem formulating: Where do good problems come from? In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 123–147). Lawrence Erlbaum Associates.
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academies Press.
- Leung, S., & Silver, E.A. (1997). The role of task format, mathematics knowledge, and creative thinking on the arithmetic Problem Posing of prospective elementary school teachers. *Mathematics Education Research Journal*, 9(1), 5-24. <https://doi.org/10.1007/BF03217299>
- Palm, T. (2006). Word problems as simulations of real-world situations: A proposed framework. *For the learning of mathematics*, 26(1), 42-47.
- Passarella, M. (2022). Real Contexts In Problem-Posing: An Exploratory Study Of Students' Creativity. *International Journal of Innovation in Science and Mathematics Education*, 30(1), 15–29. <https://doi.org/10.30722/IJISME.30.01.002>
- Polya, G. (1981). *Mathematical discovery: On understanding, learning, and teaching problem solving* (Combined ed.). John Wiley & Sons.
- PSPP (2022). GNU PSPP (v1.6.2) [software]. <http://www.pspp.org/>
- Rodríguez, F. (2012). *Resolución de problemas matemáticos: Fundamentos teóricos y prácticas de aula*. Narcea.
- Russo, J., Bobis, J., Downton, A., Hughes, S., Livy, S., McCormick, M., & Sullivan, P. (2020). Students who surprise teachers when learning mathematics through problem solving in the early primary years. *International Journal of Innovation in Science and Mathematics Education*, 28(3), 14-23. <https://doi.org/10.30722/IJISME.28.03.002>
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334–370). Macmillan.
- Silver, E. A. (1994). On mathematical problem posing. *For the learning of mathematics*, 14(1), 19-28. <https://www.jstor.org/stable/40248099>
- Stoyanova, E., & Ellerton, N. F. (1996). A framework for research into students' problem posing. In P. Clarkson (Ed.), *Technology in mathematics education* (pp. 518–525). Melbourne, Australia: Mathematics Education Research Group of Australasia.
- Torrance, E. P. (1966). Nurture of creative talents. *Theory Into Practice*, 5(4), 167–173. <https://doi.org/10.1080/00405846609542020>

- van den Heuvel-Panhuizen, M., & Drijvers, P. (2014). Realistic mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 521–525). Springer.
- Xie, J., & Masingila, J. O. (2017). Examining interactions between problem posing and problem solving with prospective primary teachers: A case of using fractions. *Educational Studies in Mathematics*, *96*, 101-118. <https://doi.org/10.1007/S10649-017-9760-9>