Educational Studies in Mathematics (2026) 122:487–510
https://doi.org/10.1007/s10649-026-10481-w

Does problem connection to reality shape interest, performance and their interplay? A comparative analysis involving students with high and average mathematical potential

Irene Ferrando[1] · **Carlos Segura**[1] · **Stanislaw Schukajlow**[2]

Received: 17 April 2025 / Accepted: 13 January 2026 / Published online: 10 February 2026
© The Author(s) 2026

Abstract
Interest and performance are important learning outcomes. However, little is known about how interest, performance and their relationship depend on students' mathematical potential and the problems' connection to reality. The problems' connection to reality ranges from intramathematical problems through word problems to modelling problems. The research questions of the study are: Does the performance gap between students with high and average mathematical potential vary with problem type? Does the interest gap between students with high and average mathematical potential vary with problem type? Does the relationship between interest and performance depend on the problems' connection to reality and the student profile? Ninth-grade students with high and average mathematical potential ($N = 126$) solved intramathematical, word and modelling problems and rated their interest regarding these problems. In modelling problems, the difference in performance between the two students' groups tended to be smaller than in intramathematical problems. Furthermore, in modelling problems, the difference in interest is higher than in word problems. The connection between interest and performance is comparable in students with high and average mathematical potential across intramathematical, word and modelling problems. This study's theoretical contribution to the theory of interest confirms the stability of the relationship between interest and performance across different types of problems. One practical implication is to offer students with high mathematical potential modelling problems in the classroom, as these problems, even when they can be solved by standard mathematical procedures, might be interesting for them.

Keywords Problem solving · Word problems · Modelling · High mathematical potential students · Performance · Interest

✉ Irene Ferrando
irene.ferrando@uv.es

Carlos Segura
carlos.segura@uv.es

Stanislaw Schukajlow
schukajlow@uni-muenster.de

[1] Department of Mathematics Education, Universitat de València, Valencia, Spain

[2] Institute of Mathematics Education and Computer Science Education, University of Münster, Münster, Germany

1 Introduction

Problem solving has long been central to mathematics education (Liljedahl & Cai, 2021). Since Polya's work (1945), there have been advances in understanding the cognitive and affective aspects of problem solving (Lester & Cai, 2016; Schoenfeld, 1985, 1992, 2013). Students who perform highly in mathematical problem solving typically excel in higher-order processes such as representation, abstraction and generalisation (Leikin, 2020; Sriraman, 2003).

Motivation and emotions are recognised as foundational factors, mediators and outcomes in the learning process (Schukajlow et al., 2023). This study focuses on the motivational construct of interest, a psychological state reflecting a positive emotional connection between an individual and an object or topic (Hidi & Renninger, 2006). Interest fosters deeper engagement in problem solving (Renninger & Hidi, 2015).

Mathematical problems vary in their connection to reality, ranging from abstract problems to those artificially linked to a real context, or directly connected to reality (Niss et al., 2007; Wijaya et al., 2014). These differences influence problem-solving processes (Niss et al., 2007; Rellensmann & Schukajlow, 2017). Problems linking mathematics to reality, like modelling problems, are widely considered essential for developing mathematical competencies (Maaß et al., 2019). Consequently, curricula increasingly promote real-world connections (Cevikbas et al., 2022). Spain's recent curriculum explicitly includes modelling as a core competence in compulsory mathematics education (MEFP, 2022). However, it is essential to consider the benefits of prioritising such problems, given students' diverse interests and cognitive profiles.

The influence of connecting mathematical problems to real life is particularly notable among students with high mathematical potential, who often show increased engagement with realistic tasks. Such real-world problem-solving tasks sustain their interest and allow them to apply their advanced abilities in meaningful ways (Leikin & Levav-Waynberg, 2007), fostering cognitive and affective development (Sriraman, 2003).

This study compares two distinct student groups. The first consists of students participating in a mathematics talent development programme who have been selected for their high performance in solving complex problems, henceforth referred to as high mathematical potential students (high-MPSs). The second group comprises students following the standard curriculum, hereafter referred to as students with average mathematical potential (average-MPSs). The designation "student profile" denotes this distinction, reflecting differences in educational opportunities rather than innate ability.

Although research has examined interest and performance in mathematics learning, little is known about how problem type influences these constructs across profiles. Moreover, the interrelation of interest and performance across problem types remains unexplored, particularly comparing high-MPSs and average-MPSs. While problems connected to reality, such as modelling tasks, can enhance engagement (Schulze Elfringhoff & Schukajlow, 2021), it is unclear whether this effect differs between students in talent programmes and their peers in regular classrooms. This study fills these gaps by examining how the problem's connection to reality influences both performance and interest, and whether the relationship between interest and performance is stable across student profiles and problem types.

2 Theoretical framework

2.1 Problem-solving performance

Success in mathematical problem solving depends on two main competencies: mastery of relevant techniques and their strategic use (Schoenfeld, 1982). The first condition pertains to domain-specific knowledge (Mayer, 1992). For example, solving a problem involving the Pythagorean Theorem requires recognising a right triangle, its hypotenuse and the relationship between sides. However, content knowledge alone is insufficient. Strategic use involves planning how to apply these techniques effectively—deciding which triangle to analyse first, whether to compute side lengths directly or in steps, and monitoring the validity of the approach. Research on problem-solving expertise emphasises the need for a coherent strategy and implementation (Mayer, 1998; Schoenfeld, 1982; Segura & Ferrando, 2023).

Problem features can further influence the complexity of strategies required, influencing overall performance. This underscores the need to consider problem characteristics in the present study.

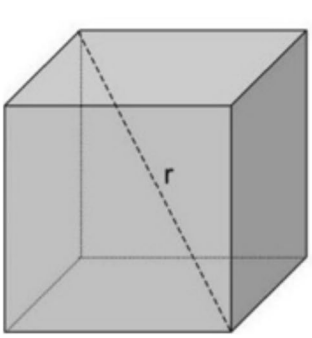
2.1.1 Influence of connection to reality in problem-solving performance

Problem-solving performance depends not only on domain knowledge and strategic skills but also on the problem's connection to reality (García-Cerdá et al., 2024; Niss et al., 2007).

Intramathematical problems (Intra-Ps) are purely mathematical, aimed at practising mathematical procedures or reinforcing core concepts. They provide all necessary data for direct calculation, sometimes involving non-routine challenges that require novel strategies (Elia et al., 2009; Pantziara et al., 2009; Zawojewski, 2013). In Fig. 1, we show an example of an intramathematical problem involving the Pythagorean Theorem.

Word problems (Word-Ps) incorporate real-world elements, requiring students to translate them into mathematical language, which adds a level of interpretation not needed in Intra-Ps. Even if word problems' statements include a real context, their relationship to reality is often artificial and structured to facilitate straightforward calculations (Niss et al., 2007; Rellensmann & Schukajlow, 2017), as we see in Fig. 2. Solving word-problems involves "undressing" the situation in the statement to extract mathematical elements and apply mathematical procedures (Copur-Gencturk & Doleck, 2021; Pollak, 1969).

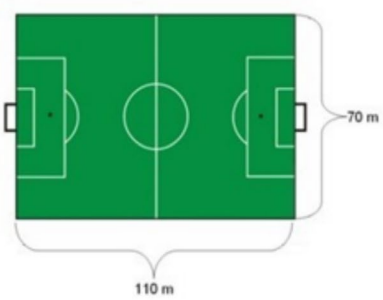
Modelling problems (Mod-Ps) require creating a mathematical model of a real situation by making assumptions, quantifying, applying procedures and interpreting results (Blum



Calculate the length of the diagonal of the cube whose sides measure 5 cm.

r =.....

Fig. 1 An example of one of the intramathematical problems included in the questionnaire



The coach of the Valazul team would like to run a race on the diagonal of the football field with his team.

To do this he needs to know how long the diagonal is.

Can you help him?

Calculate the diagonal of the football field.

Fig. 2 An example of one of the word problems included in the questionnaire

& Leiss, 2007). Unlike word problems, which provide all necessary data for direct calculation, Mod-Ps present open-ended contexts that demand realistic assumptions (Galbraith & Stillman, 2006; Krawitz et al., 2018). This distinction aligns with the Spanish curriculum (MEFP, 2022), which defines modelling as a process of mathematising real situations through interpretation and representation.

For example, in Fig. 3, we show a modelling problem involving a zip line. To solve this problem, it is crucial to assume a slope for the zip line and account for additional cable length at both ends for anchoring. First, estimate the slope and, assuming points A and B are at the same height, use the Pythagorean theorem or scaling drawing to calculate the distance between them (interpreting the measures in the image is key). Next, deduce the height of the highest point of the zip line and, using the Pythagorean theorem again, calculate its length. Finally, add extra length to account for cable anchoring at stations A and B.

The Park managers are going to put up a zip line in the Viveros Gardens linking points A and B. In the following picture we provide you with the information from GoogleMaps, can you estimate the length of the cable?



Fig. 3 An example of one of the modelling problems included in the questionnaire

Table 1 Key distinctions between intramathematical, word and modelling problems

	Contextual features	Information provided	Requires assumptions
Intramathematical (Intra-P)	Purely mathematical	All data provided	No
Word (Word-P)	Real-world context, simplified	All data provided	No
Modelling (Mod-P)	Real-world context, open ended	Partial or implicit data	Yes

While modelling problems involve the application of mathematical tools like the Pythagorean theorem, their open-endedness and realistic assumption making differentiate them from word problems, which provide all data explicitly.

Table 1 summarises the key distinctions between intramathematical, word and modelling problems.

Empirical studies confirm that problem type affects performance. Rellensmann and Schukajlow (2017) found German ninth graders performed better on real-world problems (M = 0.62) than on intramathematical problems (M = 0.49). Liebendörfer and Wagner (2024) found that German eighth graders showed better performance on Word-Ps than on Intra-Ps or Mod-Ps. In Spain, García-Cerdá et al. (2024) observed higher success on Word-Ps (40.6% correct answers) than on Intra-Ps (23.1%) and Mod-Ps (13.8%), with significant differences (χ^2 (2) = 43.61, $p < .01$) among ninth-grade students. These findings suggest that Word-Ps, despite their artificial context, are more accessible and familiar to students, while Mod-Ps pose greater challenges due to their open-endedness and assumption making.

2.1.2 High mathematical potential students' problem-solving performance

Students with high mathematical potential (high-MPSs) exhibit advanced cognitive abilities, such as strong mathematical memory, flexible reasoning and effective abstraction and generalisation skills (Leikin, 2020). They quickly recognise patterns and structures, drawing on their ability to recall relevant mathematical knowledge and manage diverse strategies (Gutiérrez et al., 2018; Leikin, 2020). Case studies show that high-MPSs attain high levels of cognitive demand and efficiency in solving rich mathematical tasks (Benedicto et al., 2018).

Sriraman (2003) observed that high-MPSs often approach problems by formulating generalisations and demonstrating creative thinking, essential for deep mathematical understanding. Leikin et al. (2016) found that high-MPSs engage in dual processing, handling multiple information sets for more efficient problem solving. These findings underscore the need for tailored educational approaches to challenge and nurture their abilities (Burney, 2008; Gutiérrez & Jaime, 2021).

Several studies highlight that high-MPSs not only excel in traditional problem-solving settings but also adapt and act creatively across different mathematical contexts (Benedicto et al., 2018; Gutiérrez et al., 2018; Leikin, 2019; Leikin & Sriraman, 2022). However, comprehensive comparative research on how these skills transfer across different problem types remains scarce (Leikin, 2019; Leikin & Sriraman, 2022).

2.2 Interest in problem solving

Interest in mathematical problem solving is understood as a psychological state reflecting a learner's positive emotional connection to the task, varying in intensity and influencing engagement (Harackiewicz et al., 2016; Hidi & Renninger, 2006). Rather than a binary state, interest operates on a continuum, shaping the depth and persistence of involvement.

Interest theory distinguishes situational interest—temporary engagement triggered by task features—from individual interest, a stable disposition reflecting long-term personal preference (Hidi & Renninger, 2006). Situational interest is particularly influential in mathematical problem solving, as it can motivate engagement with complex tasks regardless of prior interest in the subject. Schukajlow et al. (2012) emphasise that interest, particularly situational interest during problem solving, is both an outcome and a mediator that enhances cognitive engagement, supporting the transition to enduring individual interest (Hidi & Renninger, 2006).

Task-specific interest denotes a state-like situational interest arising during problem solving (Rellensmann & Schukajlow, 2017; Schukajlow et al., 2012) and refers to specific task types rather than to mathematics in general (Schukajlow et al., 2023). Students experiencing higher task-specific interest demonstrate greater concentration and employ more sophisticated problem-solving strategies, and persist longer on challenging problems, improving learning outcomes (Ainley et al., 2002; Rellensmann & Schukajlow, 2017).

Interest in problem solving is shaped by task-related features, such as the problem's connection to real life, that increase relevance and engagement (Schulze Elfringhoff & Schukajlow, 2021), and by individual factors, including prior knowledge, self-efficacy and personality traits (Geisler & Rach, 2023; Ryan & Deci, 2000).

2.2.1 Influence of connection to reality in problem-solving interest

The real-world connection of a mathematical problem can influence students' interest in varied and sometimes unpredictable ways. Realistic contexts often aim to enhance task-specific interest by making mathematics more relevant (Borromeo Ferri, 2018; Gravemeijer & Doorman, 1999; Niss & Blum, 2020).

However, empirical findings challenge the assumption that real-world connections always boost interest. Rellensmann and Schukajlow (2017) found that German secondary school students reported greater interest in intra-mathematical problems than in real-world problems (including modelling and word problems), with a significant difference, $t(99) = -2.96$, $p < .01$, $d = 0.30$. Their study involved 12 Pythagoras-based problems (4 intramathematical, 8 real-world), with interest rated after each task. The authors note that modelling steps (e.g. assumption making) may increase cognitive load (Blum & Leiss, 2007). Similarly, Krawitz and Schukajlow (2018) noted that students valued Mod-Ps less than Word-Ps or Intra-Ps. Other studies have found no significant differences across problem types (Schukajlow et al., 2012).

Liebendörfer and Wagner (2024) observed higher interest in Word-Ps compared to Intra-Ps or Mod-Ps, with medium effects ($p < .001$; $d = 0.62$ and 0.68, respectively). This suggests that tasks with accessible, superficial real-world context may better capture situational interest than deeply contextualised or purely abstract problems.

Preservice teachers tend to overestimate students' interest in Mod-Ps while underestimating their interest in Word-Ps and Intra-Ps (García-Cerdá et al., 2024; Liebendörfer

& Wagner, 2024), challenging assumptions about intrinsic motivation toward real-world problems (Gravemeijer & Doorman, 1999).

These findings highlight the need to consider individual differences. While Mod-Ps may engage some students through real-world relevance, others may prefer abstract problems that allow creative exploration. Students' mathematical potential and other learner characteristics likely moderate these effects.

2.2.2 High mathematical potential students' problem-solving interest

High-MPSs tend to engage more when tasks match their cognitive strengths, fostering exploration, creativity and challenge. Krutetskii (1976) noted their preference for problems allowing flexibility, abstraction and generalisation, while routine tasks are perceived as unchallenging (Leikin, 2020). Similarly, Sriraman (2003) observed greater interest in problems with multiple solutions or open-ended questions that offer opportunities for creative engagement and intellectual satisfaction (Leikin, 2009).

Recent perspective emphasises designing problems that require original thinking to spark high-MPSs' interest, especially in competitive contexts where creativity is valued (Freiman, 2006; Geretschläger & Donner, 2022). Mod-Ps have been identified as particularly potent for developing and identifying high-MPSs due to their open-endedness and real-world relevance (Chamberlin & Moon, 2005).

However, few studies have examined how the connection to reality influences high-MPSs' interest. Subotnik et al. (2023) argue that incorporating real-world contexts helps maintain high-MPSs' interest by linking abstract mathematics to practical outcomes. Brandl (2011) describes modelling problems tailored to high-MPSs' preferences. Yet, no research has directly compared high-MPSs' interest in Mod-Ps versus Word-Ps or Intra-Ps.

2.3 Relationship between interest and performance in problem solving

Research indicates that interest in problem solving is positively related to short-term performance and long-term achievement in mathematics. Schiefele et al. (1992) found that students with a higher interest in mathematics performed better and used deeper processing strategies, like linking new information to prior knowledge. A recent meta-analysis confirms that interest predicts achievement, persistence and satisfaction (de Vries et al., 2024), with persistence being crucial for success in complex tasks.

Students with higher interest tend to self-regulate more effectively, monitoring progress and adjusting strategies (Ainley et al., 2009). Rotgans and Schmidt (2011) suggest that situational interest needs engagement to influence performance. Thus, fostering interest supports both cognitive and affective development.

The relationship between situational interest and performance is bidirectional: perceptions of performance can trigger situational interest, and interest can enhance performance (Nuutila et al., 2020). Task characteristics and student profiles moderate this relation.

Empirical studies find that this relationship varies depending on the problem type. García-Cerdá et al. (2024) found significant positive correlations between task-specific interest and performance across problem types for secondary students ($\rho = .25$ for Intra-Ps; $\rho = .31$ for Mod-Ps). These results suggest that interest enhances cognitive engagement and strategic effort, particularly in more demanding tasks.

Other studies support this pattern. Rellensmann and Schukajlow (2017) found that interest depends on perceived difficulty, and that real-world contexts did not enhance interest

unless they were personally meaningful. Liebendörfer and Wagner (2024) observed higher interest in tasks with familiar and relevant contexts, while complex modelling tasks elicited lower interest despite deeper real-world connections.

Overall, the strength of the interest–performance link depends on initial interest, task relevance and problem complexity (Ainley et al., 2009; Nuutila et al., 2020). This relationship may vary depending on how well tasks align with students' individual interests and performance expectations. Situational interest increases when task contexts are personally relevant (Liebendörfer & Wagner, 2024), though how context familiarity connects to detailed preferences remains unclear (Schoenherr, 2024). When tasks do not match students' initial interests, the link between performance and situational interest can be weak (Jansen et al., 2016). For instance, highly motivated students may struggle to translate individual interest into task-specific interest or performance when faced with unfamiliar open-ended problems, while showing less difficulty with more structured tasks such as Word-Ps or Intra-Ps.

Prior performance also shapes individual interest over time; tasks that are either too challenging or too simple may dampen situational interest, thereby affecting performance (Wigfield & Eccles, 2000). The link between situational interest and performance, especially concerning cognitive factors such as students' profiles, remains underexplored.

3 Research question and hypotheses

Previous research shows that problem types differ considerably in their cognitive demands and motivational impact (Borromeo Ferri, 2018; Rellensmann & Schukajlow, 2017). Students with high mathematical potential (high-MPSs) and students with average mathematical potential (average-MPSs) display distinct engagement patterns, with high-MPSs often seeking complexity and novelty (Krutetskii, 1976; Leikin, 2009, 2020; Sriraman, 2003). However, few studies examine how problem type and student profile jointly influence interest and performance, or whether the interest–performance link remains stable across these variables (Ainley et al., 2002; Nuutila et al., 2020). Integrative research combining cognitive and motivational perspectives remains scarce (Middleton et al., 2017; Schukajlow et al., 2012).

This study addresses this gap by examining the role of problem type (degree of real-world connection) and student profile (high-MPSs vs. average-MPSs) on interest, performance and their interplay. The guiding research questions and hypotheses are:

RQ1: Does the performance gap between students with high and average mathematical potential vary depending on the problems' connection to reality?
H1: High-MPSs are expected to excel in tasks that challenge abstract thinking (Krutetskii, 1976; Sriraman, 2003). Mod-Ps, involving real-world contexts and assumption making (Chang et al., 2020), may narrow the performance gap, as both groups face unfamiliar demands. Word-Ps simplify real-world contexts into single-solution tasks (Pollak, 1969). High-MPSs are likely to outperform average-MPSs most in Intra-Ps, which require deep conceptual understanding (Star, 2005). We hypothesise that the performance gap will be greater in Intra-Ps than in Word-Ps and Mod-Ps. Since Mod-Ps are the most complex and Word-Ps the simplest (García-Cerdá et al., 2024), the performance gap on Word-Ps is not expected to be significantly different from the performance gap on Mod-Ps.

RQ2: Does the interest gap between students with high and average mathematical potential vary depending on the problems' connection to reality?
H2: Mod-Ps are expected to spark greater interest among high-MPSs by offering creative challenges (Leikin, 2009; Sriraman, 2003), whereas Intra-Ps and Word-Ps—common in Spanish textbooks (Vicente et al., 2018)—may elicit lower interest from high-MPSs due to perceived repetition (Leikin, 2020). Average-MPSs, who may experience similar levels of interest across problem types, generally exhibit lower interest in Mod-Ps (Liebendörfer & Wagner, 2024). Therefore, the interest gap is hypothesised to be larger in Mod-Ps than in Word-Ps and Intra-Ps, with no differences between the interest gaps for Word-Ps and Intra-Ps.
RQ3: Does the relationship between interest and performance depend on the problems' connection to reality and/or the student profile?
H3: Previous research shows that task-related interest and performance are positively related in both directions: interest in problem solving triggers engagement that can foster better student performance (Rotgans & Schmidt, 2011); but also, performance (and then, competence perception) is related to self-efficacy, which affects situational interest (Nuutila et al., 2020). Moreover, the mechanisms of these effects (throughout engagement and self-efficacy) are stable across types of problems and individuals (Ainley et al., 2002; Schiefele et al., 1992). This relationship is anticipated to hold consistently across problem types and student groups. Both high-MPSs and average-MPSs are expected to show positive, stable correlations between interest and performance within each problem type, and these correlations are hypothesised to be similar when compared between groups.

4 Methodology

4.1 Sample and data collection

The study involved two groups of ninth graders from Valencia (Spain). The first group (average-MPSs; $N=81$, mean age 15.23) followed the standard curriculum without identified special needs. Schools were selected based on accessibility and prior collaboration. All average-MPSs participated as intact classes during regular mathematics lessons. Although they had different teachers, all confirmed that the Pythagorean Theorem had been taught 1–3 months before data collection, which was conducted during the spring term of each academic year.

The second group (high-MPSs; $N=45$, mean age 15.05) comprised students who, in addition to regular schooling, attended Estalmat, an extracurricular programme selecting 25 students annually via a test that includes complex non-routine problems (Benyhe, 2011; Hernández, 2006). The Estalmat programme enriches but does not replace the standard curriculum.

Data were collected during the 2021–2022 and 2022–2023 academic years in two 60-min sessions over consecutive weeks. One researcher was present with the students' math teacher (average-MPSs) or the Estalmat coordinator (high-MPSs).

The instrument, adapted from Rellensmann and Schukajlow (2017), is based on the version described in García-Cerdá et al. (2024). Adaptations included translation to Spanish, minor cultural adjustments, and the selection of 12 problems, four of each type: Intra-P, Word-P, Mod-P (see examples in Figs. 1, 2 and 3). The full set of problems is provided

in Supplementary Material S2. All the problems can be solved by using the Pythagorean Theorem, a topic already studied by all participants (MEFP, 2022). Using a single topic controlled for content variability, ensuring differences were attributable to problem type.

4.2 Measures and data analysis

After solving each problem, the participants rated their interest using a 5-point Likert scale (1 = "not true at all", 5 = "completely true") on the statement "It was interesting to solve this problem." Internal consistency was satisfactory for Mod-Ps ($\alpha = 0.82$), Word-Ps ($\alpha = 0.91$) and Intra-Ps ($\alpha = 0.84$), comparable to or higher than Rellensmann and Schukajlow (2017).

Problem-solving performance was rated 0–2 by problem. Following the same procedure as the one used in Segura and Ferrando (2023), we have differentiated whether the solvers can find a valid solving strategy (Copur-Gencturk & Doleck, 2021) (2 points), whether they make errors when developing the strategy (1 point), or were not able to develop any valid solving strategy (0 points). Scores were averaged per problem type. Table 2 illustrates coding criteria with examples.

The first author rated all solutions, and the second author rated 10% of the solutions independently. The disagreements were discussed. Interrater agreement (Cohen's $\kappa \geq .86$) was good in all problem types. Internal consistency for performance was satisfactory ($\alpha = .82–.86$ across problem types). Both interest ratings and performance scores were treated as continuous data, justified by adequate internal consistency, aggregated measurement across multiple items, and the robustness of parametric tests to ordinal data (Norman, 2010).

To test RQ1 and RQ2, we used mixed ANOVA analysis, with one within-subject factor (problem type) and one between-subject factor (student profile). This method tests both main and interaction effects, revealing whether the differences between high-MPSs and average-MPSs vary across problem types. Main and interaction effects are discussed in terms of their theoretical and practical relevance for understanding differences between student profiles and problem types (Bakker et al., 2019; Cheung & Slavin, 2016; Simpson, 2020).

For RQ3, we computed correlations, applied Fisher's Z transformation for group comparisons, and used Meng's Z test for within-group differences (Meng et al., 1992). The full correlation matrix is provided in Supplementary Table S1.

Table 2 Performance scoring criteria

Score	Description	Example
2 points	The student correctly identifies and applies a valid strategy, leading to a correct solution	In a Mod-P (Fig. 3), the student estimates a realistic slope, interprets the image correctly, applies the Pythagorean theorem, and adds the extra cable length to calculate the total
1 point	The student identifies a valid strategy but makes calculation or interpretation errors that prevent arriving at the correct solution	The student sets up the triangle correctly and applies the theorem, but miscalculates the square root or omits the final step of adding the extra cable length
0 points	The student does not apply a valid strategy or provides an incomplete or irrelevant response	The student writes unrelated formulas, misidentifies the triangle, or leaves the problem blank

A sensitivity analysis (G*Power, $\alpha = .05$, $1 - \beta = .80$) indicated that with $N = 126$, the minimum detectable effect size for the interaction was $f = 0.28$ (partial $\eta^2 \approx .073$), a medium-to-large effect (Cohen, 1988). Therefore, the study was adequately powered to detect medium-to-large effects but underpowered for small effects, which we acknowledge as a limitation.

5 Results

5.1 Influence of the problem type on the differences in performance between students with high and average mathematical potential

Table 3 shows mean performance scores for high-MPSs and average-MPSs for Intra-Ps, Word-Ps and Mod-Ps.

High-MPSs consistently outperformed average-MPSs, with the largest gap in Intra-Ps (1.12 points) and the smallest in Word-Ps (0.91). Both samples showed the highest performance in Word-Ps and the lowest in Mod-Ps (Fig. 4).

Table 3 High-MPSs and average-MPSs mean performance score (0–2) in each problem type

	Intramathematical (Intra-P)*M (SD)*	Word (Word-P) *M (SD)*	Modelling (Mod-P) *M (SD)*
Average-MPSs	0.52 (0.59)	0.83 (0.75)	0.39 (0.52)
High-MPSs	1.64 (0.48)	1.74 (0.44)	1.37 (0.48)

M, mean; *SD*, standard deviation

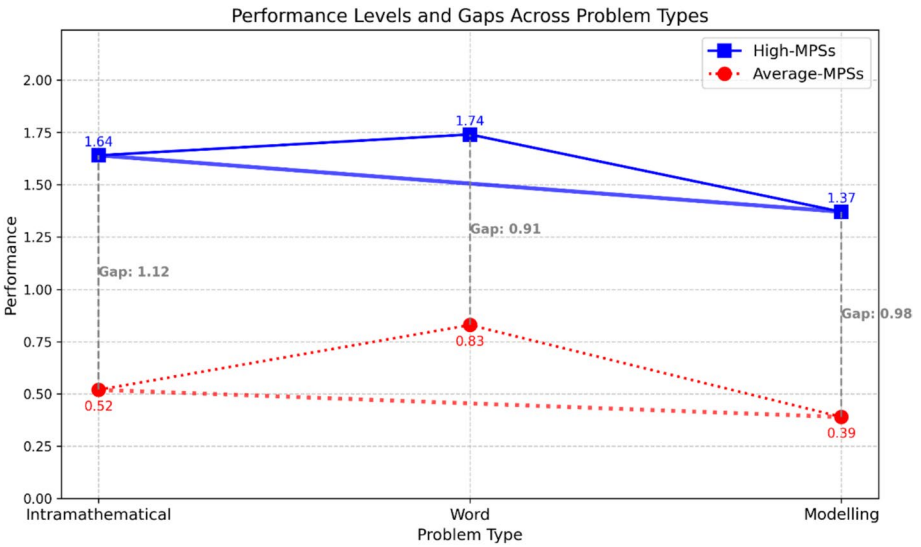


Fig. 4 Comparison of high-MPSs' performance and average-MPSs' performance across problem types

To test H1, we conducted three pairwise ANOVA comparisons, with the between-subject factor "student group" (high-MPS or average-MPS) and one within-subject factor "problem type" (Word-P or Mod-P, Intra-P or Mod-P and Intra-P or Word-P).

The first comparison examined Word-Ps and Mod-Ps. The main effect of problem type was significant (F (1, 124) = 69.63, $p < .001$, $\eta^2 = .36$), indicating that both groups performed better on Word-Ps than on Mod-Ps. However, the interaction was not significant (F (1, 124) = 0.46, $p = .25$, $\eta^2 = .004$), showing that the difference between the performance gaps for Word-Ps (0.91) and Mod-Ps (0.98) was negligible.

The second comparison contrasted Intra-Ps and Mod-Ps. The main effect was significant (F (1, 124) = 17.56, $p < .001$, $\eta^2 = .12$), confirming that performance was significantly higher on Intra-Ps than on Mod-Ps overall. The interaction approached significance (F (1, 124) = 2.34, $p = .06$, $\eta^2 = .02$), suggesting a tendency toward a larger gap for Intra-Ps (1.12) than for Mod-Ps (0.98).

Finally, the comparison between Intra-Ps and Word-Ps revealed a significant main effect of problem type (F (1, 124) = 14.04, $p < .001$, $\eta^2 = .10$), indicating that both groups performed better on Word-Ps than on Intra-Ps. The interaction was also significant (F (1, 124) = 3.53, $p < .05$, $\eta^2 = .03$). This finding confirmed that while high-MPSs may generally perform better across both problem types, they outperform average-MPSs strongly in Intra-Ps.

Following recommendations by Simpson (2020) and Bakker et al. (2019), we interpret these effects in context: interaction effects were small (partial $\eta^2 \approx .02$–$.03$) and the Word-P–Mod-P contrast was trivial ($\approx .004$), indicating that while the performance gap is consistently largest in Intra-Ps, the practical difference between Mod-Ps and Word-Ps is negligible.

5.2 Influence of the problem type on the differences in interest between students with high and average mathematical potential

Table 4 reports mean interest scores for high-MPSs and average-MPSs across Intra-P, Word-P and Mod-P.

High-MPSs reported higher interest than average-MPSs in all problem types, with the largest gap in Intra-Ps (0.88 points) and the smallest in Word-Ps (0.44) (Fig. 5).

To test H2, we conducted three pairwise ANOVA comparisons, with the between-subject factor "student group" (high-MPS or average-MPS) and one within-subject factor "type of problems" (Word-P or Mod-P, Intra-P or Mod-P and Intra-P or Word-P). We expected the interest gaps between high-MPSs and average-MPSs for Mod-Ps to be higher than for Word-Ps. Indeed, as shown in Fig. 5, the interest gap for Word-Ps is 0.44, while it is 0.83 for Mod-Ps.

Table 4 Mean interest in each type of problem

	Intramathematical (Intra-P) *M (SD)*	Word (Word-P) *M (SD)*	Modelling (Mod-P) *M (SD)*
Average-MPSs	2.69 (1.08)	2.92 (1.05)	2.74 (1.11)
High-MPSs	3.57 (0.87)	3.36 (1.17)	3.58 (0.78)

M = mean, *SD* = standard deviation

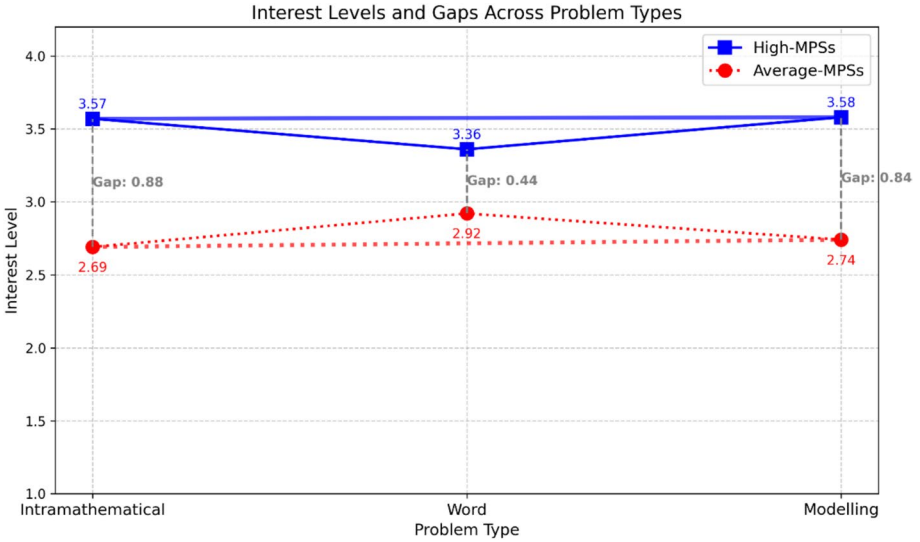


Fig. 5 Comparison of high-MPSs' and average-MPSs' interest across problem types

The first comparison examined Word-Ps and Mod-Ps. Although overall interest did not differ significantly between these types (F (1, 124) = .22, p = .64, η^2 = .002), the interaction was significant (F (1, 124) = 6.61, p < .01, η^2 = .05), indicating that the interest gap was lower for Word-Ps (0.44) than for Mod-Ps (0.83).

The second comparison contrasted Intra-Ps and Mod-Ps. The main effect of problem type was not significant (F (1, 124) = .50, p = .48, η^2 = .004), and the interaction was negligible (F (1, 124) = .08, p = .39, η^2 = .001), showing that the interest gaps for Intra-Ps (0.88) and Mod-Ps (0.83) were similar, contrary to our hypothesis.

Finally, the comparison between Intra-Ps and Word-Ps revealed no main effect of problem type on the interest of all students (F (1, 124) = .06, p = .82, η^2 = .00), but a significant interaction (F (1, 124) = 13.45, p < .001, η^2 = .10), which contradicts our expectation of equal gaps.

Following Simpson (2020) and Bakker et al. (2019), we interpret these effects in context: the Intra-P–Word-P interaction reflects a moderate effect (partial η^2 = .10), the Word-P–Mod-P interaction a small-to-moderate effect (partial η^2 = .05), and the Intra-P–Mod-P interaction is trivial (partial $\eta^2 \approx$ .001). Thus, while high-MPSs consistently reported higher interest, the relative size of the gap depends on problem type: it is largest for Intra-Ps, smallest for Word-Ps and similar for Mod-Ps and Intra-Ps.

5.3 Influence of the type of problem and the student profile on the relationship between interest and performance

Pearson's correlations were computed between interest and performance for each problem type and student group.

For average-MPSs, the relationship is significant and moderate for all problem types: intramathematical problems (Intra-Ps, $r=.30$, $p=.01$), modelling problems (Mod-Ps, $r=.40$, $p<.001$) and word problems (Word-Ps, $r=.26$, $p<.02$).

For high-MPSs, correlations were weak and not significant: Intra-Ps ($r=0.12$, $p=.43$), Mod-Ps ($r=0.23$, $p=.13$), for Word-Ps, there is almost no relationship between interest and performance ($r=0.04$, $p>.79$).

In Fig. 6, we present the dispersion plot with the regression lines for each type of problem in each sample.

Figure 6 shows that for average-MPSs, interest and performance tend to move together: those who found the problems more interesting generally achieved higher scores, and this pattern looks similar across all three problem types. For high-MPSs, however, the lines are almost flat, meaning that their performance does not seem to depend on how interested they felt in the problems. This qualifies H3: the relationship between interest and performance is stable across problem types in both groups, but it is clearly positive only for average-MPSs.

To examine whether the relationship varied by problem type within each group, we perform Meng's Z test. The results are presented in Table 5.

These results suggest that within each group, there are no statistically significant differences in the correlations between performance and interest across the different types of problems. Thus, the strength of interest–performance relationships does not significantly differ depending on the problems' context connection to reality.

To infer whether there are significant differences in the correlations between interest and performance for each problem, depending on the student group, the Fisher-Z transformation was applied, and the results are presented in Table 6.

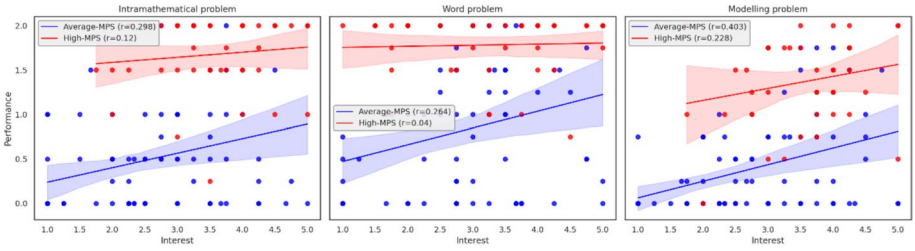


Fig. 6 Plots showing the relationship between interest (x-axis) and performance (y-axis) for each problem type. Each dot represents one student's paired scores (blue = average-MPSs, red = high-MPSs). Lines indicate regression trends for each group; shaded areas represent 95% confidence intervals

Table 5 Comparison of correlations within groups for different types of problems

	Average-MPSs			High-MPSs		
Comparison	Correlation	Z-value	p-value	Correlation	Z-value	p-value
Intramathematical vs word problems	0.30 vs 0.26	0.03	>.05	0.12 vs 0.04	0.06	>.05
Intramathematical vs modelling problems	0.30 vs 0.40	−0.09	>.05	0.12 vs 0.23	−0.08	>.05
Word vs modelling problems	0.26 vs 0.40	−0.12	>.05	0.04 vs 0.23	−0.14	>.05

Table 6 Comparison of Pearson's correlations between interest and performance in high-MPSs and average-MPSs

Problem type	Average-MPSs' correlation	High-MPSs' correlation	Fisher's Z	p-value
Intramathematical	0.30	0.12	0.98	0.33
Word	0.26	0.04	1.20	0.23
Modelling	0.40	0.23	1.02	0.31

We did not find differences in the correlation strengths between high-MPSs and average-MPSs across all problem types (all $p > .05$).

6 Discussion

This study aimed to deepen our understanding of how individual cognitive characteristics and problem types interact to affect students' performance and interest. The results have implications for designing tailored instructional strategies to support diverse learners.

6.1 Does the performance gap between students with high and average mathematical potential vary depending on the problems' connection to reality?

Both high-MPSs and average-MPSs demonstrated similar performance patterns: higher scores on Word-Ps and lower in Mod-Ps. However, the size of performance gaps—the difference between groups—varied by problem type.

As expected, there were no significant differences in the performance gaps between Word-Ps and Mod-Ps, as indicated by the non-significant interaction and negligible effect size (partial $\eta^2 = .004$). This likely reflects that both samples are more familiar with Word-Ps, common in Spanish textbooks (Vicente et al., 2018), whereas Mod-Ps remain rare in classrooms, despite their inclusion in the latest curriculum (MEFP, 2022). This suggests that the greater cognitive demands and interpretative requirements of Mod-Ps may reduce the advantage of high-MPSs over average-MPSs.

In contrast, Intra-Ps produced the largest gap, consistent with our hypothesis based on high-MPSs' strengths in handling purely mathematical tasks (Leikin, 2020; Sriraman, 2003). The difference between Intra-Ps' and Mod-Ps' gaps was small (partial $\eta^2 = .02$) and close to significance, suggesting that Mod-Ps' open-endedness and assumption making affect both groups similarly, reducing the gap. When comparing the performance gaps of Intra-Ps and Word-Ps, we confirm our expectation: the differences are close to medium (partial $\eta^2 = .03$) and significant, indicating that the high-MPSs outperform average-MPSs in Intra-Ps. While Intra-Ps, like Word-Ps, are familiar to all the students, Intra-Ps require interpreting abstract information and mobilising deep conceptual reasoning beyond the mere application of the Pythagorean theorem.

Following Bakker et al. (2019) and Simpson (2020), we interpret these effect sizes in context rather than relying solely on significance: the Word-P–Mod-P contrast is trivial (≈.004), the Intra-P–Mod-P difference is small (≈.02), and the Intra-P–Word-P difference is small-to-moderate (≈.03). Although modest, these interaction effects

are theoretically relevant because they reveal how problem characteristics interact with learner profiles to shape performance.

These findings align with previous studies. García-Cerdá et al. (2024) found that Spanish students performed significantly better on Word-Ps than on Intra-Ps and Mod-Ps, while Rellensmann and Schukajlow (2017) found German students performed better on real-world problems than on Intra-Ps. However, these studies did not differentiate student profiles, limiting their ability to capture the dynamics observed here.

Overall, problem type influences performance dynamics between high-MPSs and average-MPSs, supporting the notion that problem context affects mathematical performance (Leikin, 2013; Sriraman, 2003). Intra-Ps, which emphasise abstract reasoning, tend to amplify performance differences by leveraging the advanced cognitive skills of high-MPSs, whereas Mod-Ps reduce gaps by imposing unfamiliar demands on both groups.

This nuanced understanding is not captured in prior average-based studies (García-Cerdá et al., 2024). Our findings underscore the need to diversify problem types to equitably challenge and support learners, balancing abstract rigour with authentic contexts (Bakker et al., 2019; Middleton et al., 2017).

6.2 Does the interest gap between students with high and average mathematical potential vary depending on the problems' connection to reality?

The results (Fig. 5) show that high-MPSs and average-MPSs displayed different patterns of interest across the problem types, with notable variations in the size of the interest gaps (the difference between groups).

As expected, the interest gap is significantly larger for Mod-Ps than for Word-Ps (partial $\eta^2 = .05$). Word-Ps elicited the highest interest among average-MPSs and the lowest among high-MPSs. Although both Word-Ps and Mod-Ps include a real context, in Mod-Ps the context is functionally relevant (the solver needs to consider the characteristics of the real situation to build a model), whereas Word-Ps offer a predefined model and focus on routine procedures. Word-Ps' simplicity and prevalence in Spanish textbooks (Vicente et al., 2018) likely explain average-MPSs' higher interest (Schulze Elfringhoff & Schukajlow, 2021). Mod-Ps being demanding and rare in textbooks may attract students seeking intellectual challenge (Hidi & Renninger, 2006; Sriraman, 2003).

High-MPSs' low interest in Word-Ps may stem from recognising their repetitive procedural nature, with minimal contextual relevance (Palm, 2006; Schulze Elfringhoff & Schukajlow, 2021). For average-MPSs, however, even a superficial context can spark interest (Schulze Elfringhoff & Schukajlow, 2021). Exploring the sources of interest in problem statements based on the student profile could be a promising research direction.

These findings contrast with those of García-Cerdá et al. (2024), who found similar levels of interest across problem types in a homogeneous sample of secondary students, with only a slight preference for Word-Ps over Mod-Ps. Rellensmann and Schukajlow (2017) reported greater interest in Intra-Ps than in problems with a connection to reality, after controlling for difficulty. Neither study differentiated profiles, limiting their ability to capture the motivational dynamics observed here.

Contrary to our initial hypothesis, the interest gap between Intra-Ps and Mod-Ps did not differ (partial $\eta^2 = .001$), indicating that both problem types produce similar disparities. This pattern is explained by average-MPSs rating Intra-Ps as least interesting, while high-MPSs showed comparable interest in both. This has led to a widening of the interest gap in

Intra-Ps, approaching the interest gap in Mod-Ps. The abstract nature of the Intra-Ps may boost high-MPSs' interest (Leikin, 2020), whereas average-MPSs' lower self-efficacy may suppress theirs (Nuutila et al., 2020).

Unexpectedly, the interest gap was also significantly larger for Intra-Ps than for Word-Ps (partial $\eta^2 = .10$). This finding, combined with Fig. 5's minimal gap and opposite interest directionality for Word-Ps—highest interest among average-MPSs and lowest among high-MPSs—aligns with Krutetskii's (1976) seminal work on individual differences. The most important differences in interest emerge in cognitively demanding activities. High-MPSs show less interest in routine activities solvable through direct procedures, whereas Intra-Ps sustain their engagement by requiring deeper conceptual understanding. Unlike Word-Ps, which rely on procedural applications of the Pythagorean Theorem, Intra-Ps demand abstract reasoning and flexible thinking. For example, some Intra-Ps, such as identifying the correct placement of a right angle (see Fig. 1), involve conceptual knowledge beyond mere computation (Star, 2005).

Effect sizes interpreted in context (Bakker et al., 2019; Simpson, 2020) indicate: the Intra-P–Word-P interaction reflects a moderate effect (partial $\eta^2 \approx .10$), the Word-P–Mod-P interaction a small-to-moderate effect ($\approx .05$), and the Intra-P–Mod-P interaction is trivial ($\approx .001$). Although the overall main effect of problem type on interest was negligible (partial $\eta^2 < .01$), these interaction effects reveal that interest gaps are strongly shaped by problem features.

These results highlight an important pattern: while overall interest remains similar across problem types, the gap between profiles varies substantially with task characteristics. There appears to be a trade-off: tasks that reduce performance gaps (Mod-Ps) may simultaneously amplify interest gaps, while familiar, routine tasks (Word-Ps) engage average-MPSs but may discourage high-MPSs. This dynamic adds to prior studies that focused on interest without considering student heterogeneity.

Our findings underscore that interest is both task-dependent and profile-dependent. Designing mathematical problems that balance cognitive challenge with contextual relevance may be key to promoting equitable engagement. Such an approach could mitigate motivational disparities while preserving opportunities for deep learning, ensuring that neither group is systematically advantaged or disengaged.

6.3 Does the relationship between interest and performance depend on the type of problem and/or the type of student?

To address RQ3, we first obtained the correlation between interest and performance for each problem type and student profile. The graphs in Fig. 6 help clarify the differences in the relationship between interest and performance, both by problem type (within groups) and student profile (between groups). We observe that for average-MPSs, the regression lines have a steeper positive slope than for high-MPSs, suggesting a stronger link between interest and performance.

To identify sample-specific differences in correlations by problem type, we conducted Meng's Z-test for each sample (see Table 5). The comparison of correlation coefficients within groups confirms that, for all students, the relationship between interest and performance is positive and consistent across all problem types, validating our expectations.

For average-MPSs, interest and performance are strongly related across all problem types, meaning students who exhibit high (or low) interest or performance in one type tend

to maintain similar levels in the other. This is confirmed by Meng's *Z*-test results, which show no differences in the interest/performance correlation within the average-MPSs group across different tasks. We confirm that for average-MPSs, the positive relationship between interest and performance is strong and significant, as pointed out by previous studies (Ainley et al., 2002; Schukajlow et al., 2022). Our findings align with those of García-Cerdá et al. (2024), who reported moderate correlations between interest and performance across problem types in a homogeneous sample of secondary students ($\rho = .25–.31$). For average-MPSs, task characteristics do not alter the strength of this relationship.

In contrast, for high-MPSs, correlations between interest and performance were weak and non-significant ($r = .04–.23$, all $p > .05$), diverging from prior research and highlighting the importance of considering learner profiles. This pattern suggests that high-MPSs maintain strong performance regardless of their interest level, indicating that situational interest may have a weaker influence on their achievement compared to average-MPSs. This is a revealing result that can be contrasted with results from Brandl (2011), who recommends aligning the design of modelling tasks with high-MPSs' interests to foster deeper engagement. It also supports the theoretical perspective of Krutetskii (1976) and Leikin (2013), who argue that high-MPSs rely more on advanced cognitive abilities and intrinsic motivation than on situational interest.

To identify profile-dependent differences in correlations for each problem type, we conducted a Fisher *Z*-transformation (see Table 6). The comparison of correlation coefficients between groups suggests that interest may have a greater impact on performance for average-MPSs than for high-MPSs, particularly with Word-Ps. Across all problem types, average-MPSs exhibit stronger positive correlations between interest and performance than high-MPSs, suggesting that stimulating interest is critical for their achievement. For high-MPSs, performance appears largely independent of interest, indicating that these students' mathematical potential allows them to excel despite affective factors. This has implications for enrichment design, as it suggests that increasing task complexity or novelty may be more effective than relying on contextual relevance alone.

Overall, these findings reveal that the relationship between interest and performance is not uniform across student profiles. While average-MPSs benefit from tasks that stimulate interest—regardless of type—high-MPSs maintain high performance even in tasks they find less engaging. This underscores the need for differentiated instructional strategies: for average-MPSs, fostering situational interest may boost achievement, whereas for high-MPSs, tasks should emphasise challenge and creativity to maintain engagement without compromising performance.

7 Conclusions, implications and limitations

We have conducted a detailed study of the cognitive and affective differences of two student samples confronted with three types of mathematical problems, varying in their connection to reality. Our findings complement previous research (Ferrando et al., 2024; García-Cerdá et al., 2024; Rellensmann & Schukajlow, 2017) and are particularly relevant given the emphasis on modelling in Spain's new curriculum (MEFP, 2022).

Our first two objectives focused on the cognitive and affective differences between high-MPSs and average-MPSs, referred to as the performance and interest gaps. The performance gaps are larger in Intra-Ps, confirming that abstract mathematical tasks more strongly favour students with high mathematical potential (Krutetskii, 1976). Conversely,

Mod-Ps narrowed the performance gap, suggesting that their open-ended nature and assumption-making reduce disparities. Regarding interest, gaps were smallest for Word-Ps and largest for Mod-Ps and Intra-Ps, indicating that task characteristics strongly shape motivational differences.

Our findings offer insights into how problem types shape interest and performance among high- and average-MPSs (mathematical potential students). Mod-Ps appear promising for differentiated instruction: they sparked more interest in high-MPSs than Word-Ps and more interest in average-MPSs than Intra-Ps. This supports integrating modelling tasks to balance cognitive challenge and contextual relevance across profiles. Word-Ps—common in textbooks—remain important for sustaining average-MPSs' interest, while Intra-Ps provide opportunities for advanced reasoning.

We also confirmed that for average-MPSs' interest and performance maintain a stable positive relationship across problem type (Nuutila et al., 2020; Schiefele et al., 1992; Schukajlow et al., 2022). For high-MPSs, performance was consistently high regardless of interest, particularly in Word-Ps, suggesting that mathematical potential may buffer against motivational fluctuations. No significant differences emerged when comparing correlations between groups or across problem types, suggesting that the fundamental dynamic between interest and performance is robust across different contexts.

Several limitations constrain the generalizability of our findings. First, high-MPSs were drawn from a voluntary program for promoting mathematical talent, where selection was based on problem-solving performance. This makes it unclear whether high mathematical potential and high performance can be meaningfully distinguished. Consequently, our results may reflect characteristics of program-selected students rather than all students with high mathematical potential. Second, the study was adequately powered to detect medium-to-large interaction effects ($f \geq 0.28$, partial $\eta^2 \approx .073$) but underpowered for small effects. This constraint, common when recruiting students with high mathematical potential, should inform the interpretation of non-significant findings. Third, the study involved a restricted sample: ninth-grade students from one Spanish region, a single mathematical domain, with twelve problems using identical procedures, solved individually under time constraints. Fourth, the instrument's design may have influenced interest ratings, particularly for high-MPSs who quickly recognised problem similarity and achieved high performance on intramathematical and word problems. These factors limit transferability to other grade levels, mathematical domains, educational contexts and student populations beyond enrichment settings.

Future research should incorporate qualitative methods—such as interviews and open-ended responses—to capture cognitive and affective processes underlying students' engagement with modelling problems. Prior work with high-MPSs using qualitative approaches revealed greater engagement, nuanced motivational profiles and complex behaviours, including perseverance despite errors or boredom (Ferrando et al., 2024). Expanding to broader mathematical content and mixed-method designs would deepen understanding of how problem characteristics interact with learner profiles to shape both interest and performance.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10649-026-10481-w.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Irene Ferrando and Carlos Segura. The first draft of the manuscript was written by Irene Ferrando and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. Irene Ferrando and Carlos Segura are grateful for the support provided by the Grant No.PID2021-126707NB-I00 funded by MCIU/AEI/10.13039/501100011033 and by "ERDF A way of making Europe". They also acknowledge the support from the Special Research Actions Grant UV-INV-AE-3668116 and from Generalitat Valenciana through the project CIGE2024/64.

Data availability Data from this study are available and can be shared.

Code availability Not applicable.

Declarations

Ethics approval The students involved in the study, being minors, had the consent of their parents/guardians. Both the students and their parents/guardians were informed of the purpose of the study, how the data would be collected and the publication of the results of this research.

Consent to participate All the participants completed and informed consents before participating in this study.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Ainley, M., Buckley, S., & Chan, J. (2009). Interest and efficacy beliefs in self-regulated learning. Does the task make a difference? In M. Wosnitza, A. S. Karabenick, A. Efklides, & P. Nenniger (Eds.), *Contemporary Motivation Research: From Global to Local Perspectives* (pp. 207–228). Hogrefe & Huber Publishing.

Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning and the psychological processes that mediate their relationship. *Journal of Educational Psychology, 94*(3), 545–561. https://doi.org/10.1037/0022-0663.94.3.545

Bakker, A., Cai, J., English, L., Kaiser, G., Mesa, V., & Van Dooren, W. (2019). Beyond small, medium, or large: Points of consideration when interpreting effect sizes. *Educational Studies in Mathematics, 102*, 1–8. https://doi.org/10.1007/s10649-019-09908-4

Benedicto, C., Gutiérrez, A., & Jaime, A. (2018). Analysis of mathematically gifted students' problem solving and cognitive demand. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.), *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 5, p. 18). PME. https://ivv5hpp.uni-muenster.de/u/sschu_12/pdf/Publikationen/Krawitz_Schukajlow_2018_PME42.pdf

Benyhe, I. (2011). Links between Spanish public education and gifted education: Example of the Estalmat programme. In J. B. Györi (Ed.), *International Horizons of Talent Support* (Vol. 1, pp. 185–206). Géniusz Books.

Blum, W., & Leiss, D. (2007). How do students and teachers deal with mathematical modelling problems? The example sugarloaf and the DISUM project. In C. Haines, P. L. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical modelling: Education, engineering and economics - ICTMA 12* (pp. 222–231). Horwood.

Borromeo Ferri, R. (2018). Learning how to teach mathematical modeling in school and teacher education. *Springer International Publishing*. https://doi.org/10.1007/978-3-319-68072-9

Brandl, M. (2011). Modelling tasks at the internet portal "Program for Gifted." In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. Stillman (Eds.), *Trends in teaching and learning of mathematical modelling. International Perspectives on the Teaching and Learning of Mathematical Modelling* (pp. 551–558). Springer. https://doi.org/10.1007/978-94-007-0910-2_53

Burney, V. H. (2008). Applications of social cognitive theory to gifted education. *Roeper Review, 30*(2), 130–139. https://doi.org/10.1080/02783190801955335

Cevikbas, M., Kaiser, G., & Schukajlow, S. (2022). A systematic literature review of the current discussion on mathematical modelling competencies: State-of-the-art developments in conceptualizing, measuring, and fostering. *Educational Studies in Mathematics, 109*, 205–236. https://doi.org/10.1007/s10649-021-10104-6

Chamberlin, S. A., & Moon, S. M. (2005). Model-eliciting activities as a tool to develop and identify creatively gifted mathematicians. *Journal of Secondary Gifted Education, 17*(1), 37–47. https://doi.org/10.4219/jsge-2005-393

Chang, Y.-P., Krawitz, J., Schukajlow, S., & Yang, K.-L. (2020). Comparing German and Taiwanese secondary school students' knowledge in solving mathematical modelling tasks requiring their assumptions. *ZDM – Mathematics Education, 52*(1), 59–72. https://doi.org/10.1007/s11858-019-01090-4

Cheung, A. C. K., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher, 45*(5), 283–292. https://doi.org/10.3102/0013189X16656615

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Lawrence Erlbaum Associates.

Copur-Gencturk, Y., & Doleck, T. (2021). Strategic competence for multistep fraction word problems: An overlooked aspect of mathematical knowledge for teaching. *Educational Studies in Mathematics, 107*(1), 49–70. https://doi.org/10.1007/s10649-021-10028-1

de Vries, N., Meeter, M., & Huizinga, M. (2024). Does interest fit between student and study program lead to better outcomes? A meta-analysis of vocational interest congruence as predictor for academic success. *Educational Research Review,* Article 100619. https://doi.org/10.1016/j.edurev.2024.100619

Elia, I., van den Heuvel-Panhuizen, M., & Kolovou, A. (2009). Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM – Mathematics Education, 41*, 605–618. https://doi.org/10.1007/s11858-009-0184-6

Ferrando, I., Segura, C., & Castillo, J. (2024). Regular and talented students' behaviour when solving modelling tasks: Are there differences? In H. S. Siller, V. Geiger, & G. Kaiser (Eds.), *Researching Mathematical Modelling Education in Disruptive Times. International Perspectives on the Teaching and Learning of Mathematical Modelling* (pp. 401–411). Springer. https://doi.org/10.1007/978-3-031-53322-8_32

Freiman, V. (2006). Problems to discover and to boost mathematical talent in early grades: A challenging situations approach. *The Mathematics Enthusiast, 3*(1), 51–75. https://doi.org/10.54870/1551-3440.1035

Galbraith, P., & Stillman, G. (2006). A framework for identifying student blockages during transitions in the modelling process. *ZDM – Mathematics Education, 38*, 143–162. https://doi.org/10.1007/BF02655886

García-Cerdá, C., Segura, C., & Ferrando, I. (2024). Influence of the type of mathematical problems on students' and pre-service teachers' interest and performance. A replication and elaboration study. *Implementation and Replication Studies in Mathematics Education, 4*(1), 125–159. https://doi.org/10.1163/26670127-bja10017

Geisler, S., & Rach, S. (2023). Who benefits from modelling tasks combined with experiments? Effects of students' affective traits and learning situations on situational interest and feeling of competence. *Quadrante, 32*(2), 130–152. https://doi.org/10.48489/quadrante.31452

Geretschläger, R., & Donner, L. (2022). Writing and choosing problems for a popular high school mathematics competition. *ZDM – Mathematics Education, 54*(5), 971–982. https://doi.org/10.1007/s11858-022-01351-9

Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational Studies in Mathematics, 39*, 111–129. https://doi.org/10.1023/A:1003749919816

Gutiérrez, A., Benedicto, C., Jaime, A., & Arbona, E. (2018). The cognitive demand of a gifted student's answers to geometric pattern problems. In F. M. Singer (Ed.), *Mathematical Creativity and Mathematical Giftedness* (ICME-13 Monographs, pp. 169–198). Springer. https://doi.org/10.1007/978-3-319-73156-8_7

Gutiérrez, A., & Jaime, A. (2021). Current challenges in mathematics education. *Innovaciones Educativas, 23*(34), 198–203. https://doi.org/10.22458/ie.v23i34.3515

Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters: The importance of promoting interest in education. *Policy Insights From the Behavioral and Brain Sciences, 3*(2), 220–227. https://doi.org/10.1177/2372732216655542

Hernández, E. (2006). *Nurturing the talent of gifted children ages 12 to 15: A project of Miguel de Guzmán* [Conference presentation]. International Congress of mathematicians (ICM 2006), Madrid, Spain. https://www.estalmat.org/archivos/PresentationESTALMAT-ICM2006.pdf

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Jansen, M., Lüdtke, O., & Schroeders, U. (2016). Evidence for a positive relation between interest and achievement: Examining between-person and within-person variation in five domains. *Contemporary Educational Psychology, 46*, 116–127. https://doi.org/10.1016/j.cedpsych.2016.05.004

Krawitz, J., & Schukajlow, S. (2018). Activation and monitoring of prior mathematical knowledge in modelling processes. In E. Bergqvist, M. Österholm, M. Granberg, & L. Sumpter (Eds.), *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 243–250). PME.

Krawitz, J., Schukajlow, S., & Van Dooren, W. (2018). Unrealistic responses to realistic problems with missing information: What are important barriers? *Educational Psychology, 38*(10), 1221–1238. https://doi.org/10.1080/01443410.2018.1502413

Krutetskii, V. A. (1976). *The psychology of mathematical abilities in school children*. University of Chicago Press.

Leikin, R. (2009). Exploring mathematical creativity using multiple solution tasks. In J. H. Borwein, E. M. Liljedahl, & P. L. S. de Lange (Eds.), *Mathematical creativity and mathematical giftedness* (pp. 129–145). Sense Publishers.

Leikin, R. (2013). Evaluating mathematical creativity: The interplay between multiplicity and insight. *Psychological Test and Assessment Modeling, 55*(4), 385–400.

Leikin, R. (2019). Developing mathematical talent in schoolchildren: Who, what, and how? In R. F. Subotnik, P. Olszewski-Kubilius, & F. C. Worrell (Eds.), *The psychology of high performance: Developing human potential into domain-specific talent* (pp. 173–199). American Psychological Association. https://doi.org/10.1037/0000120-009

Leikin, R. (2020). Giftedness and high ability in mathematics. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education* (pp. 315–325). Springer. https://doi.org/10.1007/978-3-030-15789-0_65

Leikin, R., & Levav-Waynberg, A. (2007). Exploring mathematics teacher knowledge to explain the gap between theory-based recommendations and school practice in the use of connecting tasks. *Educational Studies in Mathematics, 66*, 349–371. https://doi.org/10.1007/s10649-006-9071-z

Leikin, R., & Sriraman, B. (2022). Empirical research on creativity in mathematics (education): From the wastelands of psychology to the current state of the art. *ZDM – Mathematics Education, 54*, 1–17. https://doi.org/10.1007/s11858-022-01340-y

Leikin, R., Waisman, I., & Leikin, M. (2016). Does solving insight-based problems differ from solving learning-based problems? Some evidence from an ERP study. *ZDM – Mathematics Education, 48*, 305–319. https://doi.org/10.1007/s11858-016-0767-y

Lester, F. K., & Cai, J. (2016). Can mathematical problem solving be taught? Preliminary answers from 30 years of research. In P. Felmer, E. Pehkonen, & J. Kilpatrick (Eds.), *Posing and solving mathematical problems: Advances and new perspectives* (pp. 117–135). Springer. https://doi.org/10.1007/978-3-319-28023-3_8

Liebendörfer, M., & Wagner, K. (2024). Do real-world connections make tasks catch secondary school students' interest? What preservice teachers expect and what students really experience. *Implementation and Replication Studies in Mathematics Education, 4*(2), 282–314. https://doi.org/10.1163/26670127-bja10023

Liljedahl, P., & Cai, J. (2021). Empirical research on problem solving and problem posing: A look at the state of the art. *ZDM – Mathematics Education Education, 53*(4), 723–735. https://doi.org/10.1007/s11858-021-01291-w

Maaß, K., Geiger, V., Ariza, M. R., & Goos, M. (2019). The role of mathematics in interdisciplinary STEM education. *ZDM – Mathematics EducationEducation, 51*(7), 869–884. https://doi.org/10.1007/s11858-019-01100-5

Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed). Freeman.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science, 26*(1–2), 49–63. https://doi.org/10.1023/A:1003088013286

Meng, X.-l, Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111*(1), 172–175. https://doi.org/10.1037/0033-2909.111.1.172

Middleton, J. A., Jansen, A., & Goldin, G. A. (2017). The complexities of mathematics engagement: Motivation, affect, and social interactions. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 667–699). National Council of Teachers of Mathematics.

Ministerio de Educación y Formación Profesional (MEFP) (2022). *Real Decreto 217/2022, de 29 de marzo, por el que se establece la ordenación y las enseñanzas mínimas de la Educación Secundaria Obligatoria* [Royal Decree 217/2022, of 29 March, which establishes the organisation and minimum teaching of Compulsory Secondary Education]. Boletín Oficial del Estado, 76, de 30 de marzo de 2022. https://www.boe.es/eli/es/rd/2022/03/29/217/con

Niss, M., & Blum, W. (2020). The learning and teaching of mathematical modelling. *Routledge*. https://doi.org/10.4324/9781315189314

Niss, M., Blum, W., & Galbraith, P. L. (2007). Introduction. In W. Blum, P. L. Galbraith, H. W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education: The 14th ICMI study* (pp. 3–32). Springer. https://doi.org/10.1007/978-0-387-29822-1_1

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education, 15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Nuutila, K., Tapola, A., Tuominen, H., Kupiainen, S., Pásztor, A., & Niemivirta, M. (2020). Reciprocal predictions between interest, self-efficacy, and performance during a task. *Frontiers in Education, 5*(36). https://doi.org/10.3389/feduc.2020.00036

Palm, T. (2006). Word problems as simulations of real-world situations: A proposed framework. *For the Learning of Mathematics, 26*(1), 42–47.

Pantziara, M., Gagatsis, A., & Elia, I. (2009). Using diagrams as tools for the solution of non-routine mathematical problems. *Educational Studies in Mathematics, 72*, 39–60. https://doi.org/10.1007/s10649-009-9181-5

Pollak, H. O. (1969). How can we teach applications of mathematics? *Educational Studies in Mathematics, 2*(3), 393–404. https://doi.org/10.1007/BF00303471

Polya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton University Press.

Rellensmann, J., & Schukajlow, S. (2017). Does students' interest in a mathematical problem depend on the problem's connection to reality? An analysis of students' interest and pre-service teachers' judgments of students' interest in problems with and without a connection to reality. *ZDM – Mathematics Education Education, 49*(3), 367–378. https://doi.org/10.1007/s11858-016-0819-3

Renninger, K. A., & Hidi, S. (2015). *The power of interest for motivation and engagement*. Routledge.

Rotgans, J. I., & Schmidt, H. G. (2011). Situational interest and academic achievement in the active-learning classroom. *Learning and Instruction, 21*(1), 58–67. https://doi.org/10.1016/j.learninstruc.2009.11.001

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 183–212). Lawrence Erlbaum.

Schoenfeld, A. H. (1982). Measures of problem-solving performance and of problem-solving instruction. *Journal for Research in Mathematics Education, 13*(1), 31–49. https://doi.org/10.5951/jresematheduc.13.1.0031

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Academic Press.

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.

Schoenfeld, A. H. (2013). Reflections on problem solving theory and practice. *The Mathematics Enthusiast, 10*(1 & 2), 9–34.

Schoenherr, J. (2024). Personalizing real-world problems: Posing own problems increases self-efficacy expectations, intrinsic value, attainment value, and utility value. *British Journal of Educational Psychology, 94*(2), 407–424. https://doi.org/10.1111/bjep.12653

Schukajlow, S., Krawitz, J., Kanefke, J., & Rakoczy, K. (2022). Interest and performance in solving open modelling problems and closed real-world problems. In C. Fernandez, S. Llinares, A. Gutierrez, & N. Planas (Eds.), Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education (Vol. 3, pp. 403–410). PME.

Schukajlow, S., Leiss, D., Pekrun, R., Blum, W., Müller, M., & Messner, R. (2012). Teaching methods for modelling problems and students' task-specific enjoyment, value, interest and

self-efficacy expectations. *Educational Studies in Mathematics, 79*(2), 215–237. https://doi.org/10.1007/s10649-011-9341-2

Schukajlow, S., Rakoczy, K., & Pekrun, R. (2023). Emotions and motivation in mathematics education: Where we are today and where we need to go. *ZDM – Mathematics Education Education, 55*(2), 249–267. https://doi.org/10.1007/s11858-022-01463-2

Schulze Elfringhoff, M., & Schukajlow, S. (2021). What makes a modelling problem interesting? Sources of situational interest in modelling problems. *Quadrante - Revista de Investigação em Educação Matemática, 30*, 8–30. https://doi.org/10.48489/quadrante.23861

Segura, C., & Ferrando, I. (2023). Pre-service teachers' flexibility and performance in solving Fermi problems. *Educational Studies in Mathematics, 113*(2), 207–227. https://doi.org/10.1007/s10649-023-10220-5

Simpson, A. (2020). On the misinterpretation of effect size. *Educational Studies in Mathematics, 103*(1), 125–133. https://doi.org/10.1007/s10649-019-09924-4

Sriraman, B. (2003). Mathematical giftedness, problem solving, and the ability to formulate generalizations: The problem-solving experiences of four gifted students. *Journal of Secondary Gifted Education, 14*(3), 151–165. https://doi.org/10.4219/jsge-2003-425

Subotnik, R. F., Olszewski-Kubilius, P., Corwith, S., Calvert, E., & Worrell, F. C. (2023). Transforming gifted education in schools: Practical applications of a comprehensive framework for developing academic talent. *Education Sciences, 13*(7), Article 707. https://doi.org/10.3390/educsci13070707

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36*(5), 404–411.

Vicente, S., Manchado, E., & Verschaffel, L. (2018). Solving arithmetic word problems. An analysis of Spanish textbooks. *Culture and Education, 30*, 71–104. https://doi.org/10.1080/11356405.2017.1421606

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wijaya, A., van den Heuvel-Panhuizen, M., Doorman, M., & Robitzsch, A. (2014). Difficulties in solving context-based PISA mathematics tasks: An analysis of students' errors. *The Mathematics Enthusiast, 11*(3), 555–584. https://doi.org/10.54870/1551-3440.1317

Zawojewski, J. (2013). Problem solving versus modeling. In R. Lesh, P. Galbraith, C. Haines, & A. Hurford (Eds.), *Modeling students' mathematical modeling competencies* (pp. 237–243). Springer.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.