ZDM – Mathematics Education
https://doi.org/10.1007/s11858-025-01658-3

ORIGINAL PAPER

# Does training matter? Effect of training strategies on how pre-service teachers pose and assess modelling problems

Irene Ferrando[1] · Berta Barquero[2] · Carlos Segura[1]

Accepted: 12 January 2025


## Abstract

Several studies highlight teachers' challenges when posing and analysing modelling problems, often due to the absence of tools supporting these professional tasks. Thus, an important question refers to the effect of implementing different training strategies on the ability of prospective teachers to pose and assess modelling problems. Our study involves a sample of 110 pre-service primary school teachers, divided into three groups. The first group received extensive theoretical and practical teacher training in mathematical modelling based on the proposal of the study and research paths for teacher education. The second group received intensive theoretical training based on a didactic guide outlining the characteristics of modelling problems. The third group, which acts as a reference group, solely relied on the information provided by the Spanish curriculum. All participants undertake the same experience, involving the redesign and analysis of three textbook word problems into modelling problems. The results shed light on teachers' training needs and the effectiveness of the tools in enhancing their ability to pose and assess modelling tasks at the primary school level.

**Keywords** Modelling · Teacher training · Problem posing · Task design · Pre-service teachers

## 1 Introduction

Mathematical modelling stands out in the Spanish curriculum as a competence that should be addressed in mathematics teaching throughout compulsory education (MEFP, 2022). This trend reflects a broader global movement towards an instructional approach aimed at facilitating students to connect mathematical content with concrete aspects of reality (Cevikbas et al., 2022).

Teachers are supposed to know how to select, design, and adapt appropriate problems to foster the development of modelling competence among students (Chapman, 2012). Problem-posing is thus interpreted as a crucial professional task for teachers, in which they might take an active role in proposing new problems and in re-designing existing ones (Cai et al., 2015).

Recent studies have examined teachers' difficulties when designing and analysing modelling problems (Cabassut & Ferrando, 2017; Paolucci & Wessels, 2017), and many of them allow to identify gaps or strengths in the teacher's knowledge (Montes et al., 2024). Indeed, Crespo and Sinclair (2008) observe that many teachers limit themselves to selecting and assigning mathematical problems found in textbooks. Vicente and Manchado (2017), after a systematic analysis of Spanish primary school textbooks, reveal a predominance of routine exercises and word problems that present situations in stereotypical and inauthentic contexts, which is not in line with the indications of the new educational programmes (MEFP, 2022), neither facilitating a rich initial context for the design of modelling problems.

One approach to address teachers' needs regarding posing and analysing modelling problems is to investigate the effect of different training strategies for mathematical modelling (Passarella, 2021). Therefore, this article presents the results of an empirical study with pre-service primary school teachers, aiming to explore whether different training strategies (extensive theoretical and practical, intensive theoretical and baseline) influence pre-service teachers' ability to pose modelling problems and assess their characteristics.

✉ Irene Ferrando
irene.ferrando@uv.es

[1] Universitat de València, Valencia, Spain

[2] Universitat de Barcelona, Barcelona, Spain

Published online: 28 January 2025

## 2 Theoretical framework

This section introduces some central considerations regarding problem posing, the design of modelling-oriented problems and teacher education needs to address this professional task.

### 2.1 Problem posing

Mathematical problem posing has garnered significant attention in mathematics education research for at least 40 years (Liljedahl & Cai, 2021). Problem-posing encompasses a range of processes wherein teachers and students formulate problems based on specific contexts (Cai & Hwang, 2021). Different definitions of problem posing can be found in literature (Cai et al., 2015) and it can take different forms. In the present work, we focus on problem posing as the process of generating a mathematical problem by reformulating textbook ones. The transformation of problems involves posing a new problem based on a prior one by modifying several of its elements (Klein & Leikin, 2020; Stoyanova & Ellerton, 1996).

As characterised by Chapman (2012), problem posing can be viewed as a component of the mathematics teacher's competence. Investigations focusing on problem-posing among teachers have explored various aims, including how teachers manage the dynamics of problem posing (Cai & Hwang, 2021). Complementary, Montes et al. (2024) identify four main categories reflecting teachers' aims when redesigning textbook problems focusing on: the content, the context, the complexity, or the evaluation of the solutions. In our work, we are particularly interested in the first two. We intend pre-service teachers to redesign textbook problems so that the self-generated problem focuses on the content of the original and reflects an understanding of the context, making the real situation more familiar, understandable, complete or useful than in the original problem. Moreover, we consider some other variables linked to mathematical modelling, as explained in the following.

### 2.2 Modelling problems: task variables for their design and analysis

Progress has been made in characterising modelling problems. According to Niss and Blum (2020), modelling problems need a demanding translation process between the real world and mathematics. A modelling problem should therefore encourage students to understand a real-world situation, structure and simplify the information to build a mathematical model, work mathematically, interpret their results and validate them against the real-world situation.

Modelling problems are varied in their typology and scope. Maaß (2010) proposes a classification of modelling problems according to its focus, the data provided, the relationship to reality, the type of models, its openness, its cognitive demand, and the mathematical content involved. Overall, Maaß (2006) defines modelling problems as authentic, complex, and open problems.

Problems are *authentic* when they start from questions that involve considering the real-world situation (Maaß, 2010; Turner et al., 2024). Addressing these questions may require making assumptions about the situation (Niss & Blum, 2020) or integrating data about the system to be modelled. *Openness* implies that there are different ways of approaching the problem admitting different solutions (Maaß, 2010), i.e., can be both open-start and open-end (Klein & Leikin, 2020). Finally, some studies consider the *complexity* of a problem if its solution involves several steps and demands a high cognitive involvement (Maaß, 2010; Silber & Cai, 2017).

Geiger et al. (2022) defined the Design and Implementation Framework for Mathematical Modelling Tasks (DIFMT) highlighting the importance of modelling task design as complementary to subsequent implementation. From the point of view of task design, Krawitz et al. (2024) consider these same characteristics as task variables (following Goldin & McClintock, 1979), since the values taken by these variables determine the structure of the problem. These authors refer to the modelling potential of a self-generated problem by discussing the values of the previous variables that allow them to characterize to what extent a problem is aligned with modelling. Indeed, Geiger et al. (2022) point out that teachers should anticipate the modelling potential of problems for their effective design and implementation.

Czocher (2018) stresses a complex process that should be anticipated when designing a modelling problem: the possibility of determining whether the relationships between reality and mathematics are correct. Although Krawitz et al. (2024) do not consider validation as a task variable in the design of a modelling problem, Galbraith (2007) does in his five principles for designing such problems. The first principles refer to the mathematical treatment and aspects that can be determined by the three task variables described above. The last principle refers to the possibility of *interpreting* and *validating* solutions. Galbraith (2007) describes the latter principle as "an evaluation procedure is available that enables checking for mathematical accuracy, and for the appropriateness of the solution with respect to the contextual setting" (p. 55). In our study, we integrate this *validation* principle, in addition to the previously considered —authenticity, openness, complexity—, as a task variable related to the discussion of the *modelling potential* of pre-service teachers' self-generated tasks.

### 2.3 Teacher education for problem-posing modelling-related tasks

Research indicates that a substantial number of teachers perceive mathematics primarily as the execution of algorithms and formulas (Szydlik et al., 2003). Although modelling problems are supported by national curricula, the number of such problems in textbooks is scarce (Vicente & Manchado, 2017). Some studies warn that teachers tend to use material that is easy for them to find (such as textbooks), so they have difficulties in designing and analysing modelling problems (Hansen & Hana, 2015; Paolucci & Wessels, 2017; Passarella, 2021). Hence, it has become increasingly important for teacher education programs to ensure that they equip pre-service teachers with specific pedagogical knowledge to foster mathematical modelling in their classrooms effectively.

Some works (Borromeo Ferri, 2018; Brunner et al., 2024; Geiger et al., 2022; Wess et al., 2021) have developed modelling-specific pedagogical knowledge frameworks in which the design and assessment of modelling tasks is an essential component. However, some authors (Krawitz et al., 2024; Passarella, 2021) have pointed out that more research is needed on the impact of specific teacher training strategies for modelling on how teachers progress on posing modelling problems.

In this line, research on problem posing has found that when pre-service teachers receive long-term training in designing and assessing their self-generated problems, they improve their task design concerning several task variables, such as correctness of the mathematical content or authenticity (Abu-Elwan, 2002). However, Chen et al. (2015) found that difficulties in posing high-quality problems persisted among students. After a short-term intervention (including prompts) on modelling problem design with students, Krawitz et al. (2024) also found improvements in some task variables, although not in complexity.

Paolucci and Wessels (2017) highlight the need to promote metacognitive recognition practices among pre-service teachers so that they are then able to design and assess modelling problems efficiently. In this line, Wess et al. (2021) highlight that an important aspect of specific knowledge about modelling tasks is the understanding of their characteristics. Brunner et al. (2024) point out that the ability of pre-service teachers to accurately assess the task variables of a problem is an essential component of pedagogical knowledge for teaching. However, Kaiser et al. (2010) have shown that pre-service teachers often have a rigid conception of the characteristics of mathematical modelling, which may impact the accuracy of their diagnostic judgements about problems. It is then also necessary to investigate whether specific teacher training strategies for modelling have an impact on the ability of preservice teachers when identifying the characteristics of modelling tasks, that is, in their assessment accuracy.

## 3 Research questions and hypothesis

Considering the presence of modelling in educational programmes and the difficulties of teachers when designing modelling tasks (Cabassut & Ferrando, 2017), it seems crucial to investigate which training strategies are more effective in helping pre-service teachers to pose modelling problems (unaddressed question 3 from Cai et al., 2015). Complementarily, teachers must also be able to accurately assess their self-generated problems to ensure that they meet the characteristics of a modelling activity involving the mathematical content they wish to introduce (Montes et al., 2024). It is then important to study the effect of the teacher training strategy on their accuracy when assessing their self-generated modelling problems.

Our study aims to compare the effects of different teacher training strategies on three samples of pre-service primary school teachers. In two of the samples, the training strategies have different characteristics: one is extensive with a theoretical-practical approach to modelling with fluent interaction among educators and pre-service teachers, and the other is intensive and theoretically self-taught. In the third sample, the reference group, the training strategy solely consisted of providing the participants with an extract about modelling from the Spanish mathematics curriculum. Our paper thus addresses the following research questions:

**RQ1** How do teacher training strategies influence the characteristics of the pre-service teachers' self-generated modelling problems?

**RQ2** How do teacher training strategies influence pre-service teachers' accuracy in assessing the proposed self-generated modelling problems?

Regarding the characteristics of self-generated problems, we hypothesise that the participants, who follow extensive theoretical-practical training will be able to pose problems which address most of the characteristics of modelling tasks. Participants who are introduced intensively to a theoretical approach to modelling will be more capable of focusing on the real context than those in the reference group, even if we do not expect them to be able to pose, in general, complete modelling tasks. However, we hypothesise that both group participants would have difficulties in posing authentic tasks that can be validatable since these characteristics are residual in the type of problems proposed in Spanish textbooks.

Regarding the accuracy when assessing their self-generated problems, we hypothesise that the accuracy will be higher for those participants who follow extensive practical training than for the other two groups because the training includes specific didactic tools for task analysis (such as the analysis of task variables, tools for the epistemological and didactic analysis of tasks, among others, see Barquero, 2023 for more details). However, we do not expect high results in this metacognitive aspect in any of the groups, because of the constraints derived from the conception of modelling that has already been pointed out in the literature.

## 4 Methodological design

### 4.1 Sample and design

The study comprises 110 pre-service primary school teachers enrolled at two Spanish public universities, the Universitat de Barcelona and Universitat de València. These participants are divided into three groups: TP-group (N = 26), T-Group (N = 32) and Rf-Group (N = 52). All participants, within each group, are in the same class in the Faculty of Education, where they are attending a mandatory course on Didactics of Mathematics. Even if the groups are from two different universities, all the participants are following the same undergraduate studies (regulated by a national law) at the same level. All participants are in their last year (of a 4-year degree) following the last compulsory course on Didactics of Mathematics for pre-service primary school teachers (a 4th-year course, 6 ECTS).

We employed a quasi-experimental design with a convenience sample (see Fig. 1). First, to analyse participants' prior experience in solving and posing modelling problems, an online questionnaire was administered to the whole sample, at the beginning of our study. The intervention consisted of teacher training 1 week after the participants completed the online questionnaire. The teacher training lessons in TP-group were led by the second author during eight 120-min class sessions in the regular classroom. The third author led the teacher training in T-group and Rf-group. For T-group the teacher training was developed during a 120-min session in the regular classroom, while for Rf-group it last 30 min. Just after the training was completed, each group participated in the same problem-posing experience.

The problem-posing experience was developed for the three groups, during a 120-min session under the supervision of one of the researchers. Participants worked individually. The experience consisted of two parts, in the first part, the educators presented three problems from a Spanish textbook for sixth-grader primary school students (Tomás Henao, 2023). To choose the problems, we looked at textbooks from two of the most common publishers in Spain. We aimed to put participants in the shoes of primary school teachers who are faced with redesigning activities based on those they find in the most common resources in primary classrooms. First,

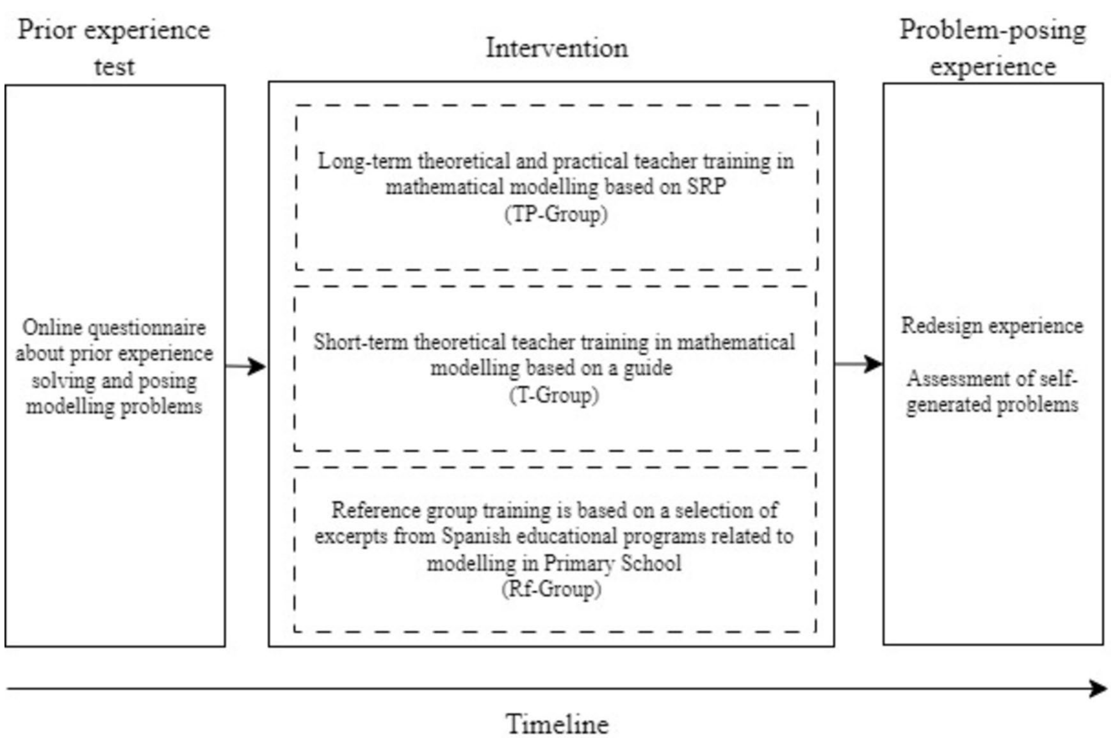


**Fig. 1** Overview of study design

**Table 1** General description of the teacher training contents addressed in each group (theoretically or practically)

| | MM in the Spanish curriculum | | Characteristics of MM tasks | | MM solving process | | MM Task design | |
|---|---|---|---|---|---|---|---|---|
| | Theory | Practice | Theory | Practice | Theory | Practice | Theory | Practice |
| TP-group | X | | X | X | X | X | X | X |
| T-group | X | | X | | X | | X | |
| Rf-group | X | | X | | | | | |

they were asked to solve the textbook problems. Then, the instruction was "From each of the three tasks, you should pose a modelling problem for 6th-grade students. The problem you pose should lead, at least, to cover the same mathematical content as the original problem." In the second part, participants were asked to solve the self-generated problems and to assess their characteristics using a questionnaire consisting of six items with dichotomous responses (yes or no).

## 4.2 Intervention

The participants in TP-group took part in an extensive course, which ran over a semester, supervised by the second author at the Universitat de Barcelona (see Barquero (2023) for a whole description of the course). The last block of the course ran over 4 weeks with two 120-min sessions per week. This block focused on transferring tools for the analysis, management and design of modelling tasks, based on the study and research paths (SRP). During this block, students initially got experience with the SRP, as proposed by the anthropological theory of the didactic for the teaching and learning of modelling (Bosch, 2018). Participants were asked to experience an SRP (for 3 sessions) that, on this occasion, was about the *cake box* modelling task (Barquero et al., 2024). Following this experience, students were introduced to specific tools for the analysis (2 sessions), such as the analysis of the dialectics questions-answers in the modelling process or the analysis of the evolution of the system and considered models. This was followed by 3 sessions where pre-service teachers worked on the design of (or adaptation of an experienced) an SRP.

The participants in T-group and Rf-group took part in a short-term teacher training supervised by the third author, who was the educator of both courses. For T-group, the intensive training consisted of introducing a didactic guide about modelling problems, which was distributed to each participant, who was asked to work on it autonomously. The guide, specially designed for this research, comprises five short sections. The first section is a summary of the Spanish mathematics curriculum. The second section focuses on realistic contexts, showing examples of authentic and non-authentic situations. The third section introduces modelling problems and explains the characterisation of such modelling problems (following Maaß, 2006) by providing examples of problems that fulfil all the characteristics and those that do not. The fourth section explains each phase of the modelling cycle through the solution of an example. Finally, the last section relates the task design guidelines outlined in the Spanish curriculum with the characteristics for designing modelling problems. The participants in the Rf-group were only provided with a printed document with extracts from the Spanish mathematics curriculum related to modelling. During the training sessions in the T-group and Rf-group, the educator oversaw a summary of the written content, which was then read and discussed by the participants in pairs.

It is worth noting that Rf-group serves as a reference group, enabling us to evaluate the influence of the training strategies employed within the other two groups. Additionally, the comparison among the groups allows us to find out how prepared primary school teachers are to understand the curriculum guidelines to design modelling problems. Table 1 summarizes the contents related to mathematical modelling (MM) addressed in each group.

## 4.3 Materials

### 4.3.1 Prior experience with modelling online questionnaire

The questionnaire[1] items were adapted from the work of Rellensman and Schukajlow (2017) and were validated by Ferrando et al. (2024). It comprises two sections, each containing six questions rated on a five-point Likert scale (where 1 is totally disagree and 5 is totally agree). The first section explores participants' previous experience in solving modelling problems, and the second their experience in designing modelling problems.

### 4.3.2 Redesign experience material

The material consisted of a booklet with three textbook problems that share some characteristics: they are formulated in a sports context, they include text and an illustrative picture, and they do not require realistic assumptions

[1] This questionnaire and the other materials used for data collection are available as electronic supplementary materials.

for their solution. Furthermore, we selected each problem to encompass varied mathematical knowledge. *Problem 1* involves measurement content (see Fig. 2), *Problem 2* involves arithmetical operations and numerical comparisons and *Problem 3* involves statistical content.

During the first phase, the participants solve the textbook tasks and work on the redesign of the tasks to get their self-generated problems.

#### 4.3.3 Self-generated problems assessment material

In the second phase, we asked the participants to solve their self-generated problems and to assess them by fulfilling six items that correspond to the different task variables.

The first two task variables, assessed in items 1 and 2, relate to the relationship in terms of content and context between the self-generated problem and the original textbook one. Item 1 examines whether the self-generated problem includes the *contents* of the original, and item 2 examines whether the problem requires consideration of the *context* in which it is formulated (following Montes et al., 2024). The other task variables are related to the modelling potential of the self-generated problem (Krawitz et al., 2024). Item 3 examines the *authenticity* of the situation presented by the problem, focusing on whether the real-world context influenced the solving process. Item 4 assesses the *openness* of the problem, specifically, the possibility of different solutions. Item 5 evaluates the *complexity* of the problem, indicating whether it requires addressing different steps for its solution. Finally, item 6 enquires about the availability of an *evaluation procedure* that, adapting the proposal of Galbraith (2007), refers to the possibility of *validating* the answers against the task reality.

### 4.4 Measures

#### 4.4.1 Prior experience with modelling

For each participant, we obtained the mean value of the answers to the six questions related to their prior experience when solving and designing modelling problems, resulting in two values up to five points. The results are presented in Table 2.

We conducted a one-way ANOVA analysis to identify differences between groups. No differences were found regarding solving modelling problems ($F = 0.57$, $p = 0.57$), but we found significant differences between the groups ($p < 0.001$) in modelling-problems design. Post-hoc test shows us that the previous experience of TP-group is significantly lower than that of the other two groups.

#### 4.4.2 Characterization of self-generated problems

To characterise the problems designed by the participants, we analysed the self-generated problems. With this aim, we relied on the same six task variables related to the six items answered by the participants (focus on the content, understanding of the context, authenticity, openness, complexity, and validation). For each task variable, we rated with 0, if it does not appear, and 1, if it does. Three raters scored the

**Table 2** Results of the questionnaire on prior modelling experience (all the values are in a range between 1 and 5 points)

| | Solving modelling, mean (SD) | Designing modelling, mean (SD) |
|---|---|---|
| TP-group | 3.34 (0.46) | 2.72 (0.93) |
| T-group | 3.49 (0.51) | 3.47 (0.69) |
| Rf-group | 3.37 (0.44) | 3.13 (0.91) |

**Fig. 2** First problem (P1) provided to participants (authors' translation from Tomás Henao, 2023, p. 205)

**YOU ARE ABLE TO... make calculations for the maintenance of a swimming pool**

A swimming school is preparing the pool for this season. They have filled it with water and have to add chlorine to make it ready for the classes to start.

The school's pool is shaped like an orthohedron and is 50 m long, 20 m wide and 2 m deep.

The school knows that they have to add 4g of chlorine per cubic metre of water. The chlorine is bought in cans of 5kg each.

- How many cubic metres is the pool? How many kilolitres is that?
- How many grams of chlorine should they put in the pool in total?
- How many bots of chlorine do they have to buy to prepare it? Will they have any chlorine left over?

self-generated problems: 5% of the sample was analysed to confirm the interpretation of each variable. Then, to assess interrater agreement, 20% of the self-generated problems were independently coded by two pairs of raters. The inter-rater agreement (Cohen's $\kappa$) for all the task variables ranged between 0.73 and 1 ($M=0.91$). The double coding was analysed to identify systematic disagreements, which were subsequently discussed and resolved.

From this analysis, we obtained, for each participant, the mean values for each task variable (between 0 and 1, the higher the value, the more this characteristic appears in the problems posed). Concerning this data, we did not have missing values (all the participants designed the three problems).

The use of these dichotomous variables for problem analysis was the key to approaching a categorisation of problems. We used presence/absence combinations of the task variables to delimit different categories. To define self-generated problems' categories, we stand on those task variables related to modelling potential: *openness* (we use C when the problem is closed and O when it is open), *complexity* (we use S when the problem is simple and K when it is complex), *authenticity* (we use A when it is an authentic problem) and *validation* (we use V when the problem is validatable). Table 3 outlines the coding and description of each category; some illustrative examples are included in the electronic supplementary material files.

Regarding this categorisation, it is worth noting that we have explicitly considered two characteristics of modelling tasks: openness and complexity The other two characteristics, authenticity and validation may or may not appear in the CS (closed and simple), CK (closed and complex) and OS (open and simple) categories but not in the OK category, as complex and authentic open problems fall into the category of modelling problems, coded as OKA(V) (Maaß, 2006).

#### 4.4.3 Scales for measuring accuracy when assessing self-generated problems

To quantify the participants' degree of accuracy in the assessment of their self-generated problems, for each task variable, we calculated the means of the absolute values of the differences between their ratings and those of the researchers (obtaining values between 0 and 1, the higher the value, the greater the gap between the student's assessment and the researchers' assessment). In case of missing values (less than 6%), we used the mean values of the corresponding group.

#### 4.4.4 Data analysis for addressing research questions

To find the effect of the training strategy on the pre-service teachers' ability to pose problems, as well as on their ability to assess their self-generated problems, we study whether there are significant differences in the scores obtained in TP-group, T-group and Rf-group for each task variable. Consequently, to address these two research questions we use one-way ANCOVA tests, considering as a covariate the pre-service teachers' prior experience posing modelling problems, as this was not the same in all three groups (see Sect. 4.4.1). This analysis requires normality and homoscedasticity, although it remains robust in cases where this is not met. In addition, the data meet the assumptions of no multicollinearity (no correlation) between the dependent variable and the covariate, and homogeneity of the regression coefficients.

**Table 3** Categorisation of the designed problems

| Category | Description |
|---|---|
| NP<br>Not a problem | Tasks that cannot be solved with mathematical procedures |
| CS<br>Closed and simple problem | Problems that can be solved with a direct mathematical procedure and with only one solution |
| CK<br>Closed and complex problem | Problems that require different mathematical procedures to reach a solution, but the solution is unique |
| OS<br>Open and simple | Problems that can be solved with a direct mathematical procedure, but they have several solutions |
| OK<br>Open and complex problem | Problems that need different mathematical procedures and they have several solutions |
| OKA(V)<br>Modelling problem | Problems that are open, complex, authentic (and validatable) |

## 5 Results

### 5.1 Characteristics of the problems designed

In Table 4 we show the mean values (and the standard deviation) for each task variable in each group, as well as the ANCOVA F-values, their significance, the effect size of the differences ($\eta^2$) and the adjusted $R^2$, which explains what percentage of the variance in the score of each task variable is explained by the model. We find significant differences in scores between groups in all task variables of the self-generated problems, controlling for prior experience with modelling problems. Moreover, the effect size of the differences is high ($\eta^2 > 0.2$) in all the variables except for the focus on the content ($\eta^2 = 0.1$), which is medium; in the rest of the task variables the training strategy received, controlled by prior experience posing modelling problems, explains between 20 and 30% of the performance.

We conducted a Sidak post-hoc analysis of pairwise comparisons to identify differences between groups The results indicate that in the variable focus on the content TP-group performs significantly better than T-group ($p < 0.01$) and Rf-group ($p < 0.05$). On the remaining five tasks variables, no significant differences are identified between TP-group and T-group, but significant differences are identified with Rf-group. Indeed, pre-service teachers in TP-group and T-group design problems that meet all five task variables significantly better ($p < 0.01$) than those in Rf-group.

To identify the differences regarding the categories described in Table 3, in Table 5, we present the proportion of each category in each group.

In Fig. 3 we present a graph of the proportion of each characteristic in each group.

To identify differences between the problems designed in each group, we conducted a two-tailed direct test of comparison of proportions. The results confirm that the ratio of non-mathematical problems designed by participants is significantly higher in T-group than in TP-group ($z = 2.51$ and $p < 0.05$) and Rf-group ($z = 3.28$ and $p < 0.01$). The ratio of closed and simple problems is significantly higher in Rf-group than in TP-group ($z = 5.34$ and $p < 0.001$) and T-group ($z = 5.41$ and $p < 0.001$). The ratio of closed and complex problems is significantly lower in T-group than in TP-group ($z = -2.35$ and $p < 0.05$) and Rf-group ($z = -2.16$ and $p < 0.05$). There are no differences between the groups in the proportion of open and simple problems and open and complex problems designed by the participants.

Concerning OKA(V) problems, namely those that meet the three task variables that allow them to be considered modelling problems, we observe that their ratio in TP-group and T-group are almost identical (51% and 52%, respectively). On the other hand, in the Rf-group, only 15% of the designed problems can be considered modelling problems. Indeed, this difference is significant ($z = 4.69$ and $p < 0.001$ for TP-group–Rf-group, $z = 5.07$ and $p < 0.001$ for T-group–Rf-group). If, within this category, we examine those more complete modelling problems, which are also validatable (OKAV), then we observe that their ratio is significantly lower in Rf-group than in TP-group ($z = -4.20$ and $p < 0.001$) and T-group ($z = -6.26$ and $p < 0.001$). However, the difference in the ratio of OKAV problems designed by TP-group and T-group pre-service teachers is not significant ($p = 0.10$).

**Table 4** Results of the ANCOVA analysis of the self-generated problems

| | $F$ | $\eta^2$ | $R^2$ (%) | TP-group | T-group | Rf-group |
|---|---|---|---|---|---|---|
| Focus on the content | $4.13^{**}$ | 0.10 | 8 | 0.95 (0.12) | 0.73 (0.32) | 0.79 (0.24) |
| Understanding of the context | $13.98^{***}$ | 0.28 | 26 | 0.65 (0.32) | 0.52 (0.41) | 0.21 (0.24) |
| Authenticity | $15.13^{***}$ | 0.29 | 28 | 0.56 (0.35) | 0.60 (0.36) | 0.19 (0.25) |
| Openness | $13.55^{***}$ | 0.27 | 26 | 0.60 (0.30) | 0.57 (0.31) | 0.25 (0.24) |
| Complexity | $9.74^{***}$ | 0.22 | 19 | 0.73 (0.28) | 0.58 (0.34) | 0.36 (0.28) |
| Validation | $10.91^{***}$ | 0.24 | 21 | 0.32 (0.26) | 0.41 (0.29) | 0.11 (0.18) |

$^*p < .05$, $^{**}p < .01$ and $^{***}p < .001$

**Table 5** Ratio of problems in each group

| | NP | CS | CK | OS | OK | OKA(V) |
|---|---|---|---|---|---|---|
| TP-group | 2 out of 78 | 18 out of 78 | 11 out of 78 | 1 out of 78 | 6 out of 78 | 40 out of 78 (21 V) |
| T-group | 13 out of 96 | 24 out of 96 | 4 out of 96 | 3 out of 96 | 2 out of 96 | 50 out of 96 (37 V) |
| Rf-group | 4 out of 156 | 94 out of 156 | 19 out of 156 | 2 out of 156 | 14 out of 156 | 23 out of 156 (11 V) |

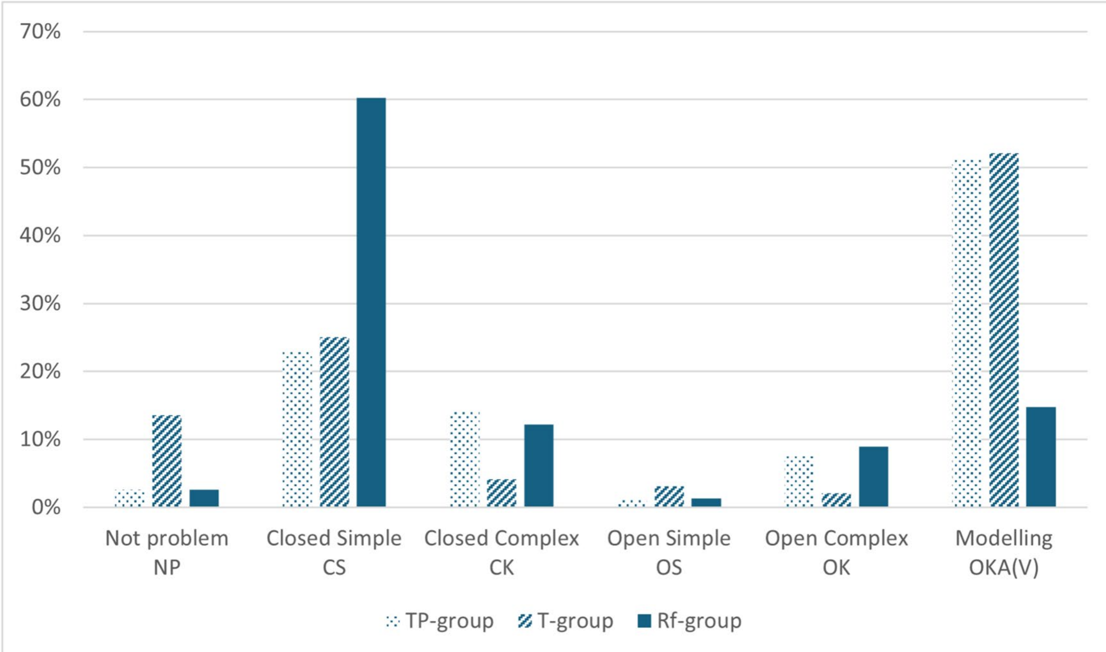


**Fig. 3** Percentage of problem categories in each group

## 5.2 Pre-service teachers assessing their self-generated problems

To analyse pre-service teachers' accuracy in assessing their self-generated problems, we calculated the mean of the absolute values of the differences between their responses to the six items and the results of the researchers' analysis of the six task variables (Table 6). Therefore, the closer the value is to zero, the more accurate the assessment of one's problems.

The results indicate that participants in the TP-group assess their self-generated problems more accurately across all task variables. However, it is necessary to analyse whether these differences are statistically significant, controlling for pre-service teachers' prior experience posing modelling problems. Using the one-way ANCOVA test we obtain that there are some significant differences between the groups in accuracy when assessing their self-generated problems (see Table 6). Specifically, these differences are significant for the task variables related to content, context, authenticity and complexity, with effect sizes being high ($\eta^2 > 0.14$) for authenticity and complexity, and medium ($\eta^2 = 0.12$) for focus on the content and low for understanding the context ($\eta^2 = 0.07$). The training strategy, controlled for prior experience posing modelling problems, explains between 2 and 17% of pre-service teachers' accuracy in assessing the characteristics of their self-generated problems.

Furthermore, a Sidak post-hoc analysis of pairwise comparisons reveals that TP-group participants assess significantly more accurately whether their designed problems are content-focused and complex than the other two groups ($p < 0.01$). Regarding the assessment of the authenticity of their self-generated problems, TP-group and T-group are significantly more accurate than Rf-group ($p < 0.01$). Finally, TP-group assesses the context of their problems significantly more accurately than Rf-group ($p < 0.05$).

**Table 6** Results of the differences for Groups 1, 2 and 3 respectively

| | $F$ | $\eta^2$ | $R^2$ (%) | TP-group | T-group | Rf-group |
|---|---|---|---|---|---|---|
| Focus on the content | 4.91** | 0.12 | 10 | 0.05 (0.12) | 0.28 (0.32) | 0.24 (0.26) |
| Understanding of the context | 2.70* | 0.07 | 5 | 0.37 (0.32) | 0.40 (0.35) | 0.49 (0.34) |
| Authenticity | 8.18*** | 0.19 | 17 | 0.36 (0.34) | 0.42 (0.36) | 0.71 (0.33) |
| Openness | 1.72 | 0.05 | 2 | 0.27 (0.23) | 0.37 (0.29) | 0.31 (0.27) |
| Complexity | 5.96*** | 0.14 | 12 | 0.28 (0.30) | 0.43 (0.33) | 0.55 (0.25) |
| Validation | 1.76 | 0.05 | 2 | 0.50 (0.25) | 0.61 (0.30) | 0.64 (0.35) |

*$p < .05$, **$p < .01$ and ***$p < .001$

## 6 Discussion

### 6.1 Influence of training strategies on the characteristics of the pre-service teachers' self-generated problems

The results shown in Table 4 reveal that the intervention in the TP-group (which includes practical training on task design) and the T-group (with the didactic guide on modelling problems) helped the participants to pose problems that met the characteristics of modelling tasks: authenticity, openness, complexity and, to a lesser extent, validation. Indeed, the training strategy received by pre-service teachers has effects of up to 30% on the success of pre-service teachers in posing problems that meet each characteristic of modelling problems. Therefore, teacher training helps to address the needs in the modelling problem-posing (Passarella, 2021), even when there is less prior experience, as was the case with the TP-group (see Table 2) or when the training is short-term and theoretical.

Nevertheless, we observe some interesting differences between the TP-group and the T-group. When we focus on the variables related to the process of re-designing a task preserving the mathematical content but improving the connection of these procedures with the real context (Montes et al., 2024), the results show significant differences between the TP-group and the other two groups. We infer that the participation of pre-service teachers in extensive training that focuses on deepening task analysis contributes to improving their ability to accurately identify the mathematical content of the problem and to be able to reflect on and improve the relationship between the content and the context of the problem when they work on the design of a modelling task. The difficulties of pre-service teachers in identifying the mathematical content when they pose modelling problems have already been pointed out in previous studies (Hansen & Hana, 2015; Paolucci & Wessels, 2017).

Analysing the task variables that determine the modelling potential of the self-generated problems is the key to drawing up the categorisation of problems, as presented in Table 3, where indeed, we combined these variables to obtain different problem categories. The results of the analysis based on this categorization presented in Table 5 and Fig. 3 show that in the TP-group and the T-group, half of the designed problems are aligned with mathematical modelling (OKA(V) category). However, in Rf-group the rate of modelling problems is barely 15% (23 out of 156). Even though the proportion of these types of problems in the TP-group and the T-group is significantly higher, we observe that pre-service teachers have difficulties when designing problems that meet the conditions to be considered modelling tasks even after specific training. These results aligned with those of the study developed by Chen et al. (2015) with primary school students, but, in our case, they are even more surprising because the participants are pre-service primary school teachers in their last year of pre-service training. Considering that modelling is a competence that is explicitly addressed in Spanish educational programs, it is revealing to observe that pre-service teachers have important difficulties in designing tasks that meet the conditions to develop modelling competence.

Moreover, within the design of modelling tasks, we consider it pertinent to differentiate between those that consider the validation aspect (OKAV problems) and those that do not. The inclusion of a new variable regarding validation (V) introduced a novel point of reflection, previously explored by Krawitz et al. (2024). We consider that taking this variable into account facilitated reflection on an essential dimension of modelling work (Czocher, 2018; Galbraith, 2007). The analysis conducted on the self-generated tasks reveals that, frequently, the proposed tasks lack or fail to anticipate validation means (OKA problems), which is in line with the results of Czocher (2018), who points out that we should afford a more prominent role to validation.

Neglecting this validation aspect results in a loss of tasks' didactic potential, often leaving them too open (see example below) and failing to provide a clear opportunity for contrasting models against any system or refining the models in this validation stages (Czocher, 2018; Galbraith, 2007; Jonassen, 1997).

> *A football team sells tickets for every match at €20 and full season tickets at different prices (grandstand: €300, side: €250 and away: €225). The team plays, on average, 50 matches per season. How much will the team earn in a season?*

It is relevant to note that the ratio of validatable modelling problems (OKAV), that is, those problems self-generated by pre-service teachers that meet all the desirable task variables, is slightly higher among participants in T-group (37 out of 96, 38%), although the difference with TP-group (21 out of 78, 27%) is not significant. Although small, the difference could be explained because the didactic guide provided to T-group participants in their training strategy included a section dedicated to Blum and Leiss' (2007) modelling cycle applied to an example problem, which included an explicit definition and exemplification of the validation process. In contrast, the TP-group placed a significant emphasis on the design of study and research paths, which commonly involve lengthy processes of modelling where both systems and models evolve, encouraging pre-service teachers to reflect on the importance of anticipating a sufficiently rich didactic means for helping modelling work to progress (Barquero, 2023). Perhaps the focus on the model evolution process led them to pose more open problems (ill-structured, as pointed out by Jonassen, 1997) without explicitly considering the possibility of validating the results.

It is important to note that Rf-group (acting as the reference group) most of the problems designed (56%, 88 out of 156) are routine problems: closed and simple and, in general, not authentic (only 6 in that category). That is, most of the problems designed by pre-service teachers after reading the indications on modelling of the Spanish educational program remain similar to those provided to them in the experience to be transformed and to those that abound in most Spanish textbooks (Vicente & Manchado, 2017). These results indicate that, especially for pre-service teachers, the curricular indications of the official documents on modelling are not clear enough and are insufficient, since they continue to tend to pose simple and closed problems, as other studies have found (Hartmann et al., 2021, 2023; Paolucci & Wessels, 2017).

Our study suggests some recommendations for pre-service teacher education. In Spain, unlike other countries, there is no specialist mathematics teacher in primary school. However, all pre-service teachers receive compulsory basic mathematical and didactic training in mathematics during their undergraduate studies (between 18 and 21 ECTS, with slight variations depending on the university). This limited training in mathematics education may explain the difficulties encountered by the participants when designing modelling problems from textbook problems.

### 6.2 Influence of training strategies on the accuracy of pre-service teachers to assess their self-generated problems

The results presented in Table 6 reveal important differences in the behaviour of TP-group when assessing their self-generated problems. On the one hand, their analysis is significantly more accurate when assessing the content and the context. This can be an effect of the training received by this group, as an important part of the intervention focused on practising analysis of the mathematical content of the tasks and their relation to the context (Barquero, 2023). The ability to adequately assess the mathematical content of their self-generated problems detected in TP-group reinforces the findings in answering the first research question, as we observed they posed problems better aligned with the content of the textbook problems. In contrast, the didactic guide received by T-group focused on modelling characteristics (Maaß, 2006; Niss & Blum, 2020) without connecting it to content analysis.

On the other hand, TP-group pre-service teachers are also the best at assessing the complexity of their self-generated problems. In contrast, the results presented in Table 6 show that both T-group and especially Rf-group overestimate the complexity of their problems. In the present work, following Krawitz et al. (2024), the complexity assessment is based on the number of solution steps, so the item related to that characteristic was: “*The problem I have posed is solved in different stages, breaking it down into subproblems*”. In this sense, we have observed that in most of the cases (especially in Rf-group, with no specific training) participants assessed problems as complex because they included different questions, without being aware that all these questions could be solved with a straightforward mathematical procedure. They therefore confuse several (simple) questions with several steps. For example, participant 301 from Rf-group proposed the following problem (based on problem 1): “*In the oceanographic there is an aquarium in the shape of an orthohedron, measuring 20 m long, 15 m wide and 3 m deep. For each cubic meter, 2 g of chlorine is needed, which is sold in 1.5 kg bags. How many cubic meters of water does the aquarium have? How many litres are there? How many grams of chlorine are needed? How many bags of chlorine do you have to buy? Is there any leftover?*”. The participant has considered this problem to be complex, authentic, open and validatable, while the problem is almost the same as problem 1 (Fig. 2). This is interesting because we observe how difficult it is for the participant to analyse the problems: we find significant confusion in terms of complexity, but also authenticity as, for instance, participant 301 who just changed the context without realising that it is not a good idea to add chlorine to the aquarium water.

Regarding authenticity, it is worth noting that both extensive (TP-group) and intensive (T-group) training interventions help to improve the pre-service teachers’ accuracy to judge it in their self-generated problems. In both training strategies, participants have gained experience of the characteristics of real and authentic contexts, and this enables them to assess adequately whether the situation posed by the problem has a real impact on the problem-solving process (Maaß, 2010; Turner et al., 2024). In contrast, as we have pointed out, pre-service teachers (Rf-group) who have consulted the Spanish curriculum documents have serious difficulties in understanding what a real and authentic context is, although it is repeatedly pointed out in the curriculum guidelines (MEFP, 2022). This can be explained by the fact that authentic contexts in problems are uncommon in conventional textbooks (Vicente & Manchado, 2017). Furthermore, it calls for a deep revision of the mathematical and didactic training offered in university to pre-service teachers (Fernández-Blanco et al., 2022), which should facilitate (at least) the experience of pre-service teachers with authentic context mathematics problems.

We also should note that training strategies have not been effective in improving pre-service teachers’ ability to assess whether their self-generated problems are open and validatable. In these aspects, their conception remains rigid (Kaiser et al., 2010). Thus, in the case of openness, pre-service teachers tend to confuse a problem that can be solved with several procedures -which refers to flexibility, as Klein and Leikin (2020) point out-, even if it has only one possible solution, with a problem that admits different solutions. Training design should focus more on this difference. Regarding whether a problem can be validated, we note that, as Jonassen (1997) points out, pre-service teachers seem to confuse the verification of a solution (checking that the calculations that have led to it are correct) with the validation, which implies that the problem provides a context for contrasting whether the solution is appropriate (Galbraith, 2007). Furthermore, these difficulties in identifying whether the validation process is possible align with the results of Czocher (2018), who found that it is the most complicated modelling process to anticipate.

## 7 Implications for teacher education

The effective introduction of modelling in the classroom requires teachers to be prepared to design modelling tasks and to accurately assess the characteristics of the activities they propose.

Regarding the design of modelling tasks, our results show that, in line with what was pointed out by Geiger et al. (2022), it is important to complement curricular changes with more accurate didactic guides (as in T-group) or with specific practical training (as in TP-group), so that pre-service and in-service teachers can gain experience and specific knowledge in modelling-oriented task design (Abu-Elwan, 2002; Passarella, 2021). Indeed, in the two groups that have received a didactic guide or have participated in specific modelling training, the ratio of routine problems is significantly lower (around 25%). Although this variation, this proportion is still high, so even with specific training, the difficulties of pre-service teachers in designing modelling problems persist, as stated by other studies (Chen et al., 2015; Hartmann et al., 2021, 2023; Krawitz et al., 2024).

Regarding the ability to accurately assess the characteristics of their self-generated problems, our results show that the training strategies proposed in our study have small effects (less than 20%). This indicates that, when it comes to modelling, it is more complicated to promote metacognitive recognition practices than problem-posing practices (Paolucci & Wessels, 2017). Therefore, the results of our study allow us to infer that there are some important aspects to consider when designing modelling training for future teachers.

First, teacher education should include the analysis of modelling tasks, with the transference and the use of specific didactic tools for the analysis, especially concerning the mathematical analysis of the tasks. This might allow pre-service teachers, when redesigning tasks, to be able to pose modelling problems related to certain real contexts and anticipate the mathematical knowledge involved in these problems (as happens in the TP-group), but also to be able to make a more accurate judgement of their tasks.

Second, pre-service teachers need to collectively share a characterisation of modelling tasks (from illustrative examples and counter-examples). In this way, teachers would have a clearer position and focus for the design and assessment of modelling tasks. However, considering the results of the TP-group and the T-group, the design of modelling tasks that accomplish some of these characteristics (such as openness, complexity, authenticity and validation) might be carefully and thoroughly addressed throughout the teacher education processes, as it involves important changes on the more dominant way to teach mathematics. In this sense, specific training more oriented towards pedagogical and didactic knowledge for teaching (Brunner et al., 2024; Geiger et al., 2022) might be needed to dismantle rigid conceptions of modelling characteristics (Kaiser et al., 2010).

Finally, teacher education should provide the opportunity not only to design problems but also to experience them and to analyse them to identify their modelling potential. In this sense, each of the characteristics of a modelling task and their relationship with the modelling phases (as is the case in the T-group about validation) must be emphasised.

## 8 Limitations

We are conscious of several limitations of our research as, on the one hand, the sample is relatively small, as data were collected from only two faculties of education. Although the socio-cultural profile of the students in both faculties is similar (both are public universities in two of the largest cities in Spain, with no significant cultural differences between the two regions), it should be noted that the design of the undergraduate studies may differ slightly between the two universities.

Moreover, it is a convenience sample since the participants were not randomly assigned to different groups. And, on the other hand, our analysis of the designed problems focused only on written production, a more in-depth analysis supplemented by interviews could undoubtedly enrich the results. Furthermore, our analysis was limited to the quantitative assessment of responses to dichotomous items. However, we also collected participants' reflections on each task. In future research, we aim to delve deeper into these reflections, as they are likely to shed light on key aspects related to pre-service modelling-specific didactic and pedagogical knowledge.

## 9 Conclusions

Our empirical study examines the effect of three teacher training approaches (extensive theoretical and practical, intensive theoretical and baseline) in three groups of pre-service primary school teachers that pose and assess modelling problems. Our findings show that each approach presents distinct advantages and disadvantages. Extensive teacher theoretical-practical training seems to help deepen the design and analysis of tasks, especially in the understanding and mathematical and didactic analysis of the tasks, while the use of a didactic guide offers a more efficient means of introducing the characteristics of modelling tasks. Thus, using well-designed teaching materials, with examples and explanations of modelling and its relation to the curriculum, coupled with practical training, seems to be an effective measure to enhance future teachers' ability to design modelling problems.

Lastly, our findings hold significant implications for public administrations tasked with updating teacher training in response to curricular reforms incorporating modelling in primary school classrooms. Thus, this study provides evidence for the need to accompany curricular changes with specific teacher training designs or, at least, by providing well-designed teaching guides that respond to teachers' professional needs for modelling.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s11858-025-01658-3.

**Acknowledgements** The work presented is supported by the PID2021-126717NB-C31 and PID2021-126707NB-I00 (MCIU/AEI/FEDER, UE) research projects.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

## Declarations

**Conflict of interest** The authors declare there are no potential conflicts of interest.



## References

Abu-Elwan, R. (2002). Effectiveness of problem posing strategies on prospective mathematics teachers' problem solving performance. *Journal of Science and Mathematics Education in Southeast Asia, 25*(1), 56–69.

Barquero, B. (2023). La modelización matemática en la formación del profesorado: experiencias con los REI-FP para educación Primaria [Mathematical modelling in teacher education: experiences with the SRP-TE for primary school education.]. *RIFOP, 98*(37.2), 165–186. https://doi.org/10.47553/rifop.v98i37.2.99313

Barquero, B., Bosch, M., & Wozniak, F. (2024). Moving beyond mute modelling praxeologies in pre-service elementary teacher education. In S. Siller, V. Geiger, & G. Kaiser (Eds.), *Researching mathematical modelling education in disruptive times.* Springer.

Blum, W., & Leiss, D. (2007). "Filling Up"—The problem of independence-preserving teacher interventions in lessons with demanding modelling tasks. In M. Bosch (Ed.), *Proceedings of CERME 4* (pp. 1623–1633). ERME.

Borromeo Ferri, R. (2018). *Learning how to teach mathematical modeling in school and teacher education*. Springer.

Bosch, M. (2018). Study and research paths: A model for inquiry. In B. Sirakov, P. N. de Souza, & M. Viana (Eds.), *International Congress of Mathematicians* (Vol. 3, pp. 4001–4022). World Scientific Publishing. https://doi.org/10.1142/9789813272880_0210

Brunner, K., Obersteiner, A., & Leuders, T. (2024). How pedagogical content knowledge sharpens prospective teachers' focus when judging mathematical tasks: An eye-tracking study. *Educational Studies in Mathematics, 115*(2), 177–196. https://doi.org/10.1007/s10649-023-10281-6

Cabassut, R., & Ferrando, I. (2017). Difficulties in teaching modelling: A French-Spanish exploration. In G. A. Stillman, W. Blum, & G. Kaiser (Eds.), *Mathematical modelling and applications: Crossing and researching boundaries in mathematics education* (pp. 223–232). Springer.

Cai, J., & Hwang, S. (2021). Teachers as redesigners of curriculum to teach mathematics through problem posing: Conceptualization and initial findings of a problem-posing project. *ZDM, 53*(6), 1403–1416. https://doi.org/10.1007/s11858-021-01252-3

Cai, J., Hwang, S., Jiang, C., & Silber, S. (2015). Problem-posing research in mathematics education: some answered and unanswered questions. In F. Singer, N. Ellerton, & J. Cai (Eds.), *Mathematical problem posing research in mathematics education.* Springer. https://doi.org/10.1007/978-1-4614-6258-3_1

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

Chapman, O. (2012). Prospective elementary school teachers ways of making sense of mathematical problem posing. *PNA, 6*(4), 135–146. https://doi.org/10.30827/pna.v6i4.6137

Chen, L., van Dooren, W., & Verschaffel, L. (2015). Enhancing the development of Chinese fifth-graders' problem-posing and problem-solving abilities, beliefs, and attitudes: A design experiment. In F. M. Singer, N. F. Ellerton, & J. Cai (Eds.), *Mathematical problem posing: From research to effective practice* (pp. 309–329). Springer. https://doi.org/10.1007/978-1-4614-6258-3_15

Crespo, S., & Sinclair, N. (2008). What makes a problem mathematically interesting? Inviting prospective teachers to pose better problems. *Journal of Mathematics Teacher Education, 11*, 395–415. https://doi.org/10.1007/s10857-008-9081-0

Czocher, J. A. (2018). How does validating activity contribute to the modeling process? *Educational Studies in Mathematics, 99*, 137–159. https://doi.org/10.1007/s10649-018-9833-4

Fernández-Blanco, T., Gorgal-Romarís, A., Núñez-García, C., & Sequeiros, P. G. (2022). Prospective primary teachers' didactic-mathematical knowledge in a service-learning project for inclusion. *Mathematics, 10*(4), 652. https://doi.org/10.3390/math10040652

Ferrando, I., Segura, C., & Castillo, J. (2024). Regular and talented students' behaviour when solving modelling tasks: are there differences? In S. Siller, V. Geiger, & G. Kaiser (Eds.), *Researching mathematical modelling education in disruptive times*. Springer.

Galbraith, P. (2007). Dreaming a "possible dream": More windmills to conquer. In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical modelling: Education, engineering and economics* (pp. 44–62). Horwood.

Geiger, V., Galbraith, P., Niss, M., & Delzoppo, C. (2022). Developing a task design and implementation framework for fostering mathematical modelling competencies. *Educational Studies in Mathematics, 109*(2), 313–336. https://doi.org/10.1007/s10649-021-10039-y

Goldin, G. A., & McClintock, C. E. (Eds.) (1979). *Task variables in mathematical problem solving*. Information Reference Center (ERIC/IRC), The Ohio State University.

Hansen, R., & Hana, G. M. (2015). Problem posing from a modelling perspective. In F. Singer, N. Ellerton, & J. Cai (Eds.), *Mathematical problem posing research in mathematics education* (pp. 35–46). Springer. https://doi.org/10.1007/978-1-4614-6258-3_2

Hartmann, L., Krawitz, J., & Schukajlow, S. (2021). Create your own problem! Do students pose modelling problems that are based on given descriptions of real-world situations? *ZDM, 53*(4), 919–935. https://doi.org/10.1007/s11858-021-01224-7

Hartmann, L. M., Krawitz, J., & Schukajlow, S. (2023). Posing and solving modelling problems—Extending the modelling process from a problem posing perspective. *Journal Für Mathematik-Didaktik, 44*(2), 533–561.

Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development, 45*(1), 65–94.

Kaiser, G., Schwarz, B., & Tiedemann, S. (2010). Future teachers' professional knowledge on modeling. In R. Lesh, P. L. Galbraith, C. R. Haines, & A. Hurford (Eds.), *Modeling students' mathematical modeling competencies* (pp. 433–444). Springer. https://doi.org/10.1007/978-1-4419-0561-1_37

Klein, S., & Leikin, R. (2020). Opening mathematical problems for posing open mathematical tasks: What do teachers do and feel? *Educational Studies in Mathematics, 105*(3), 349–365. https://doi.org/10.1007/s10649-020-09983-y

Krawitz, J., Hartmann, L., & Schukajlow, S. (2024). Do task variables of self-generated problems influence interest? Authenticity, openness, complexity, and students' interest in solving self-generated modelling problems. *The Journal of Mathematical Behavior, 73*, 101129. https://doi.org/10.1016/j.jmathb.2024.101129

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

Maaß, K. (2006). What are modelling competencies? *ZDM Mathematics Education, 38*, 113–142. https://doi.org/10.1007/BF02655885

Maaß, K. (2010). Classification scheme for modelling tasks. *Journal Für Mathematik-Didaktik, 31*(2), 285–311. https://doi.org/10.1007/s13138-010-0010-2

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

Montes, M., Chico, J., Martín-Díaz, J. P., & Badillo, E. (2024). Mathematics teachers' specialized knowledge mobilized through problem transformation. *The Journal of Mathematical Behavior, 73*, 101132. https://doi.org/10.1016/j.jmathb.2024.101132

Niss, M., & Blum, W. (2020). *The learning and teaching of mathematical modelling*. Routledge.

Paolucci, C., & Wessels, H. (2017). An examination of pre-service teachers' capacity to create mathematical modeling problems for children. *Journal of Teacher Education, 68*(3), 330–344. https://doi.org/10.1177/0022487117697636

Passarella, S. (2021). Mathematics teachers' inclusion of modelling and problem posing in their mathematics lessons: An exploratory questionnaire. *European Journal of Science and Mathematics Education, 9*(2), 43–58. https://doi.org/10.30935/scimath/10773

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, 49*(3), 367–378. https://doi.org/10.1007/s11858-016-0819-3

Silber, S., & Cai, J. (2017). Pre-service teachers' free and structured mathematical problem posing. *International Journal of Mathematical Education in Science and Technology, 48*(2), 163–184. https://doi.org/10.1080/0020739X.2016.1232843

Stoyanova, E., & Ellerton, N. F. (1996). A framework for research into students' problem posing in school mathematics. In P. C. Clarkson (Ed.), *Technology in mathematics education* (pp. 518–525). Mathematics Education Research Group of Australasia.

Szydlik, J. E., Szydlik, S. D., & Benson, S. R. (2003). Exploring changes in pre-service elementary teachers' mathematical beliefs. *Journal of Mathematics Teacher Education, 6*(3), 253–279. https://doi.org/10.1023/A:1025155328511

Tomás Henao, J. (Coord.) (2023). *Matemáticas 6 Primaria Proyecto La casa del Saber*. Santillana.

Turner, E. E., Bennett, A. B., Granillo, M., Ponnuru, N., Roth Mcduffie, A., Foote, M. Q., et al. (2024). Authenticity of elementary teacher designed and implemented mathematical modeling tasks. *Mathematical Thinking and Learning, 26*(1), 47–70. https://doi.org/10.1080/10986065.2022.2028225

Vicente, S., & Manchado, E. (2017). Dominios de contenido y autenticidad: un análisis de los problemas aritméticos verbales incluidos en los libros de texto españoles [Content domains and authenticity: an analysis of verbal arithmetic problems in Spanish textbooks]. *PNA, 11*(4), 253–279. http://hdl.handle.net/10481/47547

Wess, R., Klock, H., Siller, H.-S., & Greefrath, G. (2021). Measuring professional competence for the teaching of mathematical modelling. In F. K. S. Leung, G. A. Stillman, G. Kaiser, & K. L. Wong (Eds.), *Mathematical modelling education in east and west* (pp. 249–260). Springer. https://doi.org/10.1007/978-3-030-66996-6_21