ARTICLE

Modeling a Theoretical Construct on Pupils' Difficulties in Problem Solving

Bartolomé Vázquez-Bernal[1] (ORCID) · Roque Jiménez-Pérez[1]

Abstract

The objective of this work was the theoretical modeling of a construct based on teaching practice about the perception that pupils have of difficulties in problem solving (PS) in experimental sciences, specifically physics, to predict pupil performance in PS. The research was carried out with an incidental sample of second year of secondary education pupils, of 13 to 14 years in age. The pupils and their teachers were involved in a project in which PS, classroom inquiry, and Hooke's law were considered to be central. The information collection instrument was a Likert type questionnaire to investigate the various difficulties perceived by the pupils (Knowledge, Trust, Path, Interest,…) in the problem-solving processes applying Hooke's law. A theoretical model was estimated and not rejected that involved only one factor (a PS factor), with all of its variables strongly interrelated and a set of constraints represented by covariances between theoretical hidden variables. This model is a starting point from which to understand how pupils at this level perceive PS in experimental sciences.

1 Introduction

What makes it difficult to solve a problem of a scientific nature? Jonassen (2007) responds to this question by stating that the reason is the interaction between the participants, the activity, and the context. When school contexts are added to the research aspects (Östman & Wickman, 2014), the result is a very fruitful field in which to study the interrelationships, obstacles, and pupils' perceptions during this process.

The objective of the present study was to formulate an initial theoretical approach linked to pupils' perceptions about difficulties in problem solving (PS) in the Physics classroom, based on a set of variables considered to be relevant to the process, and then, through the existing knowledge in the scientific literature, estimate the interrelationships among those variables and submit this estimate to empirical tests by means of structural equation modeling (SEM). Underlying the set of variables is our acceptance of Bachelard's idea (Bachelard, 1983) that the obstacles are a form of knowledge which has in general been satisfactory for solving certain problems, but unsuitable for confronting new ones.

✉ Bartolomé Vázquez-Bernal
bartolome.vazquez@ddcc.uhu.es

1 Integrated Didactics Department, University of Huelva, Huelva, Spain

The importance of PS, and more specifically its collaborative version, was currently the object of special attention internationally through PISA 2015 and the program "Collaborative Problem Solving", in which the skills of providing solutions, sharing knowledge, and making efforts towards answering posed questions were assessed (OECD, 2017). Recently, PISA, 2018 has examined what students know in science and what they can do with what they know (Schleicher, 2019): 78% of students attained Level 2 (basic procedural and epistemic knowledges) or higher in science, on average across OECD countries (in Spain, 78.7% students). Only 0.8% of students attained Level 6 (highest level: students can evaluate competing designs of complex experiments, field studies, or simulations and justify their choices); remarkably very poor results in a society as technological as ours.

In light of these results, it would be interesting to know what perceptions students have about the obstacles in PS, in line with what some studies have found, for example, with experimental PS: assigned time, instruction sheets, attendance and preparedness (Deacon & Hajek, 2011), or that good performance on ill- and well-structured problems is sensitive to different social network configurations (Pulgar et al., 2020). As the linear problems are the exception and the complex ones the norm in Physics, a better understanding of students' views would help to understand the processes of learning physics from a research perspective: most influential obstacles, existence of latent variables, among others, in order to ultimately know their influence on learning outcomes and to incorporate these insights into teaching focused on these perceptions.

2 Material and Methods

In our study, each variable is associated with underlying obstacles that determine students' difficulties (Astolfi, 1999; Vázquez-Bernal & Jiménez-Pérez, 2016), as detailed in Annex 1. To begin this section, we will discuss the studies that have had a general impact on the understanding of the factors that may affect students' problem-solving in science learning and then focus on the studies that analyse their perceptions in the field of physics, and which have some kind of relationship with the variables under study (see shortened form in Annex 1). Next, we shall then discuss only those specific studies which have used SEM with their focus being on mathematics, science, and technology. Finally, we will address the role of gender in PS.

2.1 Problem-Solving in Science Learning

In an attempt to model the impact on pupils' behaviour in performing their school tasks (Tarhini et al., 2016), it was discovered that their confidence (*Di*), motivation (*In*), habit, and expectation of their performance (*Co*) have positive effects. Other research (Dunkley & Blankstein, 2000) illustrates how coping strategies (*Pa*), emotions, tasks, and evasion, associated with self-critical perfectionism can have important consequences for pupils' experience of their daily school tasks. In studies of structural relationships between the pupils' conceptions and their approaches to science learning (Lee et al., 2008), the results reveal that pupils who had constructivist conceptions of science learning (*Kn*) tended to use profound approaches to it. Another influence on the pupils seems to come from the adoption of new forms of learning, since attitude (*Ef*) is the variable that stands out most in some models (Park et al., 2012).

Some studies have modeled which factors could predict success in learning, finding that a relevant negative factor was a pupil's undesirable (*In*) and withdrawal behaviour (*Ef*) (Normandeau & Guay, 1998). In this sense, models have shown that self-efficacy directly affects pupils' achievement (Akın, 2010), as also does social support given to them (Nugent et al., 2015). Another component in PS, the speed of information processing (*Pa*), can predict school performance, but by way of other higher-level cognitive skills such as intelligence and creativity (Rindermann & Neubauer, 2004). In the case of Physics PS, and particularly equation solving, it has been found that, by the last year of primary education, most pupils already had a basic relational view of equivalence (*Ca*) and began to compare the two sides of an equation (Rittle-Johnson et al., 2011), something that they should delve into, as it is a necessity in Physics PS.

2.2 Perceptions of Students in the Field of Physics

The study of the perceptions of students in the field of Physics is not extensively investigated, being important, because according to Irving and Sayre (2015), future Physicists are characterized by their commitment to research and their interest in a profound understanding of physical phenomena (*In*). In this sense, perceptions of students are an essential element to consider, as they define the rules of interaction in a classroom largely (Turpen & Finkelstein, 2010).

Prosser et al.. (1996) determined that students perceive learning Physics as hard work (*Ef*), but unrelated to how to study the subject matter, seeking reproduction before understanding (*Un*) and combining interest in Physics (*In*) with success, but for others, not for oneself. In a study of perceptions with secondary school students developed by Fonseca and Conboy (2006), it was found that the major factors of failure were quality of teaching and previous student, as well as the low expectations of success and preparation for life. For their part, Silva et al. (2019) have shown that Hooke's law is an adequate opportunity to introduce secondary school students to true investigative methodologies in the classroom, instead of more practical ways (sequence of steps), in tune with PISA level 6.

2.3 SEM for PS in Mathematics, Science, and Technology

The following paragraphs will deal with studies that used SEM for PS in three specific domains: Mathematics, Science, and Technology. In Mathematics, SEM has been used to test the predictive and mediating role that self-efficacy beliefs play in PS, and supports the role of self-efficacy hypothesized in social cognitive theory (Pajares, 1996). It has also been used to formulate and validate four cognitive processes in PS—editing the problem, filtering important and critical information, comprehending the structural relationships in quantitative information (*Pa*), and translating the quantitative information (*Ca*) from one mode to another (Pittalis et al., 2004). In geometry, SEM has been used to examine the effects of knowledge about cognition and of the regulation of cognition on declarative, conditional, and procedural knowledge, finding that the strongest direct effect comes from knowledge about cognition (Aydin, 2007). Other applications have been: to examine the effects of two psychosocial features of the classroom environment—teacher support and personal relevance—on undergraduates' academic self-efficacy and enjoyment of mathematics classes, finding teacher support and personal relevance to be influential (Aldridge et al., 2012); to examine and validate the explanatory and predictive relationships among the variables Mathematical Problem-Solving and Reasoning Skills, Sources

of Mathematics Self-Efficacy, Spatial Ability, and Mathematics Achievements (Yurt & Sünbül, 2014); to determine the effects of mathematical PS skills (*Ca*) and prior mathematics knowledge on mathematics achievement (*Kn*) and self-efficacy (Kamalimoghaddam et al., 2016); to analyse the relationships between executive functions, IQ, and math abilities, finding that working memory (*Me*) and age predict number production and mental calculus, and shifting and sex predict arithmetical PS (Arán & Richaud, 2017); and to suggest that structured, semi-structured, and free problem-posing activities improve pupils' PS skills and metacognitive awareness (Akben, 2018). In summary, most of the previous studies have a predictive character in the performance of the student in PS based on different variables: self-efficacy, cognitive processes (*Pa*), teacher support, personal relevance, and prior mathematics knowledge (*Kn*).

In Science in general, SEM has been used to study the relationships between motivational beliefs, metacognitive strategies, and effort regulation (*Ef*) (Sungur, 2007); to explore the connections between new knowledge (*Kn*) and meaningful learning in Ausubel and Novak's construct (Brandriet et al., 2013); to assess pupils' understanding of the particulate nature of matter (*Kn*), and its collective properties and physical changes (Stamovlasis et al., 2012); to indicate that a motivating science class (*Cl*) and a family who has positive attitudes towards science and are to some degree engaged with science may influence the pupils to adopt deeper approaches to science learning (Soltani, 2018); to suggest that pupil cohesiveness, inquiry, and task orientation (*Pa*) are the most influential predictors of pupil motivation (*In*) and self-regulation in science learning (Velayutham & Aldridge, 2013); and to develop models describing academic self-concept in science (Hardy, 2014). In summary, it is mostly used to modeling some motivational aspects (*In*) and their relationship with science learning.

Finally, in Technology, SEM has been applied to analyse the antecedents affecting pupils' Web-based PS performance, finding that the task-technology fit could be the main factor affecting the pupils' intention to learn from the Web, (Kuo & Hwang, 2015); and to study the theoretical framework of technological, pedagogical, and content knowledge (TPACK), finding that pedagogy is TPACK's core element (Celik et al., 2014). That is, for predicting the Web-based PS performance and theoretical constructs about knowledge.

Concluding this review, basically the majority use the SEM in predicting the performance of the student in PS through motivational aspects (*In*), but there are aspects related to how the students perceive some of these variables in the PS, so we believe it is important to fill this gap, since the existence of possible latent variables, together with their predictive nature, can help us to improve their skills in SP.

2.4 Gender

Neither has the role of gender and motivation in PS tasks gone unnoticed. The results of a study analysing the relationship between gender and entrepreneurship in these tasks cannot be explained solely by differences in motivation (Haus et al., 2013). The role of gender and emotional intelligence in tests and examinations has also been modeled, with positive predictions that favour females (Austin et al., 2005). With regard to motivation (*In*), it is has been confirmed that pupil self-regulation is a strong predictor of science learning (Velayutham et al., 2012).
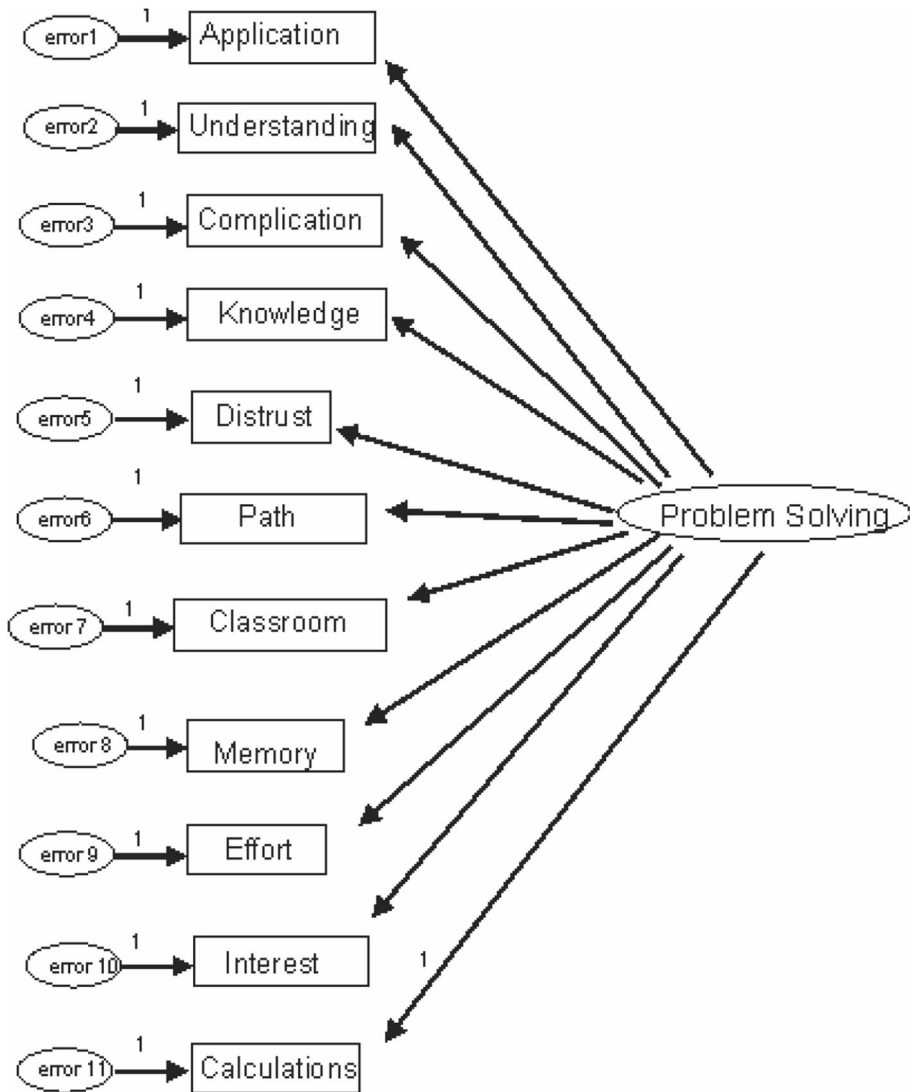
Fig. 1 Conceptual diagram of the initial Problem-Solving Difficulties Model A

2.5 Objective

The main objective of this study was the formulation of a theoretical model for PS in the Physics classroom. Figure 1 shows this model (AMOS© notation), which has been called "Conceptual diagram of the initial Problem-Solving Difficulties Model A". This model would be subjected to empirical testing (SEM) to check whether it is rejected or not. The variables that make up this model are based on a previous study that will be described in the "Measurement Instruments" subsection. Our main hypothesis is that all these difficulty variables are so closely interrelated that there is only a single latent factor, although there

could be internal correlations in aspects not directly measured by the error variables. These errors represent more than just random fluctuations, indicating unmeasured aspects of the variables in our study. In addition, we want to know whether the parameter gender has some type of influence on any of the variables in the study.

2.6 Design

According to Latorre et al. (1996), this study is an ex-post-facto non-experimental study with a quantitative methodology called SEM, a multivariate quantitative technique used to describe the relationships between observed variables, and which helps the researcher to test or validate a theoretical model for theory testing and extension (Thakkar, 2020).

Convenience sampling was used since the education centres involved formed part of a research project of the Regional Government of Andalusia (Spain) involving action research in the classroom (7 teachers specialising in Physics and Chemistry unrelated to referees). The focus of this project is the use of inquiry-based methods (Windschitl et al., 2008; Bevins & Price, 2016) that carry out practical work with pupils, examining their perceptions which affect PS, and the obstacles, emotions, self-efficacy, and self-concept, under the conceptual umbrella of the European programme "Inquiry-Based Science Education" (European Commission, 2007).

The data collection was associated with that project and lasted four school years (from 2015 to 2019). The classroom inquiry was based on the question "*What is the relationship between the length a spring stretches and the force applied*?" With this question as basis, the pupils were introduced into a process of inquiry in which the initial ideas are presented in the form of hypotheses, possible designs are put forward, and a consensus is reached between the teacher and the groups in class as to what the optimal design should be. Measurements are made, the results are debated and confronted with each pupil's initial hypotheses, and a research report is written up. Later, problems are taken on that are more closed—around the particular topic of Hooke's law. Finally, a written numerical test is given (see Annex 2).

The specified questionnaire is carried out approximately 1 month later. Under the same curriculum every year, there was a total flow of communication between researchers and teachers, with visits to classrooms, although the teaching strategies implemented by teachers could vary. We used the research report on Hooke's law to know the perceptions of the students and to discard students with a low level of commitment in the completion of the questionnaire (less than 10 in the sample). Trying to make this investigation as natural as possible, a specific test for Hooke's law was not considered but was included in the Unit of Forces and Movement, hence the month that passed.

2.7 Participants

The variables under study affect basic aspects of PS in a natural context, associated with school inquiry processes, in which a group of Physics and Chemistry teachers and their pupils took part. The study involved 528 pupils from various public secondary schools from the medium–low social level in the same geographical area of southern Andalusia (Spain). All the pupils were in the 2nd year of Lower Secondary Education (ESO, the acronym in Spain for compulsory secondary education) with an 8% bias in the distribution by gender towards more males, and mostly of 13 or 14 years in age. Table 1 lists the distribution of the participants by age and gender.

Table 1 Distribution by age and gender of the participants

Educational level	Pupils	Age distribution (%)	Gender distribution (%)
2nd ESO (lower secondary)	528	12 years: 4 (0.8%) 13 years: 332 (62.9%) 14 years: 144 (27.3%) 15 years: 46 (8.7%) 16 years: 1 (0.2%) 17 years: 1 (0.2%)	243 females: (46%) 285 males: (54%)

Table 2 Problem-solving difficulty survey

Gender: M	F	Age:	Town:

Please value from 1 to 5:
1 (completely disagree); 2 (mostly disagree); 3 (agree); 4 (mostly agree); 5 (completely agree)

Variables	When I solve problems related to Hooke's law:
Application (Ap)	I know the theory, but I do not know how to apply it to the problem
Understanding (Un)	I do not understand what the problem is saying
Knowledge (Kn)	I have a lack of knowledge, I do not know what is necessary
Memory (Me)	My memory fails, I forget what I already know
Effort (Ef)	Lack of work and study
Interest (In)	Lack of interest
Complication (Co)	They are too complicated
Classroom (Cl)	I do not understand the problems that we have done in class
Path (Pa)	It is unclear what is the way to solve them
Distrust (Di)	I know before I start that it will not turn out well
Calculations (Ca)	I fail in mathematical calculations and operations

2.8 Measurement Instruments

In the original study, the survey was developed by 15 experienced Physics teachers (referees) and was answered by 419 Spanish students between 18 and 22 years of age. Those referees used interviews with students to verify that the writing and understanding of the items was adequate. This process was verified by our group, selecting three students with different competences in PS in order, through written interviews, to effectively assess that adequacy. A doubly validated and student-centred instrument induced us to use it, although no statistical validation data was reported in the original study (Oñorbe & Sánchez, 1996).

The questionnaire (Table 2) consists of items scored on a Likert scale with a range from 1 (no agreement) to 5 (full agreement).

Another aspect was the breadth of the scale. It has been pointed out that including five to seven values in the scale increases reliability and helps the data adjust to a normal distribution (Matas, 2018).

It was decided to obtain a large sample of questionnaires in order to reduce the impact of any lack of normality (Cupani, 2012), since non-normality was expected given the experience of studies with samples of very similar characteristics (Vázquez-Bernal &

Table 3 Descriptive statistics of the variables measured

Variables	Mean	Std. deviation	Median	Variance	Confidence interval (95%)	Skewness	Kurtosis
Complication (*Co*)	2.621	1.243	2	1.470	(2.49–2.76)	0.386	- 0.743
Application (*Ap*)	2.585	1.214	3	1.432	(2.45–2.72)	0.323	- 0.685
Path (*Pa*)	2.466	1.225	2	1.412	(2.33–2.60)	0.566	- 0.464
Memory (*Me*)	2.422	1.298	2	1.699	(2.28–2.57)	0.546	- 0.817
Calculations (*Ca*)	2.383	1.262	2	1.562	(2.24–2.52)	0.606	- 0.617
Knowledge (*Kn*)	2.360	1.200	2	1.560	(2.22–2.50)	0.663	- 0.523
Classroom (*Cl*)	2.305	1.238	2	1.473	(2.17–2.44)	0.762	- 0.300
Understanding (*Un*)	2.216	1.260	2	1.408	(2.08–2.35)	0.810	- 0.194
Effort (*Ef*)	2.172	1.251	2	1.624	(2.03–2.31)	0.863	- 0.375
Distrust (*Di*)	2.127	1.265	2	1.645	(1.98–2.27)	0.965	- 0.175
Interest (*In*)	1.936	1.179	2	1.359	(1.81–2.07)	1.196	- 0.486

Jiménez-Pérez, 2016). The instrument consists of 11 variables that represent different difficulties associated with obstacles of different origins (Annex 1). With regard to the validity and reliability of the instrument, the frequency analysis showed the variables to be skewed towards the upper values of the Likert scale (3 and 4). Given then that a non-normal distribution of the population was to be expected, the non-parametric Kolmogorov–Smirnov test was applied. The null hypothesis (H_0) was rejected for all the variables as they gave values below the 5% significance level ($p < 0.000$). It was therefore assumed that the sample came from a population with the variables non-normally distributed.

McDonald's omega (McDonald, 1999) is preferred to Cronbach's alpha when the variables are ordinal or there are fewer than seven eligible alternatives. This statistic gave a value of 0.888, a very satisfactory and robust value in exploratory research (Viladrich et al., 2017). For DeCarlo (1997), Likert scales often have multiple variance problems that affect multivariate kurtosis. The kurtosis had a maximum positive value of 0.479, while negative values ranged from -0.822 to -0.180. As a value above 4 would have indicated deviation from normality, it can be argued that there is no substantial kurtosis in the distributions. Taking the variables together, the value of the critical ratio (CR) should be less than or equal to 5, which was met in all cases. The value of the multivariate CR however, 24.462, indicated a departure from normality, so that it was justified to opt for estimation with asymptotically distribution-free (ADF) techniques. In addition, it suggested that the sample size should be at least ten times the number of parameters estimated (i.e., 250, exceeded by our 528 cases).

3 Results

In the "Results" section, we will first address the descriptive statistics in relation to the study variables (difficulties) and the number of latent factors that our model must include. Secondly, we will show the SEM modeling process of the variables and, finally, the results related to gender influence.

Table 3 lists the descriptive statistics comprising the initial exploratory analysis. It can be seen that the perception of the greatest difficulty is that of Complication. In the way the pupils see the problems, in our opinion, this is a didactic obstacle and how the subject is taught in the classroom; perhaps due to the custom of using overly linear solving problems. The least difficulty is Interest (associated with intrinsic motivation, and therefore a personal obstacle). These values coincide with the values assigned in the article by Oñorbe and Sánchez (1996), so they were not surprising. Is that a sort of inherent obstacle or one that the authors think was specific to how the subject was being taught in the participating classrooms?

Regarding Interest, we can ask ourselves if the absence of this obstacle is due to the teaching of the topic related to Hooke's law or it is something more extensive. We are inclined towards the former, since the students were asked to focus on this single topic intensively when taking the questionnaire.

According to Muthén and Kaplan (1992), it is advisable to use polychoric correlation when the univariate distributions of ordinal items are asymmetric (skewed) or have excessive kurtosis, i.e., unless both indices are less than unity in absolute value. In Table 3, one sees that the skewness of the variable Interest exceeds unity, so we decided to work with polychoric correlation to decide which factors to extract and contrast with the original study (Oñorbe & Sánchez, 1996).The polychoric correlation matrix (for lack of space, not specified), computed using Bayesian modal estimation (Choi et al., 2011), presented values between 0.231 and 0.654—values that indicate strong correlation.

To determine the number of factors our model should have, we used the parallel analysis and Velicer's minimum average partial test (revised). The parallel analysis based on minimum rank factor analysis (Timmerman & Lorenzo-Seva, 2011) advised for the extraction of a single factor, giving a variance explained of 56.3%, with positive matching coefficients (Bartlett's statistic 0.000010 and KMO test 0.90265). This result was corroborated by the Velicer's minimum average partial test (with values between 4.6603 and 0.3502), and refutes the five dimensions extracted in the original study. We thus proceeded in accordance with the present outcome.

We decided to apply SEM to confirm the strong correlation of the variables and the extraction of a single factor. In this approach, a model is specified from among all the pairs of variables (Arbuckle, 2011). The analysis was carried out with the AMOS 20$^{©}$ program, obtaining significant correlations ($p < 0.01$) between all the pairs of variables. Although the degree of significance was consistent with that found in the two previous analyses (parallel and Velicer), this time, the fit of the model was based on ADF techniques. The (standardized and non-standardized) residual covariance matrices were calculated, all presenting values of 0.000, ensuring that there are no significant discrepancies.

3.1 Structural Equation Modeling for Problem-Solving Difficulties: Explanation of the Initial Model

In this subsection, we shall indicate the initial hypotheses in support of the model shown in Fig. 1. The results of these preliminary analyses and the absence in the literature of studies that deepen the perception of the difficulties that adolescent students have in PS in Physics and specifically on Hooke's law and PS, led us to think that, as a first approximation, the most plausible interpretation would be that there is only one factor underlying the exogenous (observed) variable. We therefore proposed a first-order factor or endogenous (unobserved) variable—Problem Solving Factor—that covers all the indicator variables involved in the process. Each indicator variable has a dependence on a variable termed error, designed to represent not just random fluctuations but also aspects not measured in the study on which the indicator variable could depend, assigned the arbitrary value 1 to avoid problems of identification (an indirect way of choosing a unit of measurement for the error).

While maximum likelihood (ML) estimation is efficient and unbiased when multivariate normality assumptions are met, this is precisely what our sample suffers from. In these circumstances, ADF techniques provide more benefits provided the sample is sufficiently large, as explained above. The results indicated that model A has to be rejected because the value of chi-squared (χ^2) of 164.99 for a distribution with 44 degrees of freedom (df) corresponds to a probability level less than the 0.05 threshold, thus rejecting the null hypothesis that the data of the sample coincide with those of the population from which it is drawn. Nonetheless, although model A is not accepted (probability level 0.000 and a χ^2/df ratio of 3.745), it is in terms that can be improved. We shall only present visually the standardized results for Model A (Fig. 2).

In summary, Fig. 2 shows that, for example, for the variable Understanding, the value 0.71 is the Standardised Regression Weight and expresses that when PS factor goes up by 1 standard deviation, Understanding goes up by 0.71 standard deviation. On the other hand, the value 0.51 represents the Multiple Squared Correlation (i.e., the predictors of Understanding explain 51% of its variance). In other words, the variance of the Understanding error is approximately 49% of the variance of Understanding itself. Values range reasonably well.

The first step in trying to reset the model is assessment using modification indices. We decided to construct and test a new Model B, including only those covariances which are significant and have theoretical justification (the constraints). These were error1 (*Ap*)–error2 (*Un*), error3 (*Co*)–error4 (*Kn*), error6 (*Pa*)–error7 (*Cl*), error7 (*Cl*)–error8 (*Me*), error8 (*Me*)–error9 (*Ef*), error9 (*Ef*)–error10 (*In*), and error8 (*Me*)–error11 (*Ca*). The previous justifications are supported by the original study by Oñorbe and Sánchez (1996), based on the factor analysis carried out by the authors, in which a set of underlying factors were found: Interest–Effort, on the one hand, and Calculus–Memory–Classroom–Pathway, on the other. Moreover, some works (Pozo & Gómez, 1998; Lee et al., 2008; Quílez, 2019), already link the Application–Comprehension association, which involves deep learning, its application and the scientific terminology used, especially in students with little experience in PS. Meanwhile, other works (Alpaslan et al., 2017; Jonassen & Hung, 2015; Yerushalmi & Magen, 2006), link between Complication–Knowledge, to the extent that students' perceptions are compatible with naïve epistemology without contextual appropriateness and linear problem-solving.
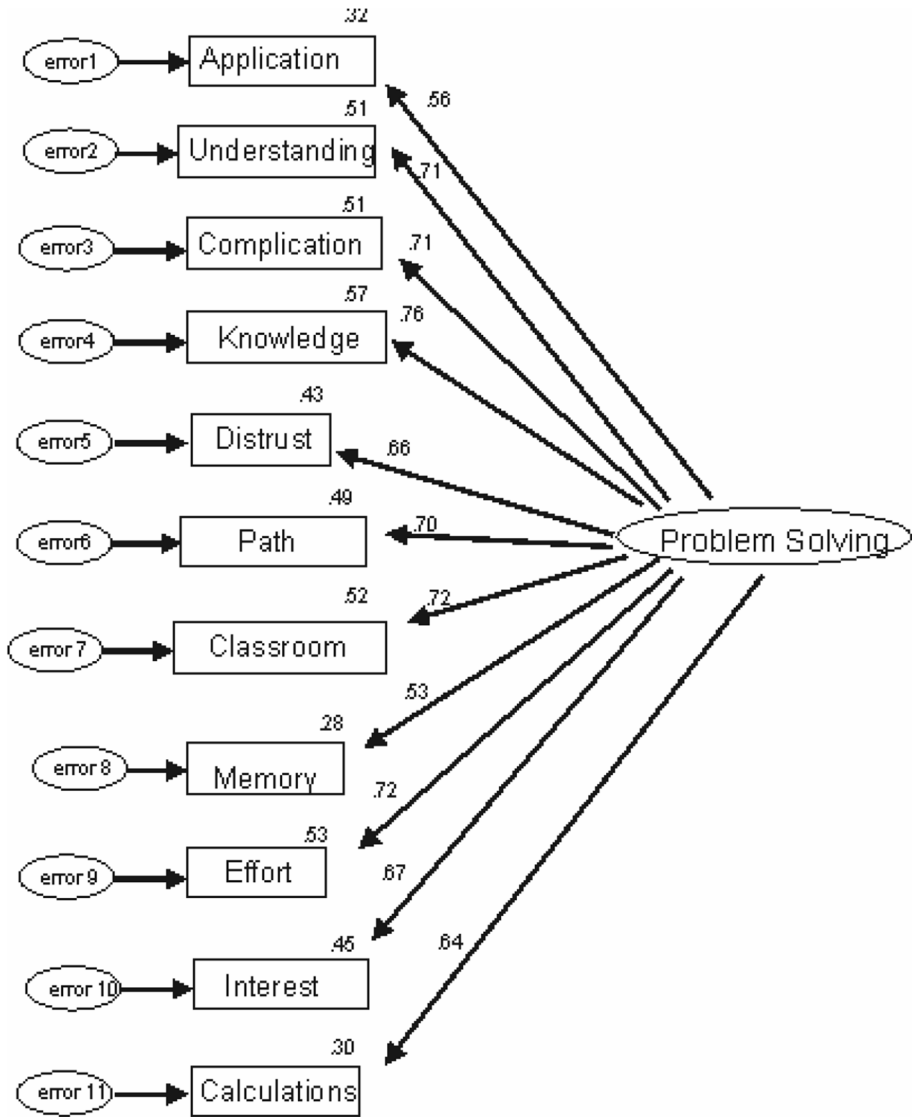
Fig. 2 Standardized outcomes for Model A (ADF estimation)

Delving into these previous ideas, the pupils' perceptions of their PS difficulties which are related to the Application and Understanding of these problems likely share hidden variables not measured in this study—in particular, the content (Hooke's law) has to be clearly understood for it to be applied in the different contexts in which the problems were set. The case is similar for the pupils' Knowledge of the concept and their perception of its Complication, which one would clearly expect to share some hidden variable. The procedure to follow in a problem's resolution (Path) is mediated by the perception of what is done in class (Classroom), hence the covariances shown. Memory plays a central role in
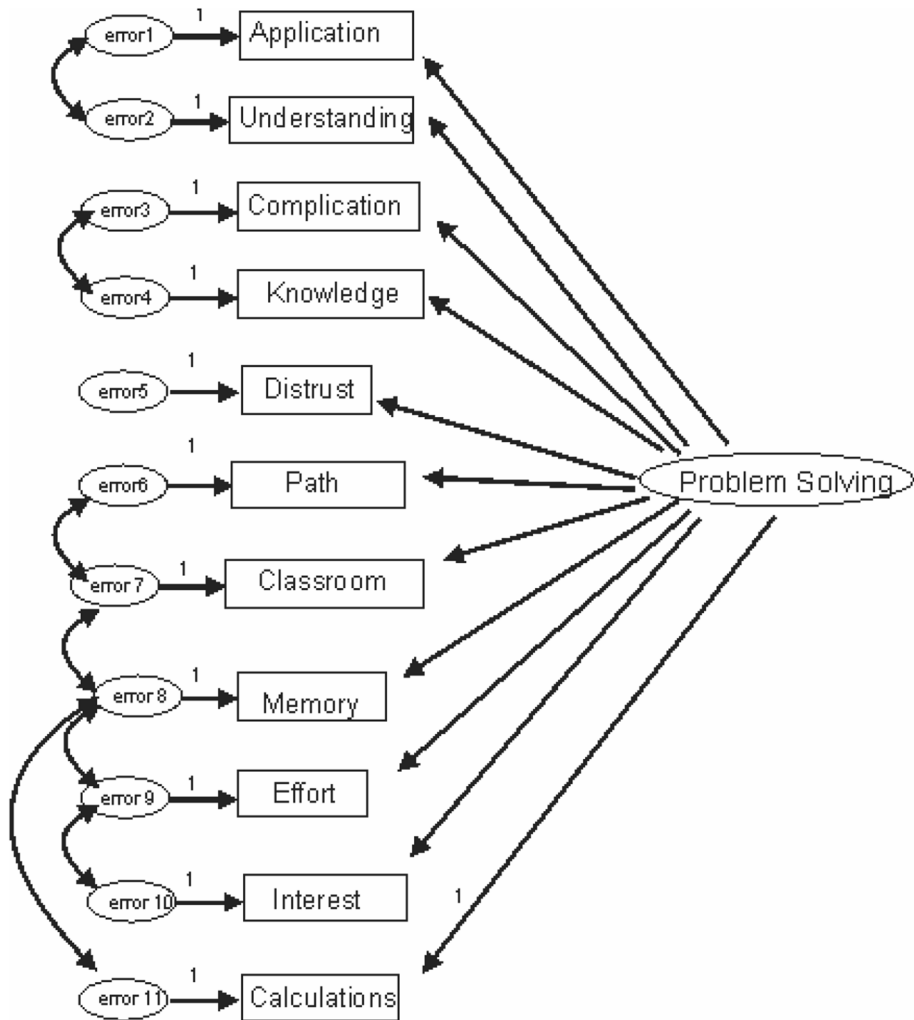
Fig. 3 Conceptual diagram of the initial Problem-Solving Difficulties Model B

covariance errors between the Effort needed for resolution, the Calculations that must be done, and the type of activities that are carried out in class (Classroom), and we found that the unmeasured variables in these processes share similar origins or causes. Finally, Effort and Interest showed error covariances, as it would seem reasonable since, if an effort is made, then there have been a minimum of motivation, so that again there might be shared hidden variables involved not measured in this study (Model B, Fig. 3).

In light of the results presented in Table 4, the introduction of restrictions led to improvement and settling of the model, since the model B has a percentage chance of 6.1%, above the minimum of 5% that is not rejected. We must remember that the higher the probability value associated with χ^2, the closer the fit between the model under hypothesis

Table 4 Comparative summary of the goodness-of-fit indices for Models A and B

Goodness-of-fit indices	Model A	Model B	Index evolution
χ^2	164.199	51.144	Null hypothesis (H_0) of Model B accepted
df	$66-22=44$	$66-29=37$	
χ^2/df	3.732	1.382	
Prob. level	0.000 (rejected)	0.061 (not rejected)	
RMSEA	0.072	0.027	Positive (lesser than 0.05)
RMR	0.123	0.048	Positive (lesser than 0.05)
GFI	0.906	0.971	Positive (higher than 0.95)
NFI	0.585	0.871	Positive (close to 0.95)
CFI	0.647	0.959	Positive (higher than 0.95)
TLI	0.559	0.938	Positive (close to 0.95)
AIC	208.199	109.144	
ECVI	0.395; ***LO 90***: 0.328; ***HI 90***: 0.477	0.207; ***LO 90***: 0.180; ***HI 90***: 0.251	Positive (below index of saturated model, 0.250)
MECVI	0.397	0.210	Positive (below index of saturated model, 0.256)
Hoelter	**0.05**: 195; **0.01**: 221	**0.05**: 538; **0.01**: 618	> 200 (adequate sample size)

(model A or B, assumed true H_0) and the perfect fit (Bollen, 1998). That is, we accept the null hypothesis that the sample data coincide with the population from which it comes. In this sense, the probability level being greater than 0.05 shows that model B is accepted (0.061).

In conclusion, all the indices evolved positively (see evolution in Table 4), validating the initial results and confirming that Model B fits the data well (Jöreskog & Sörbom, 1993; Hu & Bentler 1995; Arbuckle, 2007). Figure 4 visually presents the standardised results of model B. It can be seen that, for the variable Comprehension, the values are somewhat lower than those shown in Fig. 2 (0.71 and 0.51 respectively). The value between error1 and error2 indicates that 0.17 is the estimated correlation between error1 and error2. Overall, the statistical values are acceptable. The study of the total effects (direct and indirect effects of any given variable on others) showed that there were no inverse effects, although the direct standardized values of two variables, *Memory* and *Calculations*, were below 0.5, i.e., 0.434 and 0.472, respectively. Table 5 lists the values of the composite reliability (*CRe*) and the average variance extracted (*AVE*) calculated following Fornell and Larcker (1981). Those authors indicate recommended values of 0.7 and 0.5 for *CRe* and *AVE*, respectively, so that a good value was obtained for *CRe* (0.93), but that for *AVE* (0.46) is somewhat low. Since this is a new study, the latter value can be taken as acceptable with precaution.

3.2 Gender Influence

As the gender data were available and there was a fairly balanced proportion in the gender of the pupils, we decided to check whether there were any statistically significant differences and, if so, the corresponding statistical powers and effect sizes. The results are
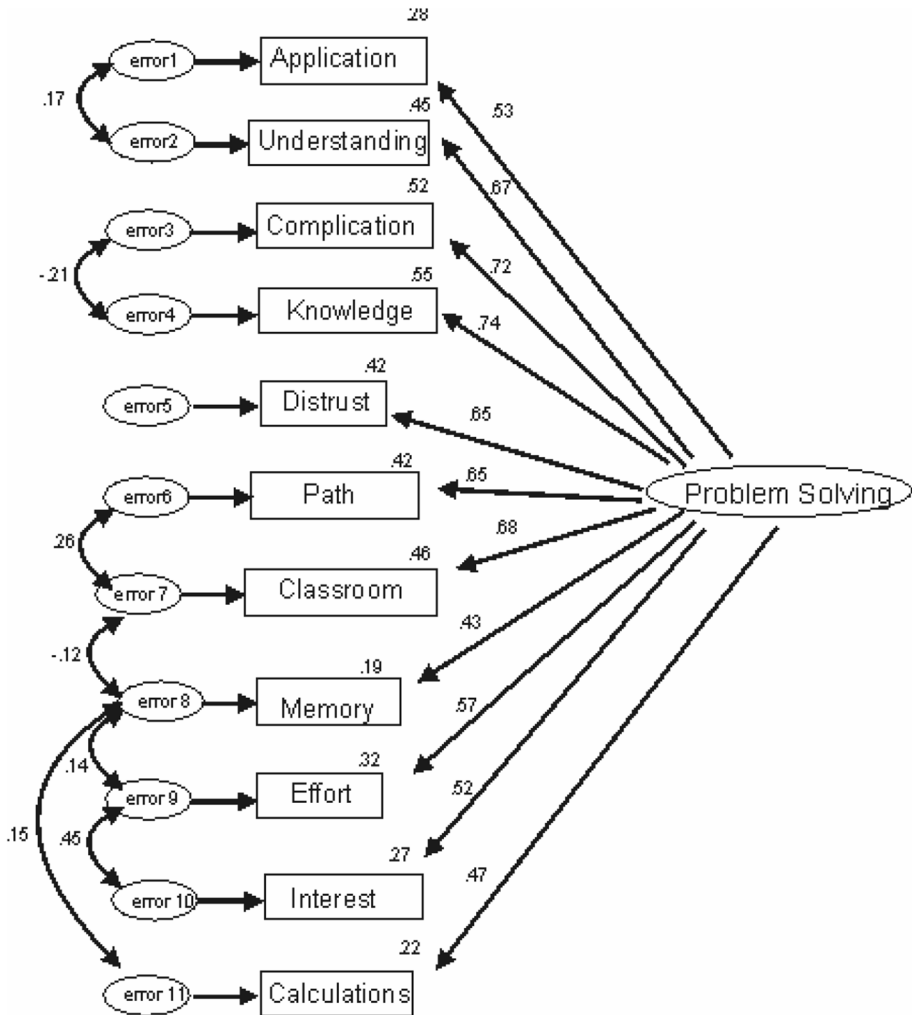
Fig. 4 Standardized outcomes for Model B (ADF estimation)

presented in Table 6. The female pupils gave higher scores to their perceptions of the variables *Complication*, *Understanding*, *Classroom*, and *Memory*—a combination of didactic and psychological obstacles (see Annex 1). Nonetheless, although the statistical powers were medium–high, the effect sizes were small. A plausible explanation for these differences is that the three didactic variables are extrinsic to the learner, and the underlying obstacles are associated with the way in which the teacher develops PS in class (*Un*), with the kind of interactions that are established in the classroom (*Cl*), and how the female pupils perceive the whole process (*Co*); intimate aspects related to the social environment that students perceive and to which they may plausibly be more sensitive, influencing their

Table 5 The *CRe* and *AVE* values for the reliability and validity of the (in-house developed) instrument of measurement (Model B—LISREL Notation for columns 1 to 4)

Endogenous latent variables	Indicator (Y_i) endogenous variable η_I	Construct variance λ_{ij}^{y}	$(\lambda_{ij}^{y})^2$	Var $\varepsilon_i = 1 - (\lambda_{ij}^{y})^2$	CRe	AVE
Problem-solving factor (η_1)	Application (*Ap*)	0.56	0.32	0.68	0.93	0.46
	Understanding (*Un*)	0.71	0.51	0.49	0.93	0.46
	Complication (*Co*)	0.71	0.51	0.49	0.93	0.46
	Knowledge (*Kn*)	0.76	0.57	0.43	0.93	0.46
	Distrust (*Di*)	0.66	0.43	0.57	0.93	0.46
	Path (*Pa*)	0.70	0.49	0.51	0.93	0.46
	Classroom (*Cl*)	0.72	0.52	0.48	0.93	0.46
	Memory (*Me*)	0.53	0.28	0.72	0.93	0.46
	Effort (*Ef*)	0.72	0.53	0.47	0.93	0.46
	Interest (*In*)	0.67	0.45	0.55	0.93	0.46
	Calculations (*Ca*)	0.64	0.30	0.70	0.93	0.46
		$\Sigma\lambda_{ij}^{y} = 7.38$ $(\Sigma\lambda_{ij}^{y})^2 = 54.46$	$\Sigma(\lambda_{ij}^{y})^2 = 4.91$	$\Sigma\text{Var}\,\varepsilon_i = 6.09$		

self-efficacy. Memory (*Me*), for its part, may not be unaffected by this amalgam of obstacles, as it is directly dependent on the task and cognitive overload.

4 Discussion

The present study has consisted of a first part directed at using descriptive and multivariate tools to explore the viability of modeling a construct that takes into account pupils' perception of PS in the Physics classroom. For some authors (Castro & Lizasoain, 2012), when such an approach is linked to SEM, it can be a powerful tool at the service of educational research, but it has to rely on a solid theoretical base.

In the second part, a model was presented, in which a single first-order factor accounts for the closely correlated set of observed variables. Kessels et al. (2006) found that the explicit attitudes of students about Physics in general, were predicted by three factors: Difficulty, Masculinity, and Heteronomy. However, we note that our study is related to only topic and PS, assuming that student perceptions may be dynamic.

Classical studies of the topic in the scientific literature (Schoenfeld, 1982) have shown that experts perceive the "deep structure" of problems, and implement much better metacognitive strategies than novices that prevent them from wasting their PS resources. As some meta-analyses have shown (Peltier & Vannest, 2017), these skills can be improved through instruction. In the present study, it is therefore interesting to find that this factor intertwines variables concerning knowledge and

Table 6 Descriptive statistics, Mann–Whitney U test, statistical power, and effect size of the differences in the female and male pupils' perceptions

Variables	Female students (N=243)			Male students (N=285)			Mann–Whitney U test (*$p<0.05$)			Cohen's d (effect size)
	M	S.D	Md	M	S.D	Md	Z	p	Statistical power	
Complication (Co)	2.76	1.200	3	2.48	1.265	2	−2.690	0.007*	66%	0.12 (small)
Application (Ap)	2.66	1.97	3	2.50	1.225	3	−1.147	0.140	-	-
Path (Pa)	2.58	1.234	2	2.45	1.217	2	−1.020	0.308	-	-
Calculations (Ca)	2.46	1.253	2	2.35	1.270	2	−1.010	0.313	-	-
Understanding (Un)	2.53	1.281	2	2.22	1.228	2	−2.596	0.009*	73%	0.15 (small)
Memory (Me)	2.54	1.228	2	2.21	1.338	2	−3.173	0.002*	77%	0.12 (small)
Classroom (Cl)	2.51	1.249	2	2.21	1.214	2	−2.608	0.009*	72%	0.11 (small)
Knowledge (Kn)	2.30	1.169	2	2.16	1.224	2	−1.665	0.096	-	-
Effort (Ef)	2.09	1.227	2	2.14	1.274	2	−0.234	0.815	-	-
Distrust (Di)	2.12	1.231	2	2.04	1.294	2	−1.078	0.281	-	-
Interest (In)	1.88	1.043	2	2.00	1.281	2	−0.166	0.868	-	-

its application, the way to solve problems, the understanding of the tasks involved, and how the classroom is perceived and its suitability for learning, together with the pupil's own confidence.

From the point of view of pedagogical potential, this study shows that PS must be approached from a holistic perspective in the classroom by teachers, since students perceive it that way. Different difficulties converge, intertwined in their perceptions, which imply obstacles of a different nature (see Annex 1), but which has the perception of the initial task (Complication—didactic obstacle) as something remarkable, and the Memory with an influence is yet to be determined.

On the one hand, the influence of the class environment on PS has been highlighted in other modeling studies, since it is a variable that indirectly influences other variables (Wüstenberg et al., 2016), and this is coherent with certain aspects of our model. On the other hand, learning opportunities have a strong dependence on the specific topic chosen (Hooke's law in the present case) which, according to some studies (Kang et al., 2016), must be firmly contextualized to provide such opportunities in PS contexts. As Aramendi et al. (2018) note, pupils' attitudes towards learning improve when the search for and management of information are encouraged, and inquiry processes linked to everyday life are developed.

A study carried out by Eseryel et al. (2014) revealed the solid link between motivation and the rest of the variables. According to Abeysekera and Dawson (2015), this motivation to commit oneself to working on PS is reinforced when one participates actively in the creation and dissemination of knowledge. This predictive power of motivation is evident even at early ages (Mercader et al., 2017) and, to a large extent, motivation impregnates personal and life satisfaction (López-Cassá et al., 2018).

Pupils perceive a close relationship between the role of memory and the mathematical calculations involved in PS. Some authors (Ramirez et al., 2016), in modeling mathematical anxiety and working memory capacity, have found that pupils with greater cognitive capacity avoid using advanced PS strategies when they are very anxious about mathematics, and as a result perform more poorly. It is still a matter of debate how this could be improved through education (Cowan, 2014). In any case, there is agreement that specific activities focused on metacognition and working memory can contribute to improving arithmetic performance in PS (Cornoldi et al., 2015).

An enormous number of variables may be involved in PS since, from the solver's perspective, there may be many types of obstacles in the way of finding the path to a solution—personal, psychological, didactic, and epistemological. For instance, in a study deriving from PISA 2012 with 85,714 participants, Philpot et al. (2017) found that reasoning skills were essential to solve complex problems. We have opted for a more pupil-centred vision, but it seems that certain findings are shared, as is reflected in the correlations between variables that are hidden in the model. Even in apparent basic phenomena, such as Hooke's law, there are simultaneous complex causal models which, according to Perkins and Grotzer (2005), are related to four

categories—mechanism, interaction pattern, probability, and agency. Recently, problem-solving experiments in Physics in collaborative groups of students, show that students design, share, rethink, and evaluate their thought experiments, highlighting the importance of conceptual understanding, past–daily experience, logical reasoning, and conceptual–logical inference (Bancong & Song, 2020). This could explain the close relationship we found between the variables of the model and the error covariances: Memory (*Me*) as an obstacle binder, which would imply the redesign of problems (Cook, 2006), and the improvement of understanding (*Un*) through shared learning (*Co*), to facilitate the learning of mental models of Hooke's law and its application (*Ap*), in our case to Elasticity (Batlolona et al., 2020).Therefore, we sincerely believe that this pupil-centred view can complement other perspectives.

As we have seen, the study of students' perceptions of PS and its difficulties are not common in the field of Physics, limiting itself to finding links between memory, calculations, and anxiety. Our study delves into these difficulties from the point of view of the obstacles (Bachelard, 1983), contributing that these perceptions of the difficulties do so holistically and that attention should be paid to the underlying didactic obstacles. This is important, because very recently, Dávila-Acedo et al. (2021) have found that in young students, the study of Forces, among other topics, mostly experience negative emotions, possibly because solving Physics problems causes boredom, nervousness, and worry. Furthermore, we know that negative emotions are a good predictor of self-efficacy (Akin & Kurbanoglu, 2011). Based on these results, more studies on student perceptions are needed.

There are more than 11 variables in the student's perceptions of difficulties in SP we are aware of. A recurring variable that appears in the examined literature is self-efficacy, and that, due to complexity, as we will see at the end, deserves another separate construct. Our approach is quite unusual in the literature and is tentative in nature. In addition, these difficulties arise from classroom practice via teachers, at a critical age when students are introduced to PS and which will determine, to a large extent, their liking or rejecting physics.

After these considerations, an important objection persists; if students' perceptions are changing, can we really develop a static, quantitate model of them? At present, we have concluded another investigation on PS, with students of the same age range, but on another topic (Ponderal Laws and Stoichiometry), within Chemistry. It will be interesting to see how much dynamic there is in the model and what remains static, that is, how the content and topic influence the students' perceptions of PS.

Finally, with regard to gender as a discriminatory variable, although we found differences in some of the endogenous variables, those differences only had a small effect size. It is interesting to note, however, that the underlying obstacles were preponderantly didactic in nature. This finding is compatible with studies that have looked at the importance of self-efficacy. In one (Marshman et al., 2018), it was found that female pupils' lower self-efficacy compared with similarly performing male pupils can result in detrimental short-term and long-term impacts. In another (Short-Meyerson et al., 2016), there was found to be an influence of the parent–child interactions, with more encouragement being given to boys than to girls. There is clearly a need for further research in this area.

5 Limitations

The main limitation of the study is that the modeling is formulated from a confirmatory perspective when there is no robust supporting theoretical basis. Nevertheless, to the best of the authors' knowledge, there have as yet been no holistic approaches to studying secondary education pupils' PS difficulties in science. Hence, we sincerely believe that the proposal to consider all the variables of the study interrelated in a single factor, with covariances between hidden factors (errors) as constraints, constitutes a plausible starting point for future developments. It should help to be able to start with a set of strongly related variables from which to delve into the model's internal structure in search of latent structures supported by theory. We point out that some of these restraints imply, with a fine-grained vision, a role relevant to Memory (cognitive overload–psychological obstacle) and an opportunity for further research.

There are currently calls for the reliability of statistical studies in education to be improved. This is one of the purposes of using the statistical power or probability $(1 - \beta)$ of not committing type II errors (Aron et al., 2013; Kraemer & Blasey, 2015). For example, in the next model, we computed Model C; the RMSEA limits were *LO 90* = 0.000 and *HI 90* = 0.044, and, considering a value of α of 0.05, our study sample size ($N = 528$), and the degrees of freedom ($df = 38$), we get 92.51% for the power—a fairly positive value (Preacher & Coffman, 2006).

The imposition of constraints on the variances and covariances can affect errors and residues, and is undesirable as it distorts the model (Hoyle, 2011). They can be accepted, however, especially in emerging studies, as long as they are theoretically justified. On the other hand, we believe that the choice of a 5-point Likert scale was positive, even though it meant that the concomitant problems of normality required that ADF techniques be used, and the samples that are necessary for them are somewhat complicated to obtain in educational settings that demand strong and sustained commitment over the years on the part of the teachers involved who work with a particular set of methods.

Another limitation is the somewhat low value of the *AVE* (Table 5), as against the high value of *CRe*. This disadvantage may come from the fact that *AVE* only explains all the variables as a group. West et al. (2012) argue that a model must have the results of research as its basis, capturing the features of interest and having a substantive theoretical underpinning (Montesinos & Backhoff, 2010). As required by Kline (2012), the assumptions made appear to be plausible, and the findings we present appear to be within what the scientific literature considers acceptable in PS. About the near future, our intention is to use this PS model, together with a construct both already completed (Vázquez-Bernal & Jiménez-Pérez, 2016) and others still under development (Emotions, Self-regulation / Self-efficacy,…), to predict pupil performance in PS, whose data are currently available.

Annex 1

See Table 7.

Table 7 Summary table of the variables studied and description, nature, and origin of the obstacles

Variables	Brief description	Nature of difficulty	Origin of obstacle	Theoretical support
Application (*Ap*)	Application of theory	Nonsignificant learning (includes alternative pupil conceptions)	Epistemological	This is related to the application of theoretical knowledge to the demands of problem solving, i.e., to transfer this knowledge to the task set. Several obstacles converge at this point: the pupils' alternative conceptions, superficial acquisition of school knowledge, framed in a significant lack of learning (Pozo & Gómez, 1998). Appropriate support scaffolding can help pupils correctly solve problems that involve strong alternative conceptions (Lin & Singh, 2015)
Understanding (*Un*)	Incomprehension of the statement	Writing down and comprehension of the instructions	Didactic	When pupils are inexperienced and the information is complex, it is suggested (Haslam & Hamilton, 2010) that physically integrating mutually referring sources of information reduces cognitive load (texts and integrated illustrations), and therefore makes practical work instructions easier to understand. Some authors stress the self-generated explanations (from talk-aloud protocols) that pupils produce while studying worked-out examples of mechanics problems, and their subsequent reliance on examples during problem solving (Chi et al., 1989). Complementarily, it has been found that terminology in common scientific areas, or non-technical terms, are related to (a) general non-technical words in science school settings; (b) meta-representational verbs; (c) linking items; (d) teleological or intentional terminology (Quílez, 2019)

Table 7 (continued)

Variables	Brief description	Nature of difficulty	Origin of obstacle	Theoretical support
Knowledge (*Kn*)	Lack of knowledge	The content's innate complexity	Epistemological	Some results suggest that reminding pupils to follow typical problem-solving frameworks limits their views of what counts as good problem solving (Kuo et al., 2017). Another study clearly suggests that, in teaching the concept of force, the systematic use of interaction diagrams significantly fostered pupils' understanding even with teachers who have no intensive tutoring or research background (Savinainen et al., 2017). The emphasis of Physics teaching research should no longer be on the compilation of a list of misconceptions that have to be remedied or replaced, but on the conceptual connections pupils make and their associative reasoning patterns. It is still up to the teacher to use the complex systems perspective as a framework to facilitate pupils' conceptual development and understanding, building on their existing knowledge systems (Lemmer et al., 2018). Pupils use several epistemological resources to decide on the accuracy of their data, including accuracy via following the right procedure and accuracy via what their classmates find. Traditional, formulation-based Physics teaching might have led pupils to activate naive epistemological resources that prevent them participating in the practice of science in more meaningful ways (Alpaslan et al., 2017)
Memory (*Me*)	Lack of memory	Cognitive overload	Psychological	Domain-general cognitive knowledge has frequently been used to explain skill, but domain-specific knowledge held in long-term memory may provide a better explanation, and if the importance of domain-specific knowledge is accepted then instructional design theories and processes will be transformed (Tricot & Sweller, 2014). Cognitive-load research has shown that learning from worked-out examples, rather than problem solving, is very effective during the initial stages of cognitive skill acquisition. In later stages, however, problem-solving is superior. In this regard, some authors propose a fading procedure in which problem-solving elements are gradually integrated into study examples (Renkl & Atkinson, 2003)

Table 7 (continued)

Variables	Brief description	Nature of difficulty	Origin of obstacle	Theoretical support
Effort (*Ef*)	Lack of work	Absence of expectations of success	Personal (epistemo-logical component)	The perceptions of scientists as hard-working individuals who have struggled to make scientific progress help increase pupils' interest in science, their recall of key science concepts, and their abilities to solve complex problems (Hong & Lin-Siegler, 2012). Of the school-related factors influencing experiences of school science, the cur-ricular content was the most important, followed by motivation and interest in the subject. There is evidence that interest and motivation in science depend on the teachers' practice and on the perception of science as a difficult subject (Shirazi, 2017)
Interest (*In*)	Lack of interest	Intrinsic demotivation	Personal (didactic component)	Some results underscore the key role that practice communities play in career and identity development, and suggest a need for interventions to help socializers better understand the value and purpose of science literacy themselves so as to encourage pupils to appreciate science, be aware of possible career options in science, and enjoy learning and doing science (Aschbacher et al., 2010). Pupils ranked enjoy-ment, interest, and skill in a subject and its perceived need in their future study or career plans as the most important factors in both choosing and rejecting subjects (Palmer et al., 2017). According to previous theories and studies, interest and self-concept at an early age have played a pivotal role in career choices, especially in the fields of science, technology, engineering, and mathematics. However, there were no reciprocal relationships between interest and self-concept in predicting pupils' science aspirations (Kang et al., 2019)

Table 7 (continued)

Variables	Brief description	Nature of difficulty	Origin of obstacle	Theoretical support
Complication (*Co*)	Initial perception of task	Misinterpretation of expectations	Didactic	The difficulty of problem-solving is also attributable, among others, to external factors such as abstraction. Abstraction refers to the representation of the content and context of a problem that either facilitates or impedes analogical transfer of one problem to another. Most Physics problems are more abstract than everyday problems which are embedded in various contexts (Jonassen & Hung, 2015). Pupils frequently misconceive the process of problem-solving, expecting a linear process to be required to solve an exercise, rather than the convoluted search process needed to solve a genuine problem (Yerushalmi & Magen, 2006)
Classroom (*Cl*)	Incomprehension in class	School habits	Didactic	Science teachers can always benefit from efficient tools that help pupils engage with the subject and understand it better without significantly adding to the teacher's workload or requiring too much class time to manage. Reflective writing, being low impact and high return, is such a tool (El-Helou & Kalman, 2018)
Path (*Pa*)	No known path	Adopted processes	Psychological	Pupils who had been exposed to a structured problem-solving strategy were found to show better conceptual understanding of Physics and to tend to adopt a conceptual approach to problem-solving (Gaigher et al., 2007). The explicit strategy improves the quality and completeness of pupils' Physics representations more than traditional textbook-based strategies (Huffman, 1997)
Distrust (*Di*)	Self-reliance	Fear of failure	Personal level (psychological component)	In self-directed success, if pupils have had insufficient practice to convince themselves of their new effectiveness, they apply the skills they have been taught weakly and inconsistently, and quickly stop using those skills when they experience difficulties (Bandura, 1988). Problem-solving strategies, including pupils' own confidence, influence their success more than the number of problems they had solved (Byun & Lee, 2014). Pupils' use of problem-solving strategy steps contributes to their critical and analytical skills, with trust playing a relevant role (Gok, 2014)

Table 7 (continued)

Variables	Brief description	Nature of difficulty	Origin of obstacle	Theoretical support
Calculations (*Ca*)	Failure in the calcula-tions	Intellectual operations needed in the transfer from mathematics	Epistemological and psychological	Learning and retention are concepts subordinate to that of transfer with emphasis on the individual's prior knowledge, skills, attitudes, and other characteristics (Voss, 1987). Mathematical complexity is defined operationally by the number and type of equations to be manipulated concurrently due to the number of unknowns in each equation. This complexity negatively affects pupils' performance in mathematics (Ibrahim et al., 2017)

Annex 2

See Table 8.

Table 8 Problem[*] about Hooke's law included in the numerical test. *A spring is stretched by applying a force according to the data collected in the following table:*

F (N)	0 N	1.5 N	3 N	4.5 N	6 N	7.5 N
L (cm)	0 cm	1 cm	2 cm	3 cm	4 cm	5 cm

a) What is the spring's elastic constant, and what is its unit? b) What does this number mean? c) Do you know of any kind of relationship, law, or principle that explains this phenomenon? d) Imagine you were given a spring, for example from a biro. Describe what steps you would take to calculate its K? e) Why do you think springs are fitted to the body of most motor vehicles?

[*]All other problems are variations of this particular one

Declarations

Ethics Approval The studio has implemented all necessary ethical approvals.

Consent to Participate The authors consent to participate.

Consent for Publication The authors consent to the publication.

Conflicts of Interest The authors declare that they have no conflict of interest.

References

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *Higher Education Research & Development, 34*(1), 1–14.

Akben, N. (2018). Effects of the problem-posing approach on students' problem solving skills and meta-cognitive awareness in science education. *Res Sci Educ, 50*(3), 1143–1165. https://doi.org/10.1007/s11165-018-9726-7

Akın, A. (2010). Achievement goals and academic locus of control: Structural equation modeling. *Eurasian Journal of Educational Research (EJER), 38*, 1–18.

Akin, A., & Kurbanoglu, I. N. (2011). The relationships between math anxiety, math attitudes, and self-efficacy: A structural equation model. *Studia Psychologica, 53*(3), 263–273.

Aldridge, J., Afari, E., & Fraser, B. (2012). Influence of teacher support and personal relevance on academic self-efficacy and enjoyment of mathematics lessons: A structural equation modeling approach. *The Alberta Journal of Educational Research*, *58*(4), 614–633. Retrieved 25 August from http://hdl.handle.net/20.500.11937/47820

Alpaslan, M. M., Yalvac, B., & Loving, C. (2017). High school Physics students' personal epistemologies and school science practice. *Science & Education, 26*(7–9), 841–865. https://doi.org/10.1007/s11191-017-9930-2

Aramendi, P., Arburua, R. M., & Buján, K. (2018). El aprendizaje basado en la indagación en la enseñanza secundaria [Inquiry-based learning in secondary education]. *Revista de Investigación Educativa, 36*(1), 109–124. https://doi.org/10.6018/rie.36.1.278991

Arán, V., & Richaud, M. C. (2017). A structural equation modeling of executive functions, IQ and mathematical skills in primary students: Differential effects on number production, mental calculus and arithmetical problems. *Child Neuropsychology, 23*(7), 864–888. https://doi.org/10.1080/09297049.2016.1199665

Arbuckle, J. L. (2007). *Amos 16 user's guide*. SPSS.

Arbuckle, J. L. (2011). *IBM SPSS AMOS 20 user's guide*. IBM Corporation.

Aron, A., Coups, E. J., & Aron, E. N. (2013). *Statistics for psychology* (6th ed.). Pearson.

Aschbacher, P. R., Li, E., & Roth, E. J. (2010). Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine. *Journal of Research in Science Teaching, 47*(5), 564–582. https://doi.org/10.1002/tea.20353

Astolfi, J. P. (1999). *El error, un medio para* enseñar [Error, a method of teaching]. Sevilla: Díada.

Austin, E. J., Evans, P., Goldwater, R., & Potter, V. (2005). A preliminary study of emotional intelligence, empathy and exam performance in first year medical students. *Personality and Individual Differences, 39*(8), 1395–1405. https://doi.org/10.1016/j.paid.2005.04.014

Aydin, U. (2007). A structural equation modeling study: The metacognition-knowledge model for geometry. *Unpublished Master Thesis. Middle East Technical University, Ankara*. Retrieved 7 July from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.633.9073&rep=rep1&type=pdf

Bachelard G. (1983). *La Formation de l'esprit Scientifique* [The Formation of the Scientific Spirit]. Paris, France: J. Vrin.

Bancong, H., & Song, J. (2020). Exploring how students construct collaborative thought experiments during physics problem-solving activities. *Science & Education, 29*(3), 617–645. https://doi.org/10.1007/s11191-020-00129-3

Bandura, A. (1988). Organisational applications of social cognitive theory. *Australian Journal of Management, 13*(2), 275–302.

Batlolona, J. R., Diantoro, M., & Leasa, M. (2020). Students' mental models of solid elasticity: Mixed method study. *Journal of Turkish Science Education*, *17*(2), 200–210. https://doi.org/10.36681/tused.2020.21

Bevins, S., & Price, G. (2016). Reconceptualising inquiry in science education. *International Journal of Science Education, 38*(1), 17–29.

Bollen, K. A. (1998). *Structural equation models*. John Wiley & Sons Ltd.

Brandriet, A. R., Ward, R. M., & Bretz, S. L. (2013). Modeling meaningful learning in chemistry using structural equation modeling. *Chemistry Education Research and Practice, 14*(4), 421–430. https://doi.org/10.1039/C3RP00043E

Byun, T., & Lee, G. (2014). Why students still can't solve physics problems after solving over 2000 problems. *American Journal of Physics, 82*(9), 906–913. https://doi.org/10.1088/0031-9120/29/6/006

Castro, M., & Lizasoain, L. (2012). Las técnicas de modelización estadística en la investigación educativa: Minería de datos, modelos de ecuaciones estructurales y modelos jerárquicos lineales [Statistical

modeling techniques in educational research: Data mining, structural equation models and hierarchical linear models]. *Revista Española De Pedagogía, 70*(251), 131–148.

Celik, I., Sahin, I., & Akturk, A. O. (2014). Analysis of the relations among the components of technological pedagogical and content knowledge (TPACK): A structural equation model. *Journal of educational computing research*, *51*(1), 1–22. https://doi.org/10.219/FEC.51.1.a

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: how students study and use examples in learning to solve problems. *Cognitive science, 13*(2), 145–182. https://doi.org/10.1016/0364-0213(89)90002-5

Choi, J., Kim, S., Chen, J., & Dannels, S. (2011). A comparison of maximum likelihood and Bayesian estimation for polychoric correlation using Monte Carlo simulation. *Journal of Educational and Behavioral Statistics*, *36*(4), 523–549. https://doi.org/10.3102/2F1076998610381398

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education, 90*(6), 1073–1091. https://doi.org/10.1002/sce.20164

Cornoldi, C., Carretti, B., Drusi, S., & Tencati, C. (2015). Improving problem solving in primary school students: The effect of a training programme focusing on metacognition and working memory. *British Journal of Educational Psychology, 85*(3), 424–439.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review, 26*(2), 197–223.

Cupani, M. (2012). Análisis de ecuaciones estructurales: Conceptos, etapas de desarrollo y un ejemplo de aplicación [Structural equation analysis: Concepts, development stages and an example of application]. *Revista Tesis, 2*(1), 186–199.

Dávila-Acedo, M. A., Cañada, F., Sánchez-Martín, J., Airado-Rodríguez, D., & Mellado, V. (2021). Emotional performance on physics and chemistry learning: The case of Spanish K-9 and K-10 students. *International Journal of Science Education, 43*(6), 823–843. https://doi.org/10.1080/09500693.2021.1889069

De Carlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods, 2*(3), 292–307.

Deacon, C., & Hajek, A. (2011). Student perceptions of the value of physics laboratories. *International Journal of Science Education, 33*(7), 943–977. https://doi.org/10.1080/09500693.2010.481682

Dunkley, D. M., & Blankstein, K. R. (2000). Self-critical perfectionism, coping, hassles, and current distress: A structural equation modeling approach. *Cognitive Therapy and Research, 24*(6), 713–730.

El-Helou, J., & Kalman, C. S. (2018). Reflective writing for a better understanding of scientific concepts in high school. *The Physics Teacher, 56*(2), 88–91. https://doi.org/10.1119/1.18279

Eseryel, D., Law, V., Ifenthaler, D., Ge, X., & Miller, R. (2014). An investigation of the interrelationships between motivation, engagement, and complex problem solving in game-based learning. *Journal of Educational Technology & Society, 17*(1), 42–53.

European Commission (2007).*Science education NOW: A renewed pedagogy for the future of Europe*. Brussels: Office for Official Publications of the European Communities.

Fonseca, J. M. & Conboy, J. E. (2006). Secondary student perceptions of factors effecting failure in science in Portugal. *Eurasia Journal of Mathematics, Science and Technology Education*, *2*(2), 82–95. https://doi.org/10.12973/ejmste/75455

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50.

Gaigher, E., Rogan, J. M., & Braun, M. W. H. (2007). Exploring the development of conceptual understanding through structured problem-solving in Physics. *International Journal of Science Education, 29*(9), 1089–1110. https://doi.org/10.1080/09500690600930972

Gok, T. (2014). Students' achievement, skill and confidence in using stepwise problem-solving strategies. Eurasia Journal of Mathematics, *Science & Technology Education*, *10*(6), 617–624. https://doi.org/10.12973/eurasia.2014.1223a

Hardy, G. (2014). Academic self-concept: Modeling and measuring for science. *Research in Science Education*, *44*(4), 549–579. https://doi.org/10.1007/s11165-013-9393-7

Haslam, C. Y., & Hamilton, R. J. (2010). Investigating the use of integrated instructions to reduce the cognitive load associated with doing practical work in secondary school science. *International Journal of Science Education, 32*(13), 1715–1737. https://doi.org/10.1080/09500690903183741

Haus, I., Steinmetz, H., Isidor, R., & Kabst, R. (2013). Gender effects on entrepreneurial intention: A meta-analytical structural equation model. *International Journal of Gender and Entrepreneurship*, *5*(2), 130–156. https://doi.org/10.1108/17566261311328828

Hong, H. Y., & Lin-Siegler, X. (2012). How learning about scientists' struggles influences students' interest and learning in physics. *Journal of Educational Psychology, 104*(2), 469–484. https://doi.org/10.1037/a0026224

Hoyle, R. H. (2011). *Structural equation modeling for social and personality psychology*. SAGE Publications Ltd.

Huffman, D. (1997). Effect of explicit problem solving instruction on high school students' problem-solving performance and conceptual understanding of physics. *Journal of Research in Science Teaching, 34*(6), 551–570. https://doi.org/10.1002/(SICI)1098-2736(199708)34:6/3C551::AID-TEA2/3E3.0.CO;2-M

Ibrahim, B., Ding, L., Heckler, A. F., White, D. R., & Badeau, R. (2017). How students process equations in solving quantitative synthesis problems? Role of mathematical complexity in students' mathematical performance. *Physical Review Physics Education Research, 13*(2), 020120. https://doi.org/10.1103/PhysRevPhysEducRes.13.020120

Irving, P. W., & Sayre, E. C. (2015). Becoming a physicist: The roles of research, mindsets, and milestones in upper-division student perceptions. *Physical Review Special Topics-Physics Education Research, 11*(2), 020120. https://doi.org/10.1103/PhysRevSTPER.11.020120

Jonassen, D. H. (2007). What makes scientific problems difficult? In D. H. Jonassen (Ed.), *Learning to Solve Complex Scientific Problems* (pp. 3–24). Routledge.

Jonassen, D. H., & Hung, W. (2015). All problems are not equal: Implications for problem-based learning. In A. Walker, H. Leary, C. E. Hmelo-Silver, & P. A. Ertmer (Eds.), *Essential Readings in Problem-Based Learning* (pp. 7–41). Purdue University Press.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software International.

Kamalimoghaddam, H., Tarmizi, R. A., Ayub, A. F. M., & Jaafar, W. M. W. (2016). Confirmatory model of mathematics self-efficacy, problem solving skills and prior knowledge on mathematics achievement: A structural equation model. *Malaysian Journal of Mathematical Sciences, 10*(2), 187–200. Retrieved 1 June from http://einspem.upm.edu.my/journal/fullpaper/vol10saugust/14.%20Hajar.pdf

Kang, H., Windschitl, M., Stroupe, D., & Thompson, J. (2016). Designing, launching, and implementing high quality learning opportunities for students that advance scientific thinking. *Journal of Research in Science Teaching, 53*(9), 1316–1340.

Kang, J., Keinonen, T. & Salonen, A. (2019). Role of interest and self-concept in predicting science aspirations: gender study. *Res Sci Educ*. https://doi.org/10.1007/s11165-019-09905-w

Kessels, U., Rau, M., & Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology, 76*(4), 761–780. https://doi.org/10.1348/000709905X59961

Kline, R. B. (2012). Assumptions in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. 111–125). Guilford Publications.

Kraemer, H. C., & Blasey, C. (2015). *How many subjects? Statistical power analysis in research*. Sage Publications.

Kuo, E., Hallinen, N. R., & Conlin, L. D. (2017). When procedures discourage insight: Epistemological consequences of prompting novice physics students to construct force diagrams. *International Journal of Science Education, 39*(7), 814–839. https://doi.org/10.1080/09500693.2017.1308037

Kuo, F. R., & Hwang, G. J. (2015). A structural equation model to analyse the antecedents to students' web-based problem-solving performance. *Australasian Journal of Educational Technology*, *31*(4). https://doi.org/10.14742/ajet.284

Latorre, A., Rincón, D., & del Arnal, J. (1996). *Bases metodológicas de la investigación educativa [Methodological bases of educational research]*. Hurtado.

Lee, M. H., Johanson, R. E., & Tsai, C. C. (2008). Exploring Taiwanese high school students' conceptions of and approaches to learning science through a structural equation modeling analysis. *Science Education, 92*(2), 191–220. https://doi.org/10.1002/sce.20245

Lemmer, M., Kriek, J., & Erasmus, B. (2018). Analysis of students' conceptions of basic magnetism from a complex systems perspective. *Research in Science Education*, 1–18. https://doi.org/10.1007/s11165-018-9693-z

Lin, S. Y., & Singh, C. (2015). Effect of scaffolding on helping introductory physics students solve quantitative problems involving strong alternative conceptions. *Physical Review Special Topics-Physics Education Research, 11*(2), 020105. https://doi.org/10.1103/PhysRevSTPER.11.020105

López-Cassà, E., Pérez-Escoda, N., & Alegre, A. (2018). Competencia emocional, satisfacciónen contextos específicos y satisfacción con la vida en la adolescencia [Emotional competence, satisfaction in specific contexts and satisfaction with life in adolescence]. *Revista de Investigación Educativa/Journal of EducationalResearch, 36*(1), 57–73. https://doi.org/10.6018/rie.36.1.273131

Marshman, E. M., Kalender, Z. Y., Nokes-Malach, T., Schunn, C., & Singh, C. (2018). Female students with A's have similar physics self-efficacy as male students with C's in introductory courses: A cause for

alarm? *Physical Review Physics Education Research, 14*(2), 020123. https://doi.org/10.1103/PhysRevPhysEducRes.14.020123

Matas, A. (2018). Diseño del formato de escalas tipo Likert: Un estado de la cuestión [Designing the Likert scale format: A state of the art]. *Revista Electronica De Investigacion Educativa, 20*(1), 38–47.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates Inc.

Mercader, J., Presentación, M.-J., Siegenthaler, R., Molinero, V., & Miranda, A. (2017). Motivación y rendimiento académico en matemáticas: Un estudio longitudinal en las primeras etapas educativas [Motivation and academic performance in mathematics: A longitudinal study in the early stages of education]. *Revista De Psicodidáctica, 22*(2), 157–163.

Montesinos, M.-J., & Backhoff, E. (2010). Validación de un cuestionario de contexto para evaluar sistemas educativos con Modelos de ecuaciones estructurales [Validation of a context questionnaire to evaluate education systems with Structural Equation Models]. *Relieve*, *1*(2), 1–17. Retrieved April 2020 from http://www.uv.es/RELIEVE/v16n2/RELIEVEv16n2_1.htm

Muthén, B., & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. *British Journal of Mathematical and Statistical Psychology, 45*, 19–30. https://doi.org/10.1111/j.2044-8317.1992.tb00975.x

Normandeau, S., & Guay, F. (1998). Preschool behavior and first-grade school achievement: The mediational role of cognitive self-control. *Journal of Educational Psychology, 90*(1), 111–121. https://doi.org/10.1037/0022-0663.90.1.111

Nugent, G., Barker, B., Welch, G., Grandgenett, N., Wu, C., & Nelson, C. (2015). A model of factors contributing to STEM learning and career orientation. *International Journal of Science Education*, (ahead-of-print), 1–22. https://doi.org/10.1080/09500693.2015.1017863

OECD (2017). *PISA 2015. La resolución colaborativa de problemas. Resumen ejecutivo* [Collaborative problem solving. Executive Summary]. Madrid: Instituto Nacional de Evaluación Educativa.

Oñorbe, A. & Sánchez, J. M. (1996). Dificultades en la enseñanza-aprendizaje de los problemas de Física y Química. I. Opiniones de los alumnus [Difficulties in teaching-learning the problems of Physics and Chemistry I. Students' opinions. Teaching of Science], *Enseñanza de las Ciencias 14*(2), 165–170.

Östman, L., & Wickman, P. O. (2014). A pragmatic approach on epistemology, teaching, and learning. *Science Education, 98*(3), 375–382. https://doi.org/10.1002/sce.21105

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary educational psychology*, *21*(4), 325–344. https://doi.org/10.1006/ceps.1996.0025

Palmer, T. A., Burke, P. F., & Aubusson, P. (2017). Why school students choose and reject science: a study of the factors that students consider when selecting subjects. *International Journal of Science Education, 39*(6), 645–662. https://doi.org/10.1080/09500693.2017.1299949

Park, S. Y., Nam, M. W., & Cha, S. B. (2012). University students' behavioral intention to use mobile learning: Evaluating the technology acceptance model. *British Journal of Educational Technology, 43*(4), 592–605. https://doi.org/10.1111/j.1467-8535.2011.01229.x

Peltier, C., & Vannest, K. J. (2017). A meta-analysis of schema instruction on the problem-solving performance of elementary school students. *Review of Educational Research, 87*(5), 899–920. https://doi.org/10.3102/2F0034654317720163

Perkins, D. N., & Grotzer, T. A. (2005). Dimensions of causal understanding: The role of complex causal models in students' understanding of science. *Studies in Science Education, 41*(1), 117–166. https://doi.org/10.1080/03057260508560216

Philpot, R., Ramalingam , D., Dossey , J. A. & McCrae, B. (2017). Factors that influence the difficulty of problem-solving items. In B. Csapó & J. Funke (Eds.) *The Nature of Problem Solving: Using Research to Inspire 21st Century Learning* (141–158). OECD Publishing: Paris. https://doi.org/10.1787/9789264273955-en

Pittalis, M., Christou, C., Mousoulides, N., & Pitta-Pantazi, D. (2004). A structural model for problem posing. In *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 49–56).

Pozo, J. I., & Gómez Crespo, M. A. (1998). *Aprender y enseñar ciencias [Learning and Teaching Science]*. Morata.

Preacher, K. J., & Coffman, D. L. (2006, May). *Computing power and minimum sample size for RMSEA* [Computer software]. Retrieved 15 March 2020 from http://quantpsy.org/rmsea/rmsea.htm

Prosser, M., Walker, P., & Millar, R. (1996). Differences in students' perceptions of learning physics. *Physics Education, 31*(1), 43.

Pulgar, J., Candia, C., & Leonardi, P. M. (2020). Social networks and academic performance in physics: Undergraduate cooperation enhances ill-structured problem elaboration and inhibits

well-structured problem solving. *Physical Review Physics Education Research, 16*(1), 010137. https://doi.org/10.1103/PhysRevPhysEducRes.16.010137

Quílez, J. (2019). A categorisation of the terminological sources of student difficulties when learning chemistry. *Studies in Science Education*. https://doi.org/10.1080/03057267.2019.1694792

Ramirez, G., Chang, H., Maloney, E. A., Levine, S. C., & Beilock, S. L. (2016). On the relationship between math anxiety and math achievement in early elementary school: The role of problem solving strategies. *Journal of Experimental Child Psychology, 141*, 83–100. https://doi.org/10.1016/j.jecp.2015.07.014

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist, 38*(1), 15–22. https://doi.org/10.1207/S15326985EP3801_3

Rindermann, H., & Neubauer, A. C. (2004). Processing speed, intelligence, creativity, and school performance: Testing of causal hypotheses using structural equation models. *Intelligence, 32*(6), 573–589. https://doi.org/10.1016/j.intell.2004.06.005

Rittle-Johnson, B., Matthews, P. G., Taylor, R. S., & McEldoon, K. L. (2011). Assessing knowledge of mathematical equivalence: A construct-modeling approach. *Journal of Educational Psychology, 103*(1), 85–104. https://doi.org/10.1037/a0021334

Savinainen, A., Mäkynen, A., Nieminen, P., & Viiri, J. (2017). The effect of using a visual representation tool in a teaching-learning sequence for teaching Newton's third law. *Research in Science Education, 47*(1), 119–135. https://doi.org/10.1007/s11165-015-9492-8

Schleicher, A. (2019). *PISA 2018: Insights and interpretations*. OECD Publishing.

Schoenfeld, A. H. (1982). Expert and novice mathematical problem solving. *Journal for Research in Mathematics Education, 13*(1), 31–49.

Shirazi, S. (2017). Student experience of school science. *International Journal of Science Education, 39*(14), 1891–1912. https://doi.org/10.1080/09500693.2017.1356943

Short-Meyerson, K., Sandrin, S., & Edwards, C. (2016). Gender influences on parent-child science problem-solving behaviors. *Journal of Research in Childhood Education, 30*(3), 334–348. https://doi.org/10.1080/02568543.2016.1178194

Silva, O. H. M., Laburú, C. E., Camargo, S., & Christófalo, A. A. C. (2019). Epistemological contributions derived from an investigative method in an experimental class in the study of Hooke's law. *Acta Scientiae*, *21*(2), 110–127. https://doi.org/10.17648/acta.scientiae.v21iss2id4695

Soltani, A. (2018). Influence of motivating science class, family, and peer models on students' approaches to learning science: A structural equation modeling analysis. *Res Sci Educ*. https://doi.org/10.1007/s11165-018-9748-1

Stamovlasis, D., Tsitsipis, G., & Papageorgiou, G. (2012). Structural equation modeling in assessing students' understanding of the state changes of matter. *Chemistry Education Research and Practice, 13*(3), 357–368. https://doi.org/10.1039/C2RP20031G

Sungur, S. (2007). Modeling the relationships among students' motivational beliefs, metacognitive strategy use, and effort regulation. *Scandinavian Journal of Educational Research, 51*(3), 315–326. https://doi.org/10.1080/00313830701356166

Tarhini, A., Mohammed, A. B., & Maqableh, M. (2016). Modeling factors affecting student's usage behaviour of e-learning systems in Lebanon. *International Journal of Business and Management, 11*(2), 299–312. https://doi.org/10.5539/ijbm.v11n2p299

Thakkar, J. J. (2020). *Structural equation modelling*. Springer.

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods, 16*, 209–220. https://doi.org/10.1037/a0023353

Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational psychology review, 26*(2), 265–283. https://doi.org/10.1007/s10648-013-9243-1

Turpen, C., & Finkelstein, N. D. (2010). The construction of different classroom norms during peer instruction: Students perceive differences. *Physical Review Special Topics-Physics Education Research, 6*(2), 020123. https://doi.org/10.1103/PhysRevSTPER.6.020123

Vázquez-Bernal, B. & Jiménez-Pérez, R. (2016). Modeling of a theoretical construct on the perception of students in inquiry processes in science. *Revista de Psicodidactica/ Journal of Psychodidactics*, *21*(1), 25-44. https://doi.org/10.1387/RevPsicodidact.13000.

Velayutham, S., & Aldridge, J. M. (2013). Influence of psychosocial classroom environment on students' motivation and self-regulation in science learning: A structural equation modeling approach. *Res Sci Educ, 43*(2), 507–527. https://doi.org/10.1007/s11165-011-9273-y

Velayutham, S., Aldridge, J. M., & Fraser, B. (2012). Gender differences in student motivation and self-regulation in science learning: A multi-group structural equation modeling analysis. *International*

journal of science and mathematics education, 10(6), 1347–1368. https://doi.org/10.1007/s10763-012-9339-y

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). Un viaje alrededor de alfa y omega para estimar la fiabilidad de consistencia interna [A trip around alpha and omega to estimate internal consistency reliability]. *Anales De Psicología/annals of Psychology, 33*(3), 755–782.

Voss, J. F. (1987). Learning and transfer in subject-matter learning: a problem-solving model. *International Journal of Educational Research, 11*(6), 607–622. https://doi.org/10.1016/0883-0355(87)90005-X

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.) *Handbook of structural equation modelling* (209–231). New York: Guilford.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, *92*(5), 941–967. https://doi.org/10.1002/sce.20259

Wüstenberg, S., Greiff, S., Vainikainen, M. P., & Murphy, K. (2016). Individual differences in students' complex problem solving skills: How they evolve and what they imply. *Journal of Educational Psychology*, *108*(7), 1028–1044. https://doi.org/10.1037/edu0000101

Yerushalmi, E., & Magen, E. (2006). Same old problem, new name? Alerting students to the nature of the problem-solving process. *Physics Education, 41*(2), 161. https://doi.org/10.1088/0031-9120/41/2/007

Yurt, E., & Sünbül, A. M. (2014). A structural equation model explaining 8th grade students' mathematics achievements. *Educational Sciences: Theory and Practice*, *14*(4), 1642–1652. https://doi.org/10.12738/estp.2014.4.2193