



# A diagnostic comparison of Spanish and Turkish middle school students' proportional reasoning

Sergio Martínez-Juste<sup>1</sup> , Muhammet Arıcan<sup>2</sup> ,  
José M. Muñoz-Escolano<sup>1</sup> , and  
Antonio M. Oller-Marcén<sup>3</sup> 

## Abstract

This small-scale comparative study investigated Spanish and Turkish middle school students' mastery of four core cognitive skills required for proportional reasoning using the log-linear cognitive diagnosis model. We compared students in the two samples for their proportional reasoning and provided a diagnostic assessment of their strengths and weaknesses. The study included 314 Spanish eighth grade and 282 Turkish seventh grade students. The students were administered a proportional reasoning test with 22 multiple-choice items. We found significant differences between the two samples in two of the cognitive skills. While the Spanish sample were better at understanding nonproportional relationships, the Turkish sample were better at understanding the concept of ratio. Furthermore, while the Spanish sample had a more dispersed distribution of students among different attribute profiles, the Turkish sample mostly consisted of either low-performing or high-performing students. To better understand the similarities and differences between the two samples, we compared students' performance on some specific test items. The study shows how cognitive diagnostic models, which are not widely used in small-scale comparative studies, can be very useful in revealing similarities and differences between two samples that cannot be detected by traditional psychometric models.

## Keywords

cognitive diagnostic models, comparative study, middle school students, proportional reasoning, ratio and proportion

Date received: 24 August 2022; accepted: 8 March 2023

---

<sup>1</sup>Departamento de Matemáticas, Universidad de Zaragoza, Zaragoza, Spain

<sup>2</sup>Department of Mathematics and Science Education, Boğaziçi University, Istanbul, Turkey

<sup>3</sup>Centro Universitario de la Defensa de Zaragoza, Zaragoza, Spain

## Corresponding Author:

Muhammet Arıcan, Department of Mathematics and Science Education, Boğaziçi University, Istanbul, Turkey.  
Email: muhammetarican@gmail.com

## I. Introduction

Manzon (2011) defines comparative education as “an interdisciplinary subfield of education studies that systematically examines the similarities and differences between educational systems in two or more national or cultural contexts” (p. 215). Postlethwaite (1988) provides four different purposes of comparative education with one of them being “describing similarities and differences in educational phenomena between systems of education and interpreting why they exist” (pp. xix–xx). Stigler et al. (2000) state that “cross-cultural comparison is a powerful way to unveil unnoticed but ubiquitous practices” (p. 88). Some researchers (e.g., Cai, 1995; Cai et al., 2016) also pointed out the importance of conducting comparative studies in the domain of mathematics education.

While some comparative studies use results from large-scale international studies such as Trends in International Mathematics and Science Study (TIMSS) (e.g., Şen & Arıcan, 2015), others use these results in order to compare instructional practices among different countries (e.g., Bokhove, 2022). Moreover, some studies can be narrower oriented, focusing only on particular topics or concepts (e.g., Chin & Lin, 2009). Proportional reasoning is one of the important skills in mathematics education. The use of proportional reasoning techniques for problem-solving is present (sometimes without any explanation) in many ancient mathematical texts (Oller-Marcén & Gairín, 2015). Furthermore, the relevance of proportional reasoning is justified by the important and extensive real-life applications (Cramer & Post, 1993) and by the presence of this skill in the official curricula and textbooks of primary and secondary education throughout the world (Lo et al., 2001; Lundberg, 2011; Martínez-Juste et al., 2017; National Council of Teachers of Mathematics [NCTM], 2000).

Considering the importance and ubiquity of proportional reasoning, it is important to conduct international comparative studies focusing both on the treatment of this skill in different countries and on the students’ possible similarities or differences according to their performance, developmental trends, etc. In relation to the former, there are several cross-national comparisons of textbooks focusing on proportional reasoning. For instance, Lo et al. (2001) compared textbooks from Japan, China, Chinese Taipei, and the United States finding many differences in terms of definitions of concepts, types of problems, and sequencing. Meanwhile, Ponte and Marques (2011) compared some grades 6 and 7 textbooks from Portugal, Brazil, Spain, and the United States, and among other results, they obtained that “in the Portuguese, Brazilian, and American textbooks the notion of ratio plays an important role, but it does not happen in the Spanish textbook” (p. 46). Similarly, Incikabi and Tjoe (2013) compared the types of proportion and ratio problems presented in Turkish and American textbooks and reported that “the US textbooks included fewer pure mathematics problems but more real-life-application problems than Turkish textbooks” (p. 8).

In terms of students’ proportional reasoning, there are studies showing that certain phenomena, like the so-called “illusion of linearity,” are common in many different countries (e.g., Fernández et al., 2012; Modestou & Gagatsis, 2007; Van Dooren et al., 2009). However, studies explicitly comparing the performance of students from different countries in this context are less common. In this regard, it is worth mentioning the work by Jiang et al. (2017) who compared 925 Chinese and Spanish fourth to eighth graders’ responses to a test of 8 missing-value word problems (4 proportional and 4 additive). Their findings indicated differences between Chinese and Spanish students’ development of additive and proportional reasoning. Moreover, Jiang et al. (2017) reported that “number structure and nature of quantities had a different impact on each country” (p. 551).

The proportional reasoning research conducted in Spain and Turkey (e.g., Arıcan, 2018; Arıcan, 2019; Atabaş & Öner, 2016; Valverde Soto & Castro Martínez, 2009) shows that both Turkish and Spanish middle school students have similar difficulties understanding ratio, proportion, and proportional relationship concepts. In fact, as we will later show, the treatment of ratio and proportion topics in middle school curricula is very much similar in both countries. However, according to Schubring (1987), “teaching practice is not so much determined by ministerial decrees and official syllabuses as

by the textbooks” (p. 41). In this respect, we have pointed out above that an indirect comparison of Turkish and Spanish textbooks reveals some similarities, but also certain differences, in their treatment of proportionality. This apparently conflicting situation motivates the interest of comparing Spanish and Turkish students’ proportional reasoning.

In the current study, Spanish and Turkish middle school students’ proportional reasoning was compared through applying a multidimensional test with 22 multiple-choice items that examined their mastery of four key cognitive skills: understanding the concept of ratio; understanding directly proportional relationships; understanding inversely proportional relationships; and understanding nonproportional relationships. In addition to being important for solving ratio and proportion problems, these four skills are the central aspects on which the mathematics curriculum is built around proportional reasoning in both countries. Moreover, the proportional reasoning test was designed from a log-linear cognitive diagnostic model (LCDM) perspective (Henson et al., 2009), which is a form of cognitive diagnostic models (CDMs). The study shows how CDMs, which are not widely used in small-scale comparative studies, can be very useful in linking similarities and differences in distribution of students’ profiles to similarities and differences in national curricula of two countries. Investigating students’ proportional reasoning by applying CDMs could demonstrate how multidimensional psychometric methods are more sensitive to these similarities and differences than traditional unidimensional models. Additionally, CDMs allow us to provide a diagnostic assessment of students’ strengths and weaknesses in terms of their proportional reasoning.

To accomplish the abovementioned objectives, this study provides a two-level comparison of Spanish and Turkish middle school students’ proportional reasoning:

- First, we compare Spanish and Turkish students’ mastery of four core cognitive skills required to solve ratio and proportion problems by applying a proportional reasoning test developed for diagnostic assessment.
- Second, we compare the performance of students in the two countries on some specific test items. To do this, we provide possible interpretations of the similarities and differences of the two samples in terms of some item characteristics such as context, type of the items, and representations.

## 2. Conceptual framework

In this section, we synthesize some key aspects of proportional reasoning, provide a description of CDMs, and present our purpose for using the LCDM in analyzing students’ proportional reasoning.

### 2.1 Proportional reasoning

The development of proportional reasoning covers a long-time interval and involves both qualitative and quantitative methods of thought as well as the comprehension of the sense of covariation between quantities (Lesh et al., 1988). Proportional reasoning is also related to the understanding of concepts such as whole-number multiplication (Hino & Kato, 2019), rational numbers (Lamon, 2007), multiplicative structures (Vergnaud, 1983), and functional modeling (De Bock et al., 2017; García et al., 2006).

In order to demonstrate competence in proportional reasoning, students should be able to identify multiplicative relationships (direct and inverse) presented in proportion problems, as well as to distinguish them from nonproportional relationships (Lim, 2009; Orrill et al., 2017). They should also have the ability to solve proportion tasks and aptitude to make arguments and deductions in a comprehensive way in proportional situations (Lamon, 1993, 2007). In addition, students should be able to represent multiplicative relationships using a variety of mathematical representations such as verbal (2 dogs for every 7 cats), symbolic–algebraic ( $2:7$ ,  $2/7$ ,  $f(x) = 2x/7$ ), iconic, tabular and/or

pictorial (including function graphs), and so forth (Arıcan, 2021; Bayazit, 2013; Martínez-Juste et al., 2015). The current study examined students' mastery of the four key cognitive skills cited above due to their crucial role in reflecting proportional reasoning competence.

Below, we elaborate on some central concepts related to proportional reasoning. As we will elaborate on the Methods section, the four key cognitive skills were developed in connection with these central concepts.

**2.1.1 Ratio and rate.** Lobato et al. (2010) define ratio as a “multiplicative comparison of two quantities, or it is a joining of two quantities in a composed unit” (p. 12). For instance, if we compare heights of two children by forming a ratio, the multiplicative relationship between their heights shows how much taller or shorter one child is than the other. Moreover, a ratio also can be formed by joining two quantities. For example, in the context of *an athlete running 100 m in 10 s*, the statement *100 m in 10 s* forms a ratio that has a new unit namely speed. On the other hand, *rate* is defined as “a set of infinitely many equivalent ratios” (Lobato et al., 2010, p. 42). Contrary to the common definition of rate, which is comparison of two quantities with different units, quantities forming a rate can have the same or different units. For instance, in the context of *mixing 3 cups of water and 2 cups of lemon juice to make lemonade*, calculating the amount of lemon juice concentrate for any amount of water demonstrates comprehension of the rate concept.

Lamon (1993) distinguishes two types of semantic problems: *well-chunked* measures (if the rate has its own meaning as a quantity of an intensive magnitude) and *associated sets* (otherwise). Moreover, regarding to the concept of *ratio*, Lamon (1993) also distinguishes two semantic types of problems: *stretchers and shrinkers* (usually related to geometric contexts) and *part-part-whole* (in which two, or more, additive homogeneous magnitudes combined to add a whole). In this last case, we can consider *part-part* and *part-whole ratios*.

**2.1.2 Proportional relationships.** Another key element of proportional reasoning is the multiplicative structure underlying the problem situations (Vergnaud, 1983). The multiplicative structure includes simple direct proportion situations (Karplus et al., 1983; Lamon, 1993; Singh, 2000b; Van Dooren et al., 2008) but also inverse and multiple proportion situations (Arıcan, 2018; Levain & Vergnaud, 1995). There are two multiplicative structures that lead to the understanding of directly and inversely proportional relationships. The first structure is isomorphism of measures, which appears in situations of direct simple proportion and provides meaning to the quotient of magnitudes. The second is product of measures that provides meaning to the product of magnitudes, especially if they are extensive (Vergnaud, 1983). These two elements are, in fact, essential to describe the situations of inverse proportion and, more generally, those of multiple proportions (Arıcan, 2018).

From a mathematical point of view, proportionality organizes phenomena modeled by relatively simple functions such as  $f(x) = kx$  in the case of direct proportion and  $f(x) = k/x$  in the case of inverse proportion. Generally, situations of multiple proportions (that may involve direct and inverse relationships between pairs of quantities) can be modeled by homogeneous functions of the form

$$f(x_1, \dots, x_n) = k \prod_{i=1}^n x_i^{a_i}, \quad a_i \in \{-1, \dots, 1\} \quad (\text{Martínez-Juste et al., 2017, 2019}).$$

Once a directly proportional relationship has been formed between two quantities, we may distinguish between *within* or *between ratios* (Freudenthal, 1983), that are central to characterize and recognize proportional relationships in terms of the processes of *invariance* and *constancy*. For example, let us consider a linear function,  $f$ , between the quantities  $x$  and  $y$ , such as  $y = f(x) = kx$ . Then, a within ratio  $x_1/x_2$  remains invariant ( $x_1/x_2 = f(x_1)/f(x_2)$ ), and the between ratio  $f(x)/x$  equals to a constant ( $f(x)/x = k$ ). Therefore, a between ratio in a linear function,  $f(x) = kx$ , generates the *constant of proportionality*,  $k$ .

To promote the development of proportional reasoning, students must face different types of rich and nonroutine tasks (Singh, 2000a). A widely used classification of proportionality tasks (Cramer &

Post, 1993) is based on the distinction between missing-value problems and comparison problems (either quantitative or qualitative). This classification is usually restricted to direct proportion situations but can be extended not only to inverse proportion situations but also to multiple proportion situations (Martínez-Juste et al., 2017, 2019). In missing-value problems, three of the four values are given and students were asked to determine the missing value (Lamon, 2007). In comparison problems, two ratios are compared to determine if they are equal, or whether one is greater or less (Lobato et al., 2010).

**2.1.3 Nonproportional relationships.** Teaching should incline students toward multiplicative reasoning by making them work within different proportional structures, but also should provide them the opportunity to apply multiplicative reasoning critically. One way to achieve this goal is to allow students to work with nonproportional relationships (Van Dooren et al., 2008).

Modestou et al. (2008) point out the tendency toward the inadequate use of linear models in perimeter, area, and volume problems, which involve not only linear, but also affine, quadratic, and cubic relationships among quantities. This overuse of linear models is rather common, and it is usually referred to as “illusion of linearity” in the literature (Van Dooren et al., 2005, 2009). In order to control this phenomenon, working with affine relationships can be helpful, especially in the case of middle school students due to their cognitive stage (Fernández et al., 2012).

## 2.2 Cognitive diagnostic models

Traditional psychometric models provide a single score for a continuous estimation of examinees’ overall ability of a subject such as mathematics (Bradshaw & Templin, 2014). However, these methods are limited in the sense of providing diagnostic feedback on multiple dimensions of examinees’ cognitive ability. Hence, in the field of educational and psychological measurement, there has been a continuously growing attention on statistical models with latent variables that allow “multi-dimensional classifications of respondents for the purpose of a fine-grained diagnosis” (Rupp & Templin, 2008, p. 220). CDMs are a family of psychometric models that have been designed for this diagnostic assessment purpose (Rupp et al., 2010).

Instead of measuring an overall mathematics ability, CDMs break mathematics down into a set of skills (usually called attributes) and classify examinees as master or nonmaster of these skills based on their responses to the given items. An *attribute* is a categorical latent trait that we want to estimate. The critical step in CDMs is to specify attributes. Tatsuoaka et al. (2016) stated that two conditions should be satisfied when specifying attributes: attributes should capture fine-grained understanding, and they should be workable. The four core cognitive skills that were presented in the introduction are critical for developing proportional reasoning (as shown in the theoretical framework section). Hence, the attributes provide fine-grained understanding for students’ proportional reasoning, and they are workable since there exists many researches on them. Therefore, the purpose of CDMs is to offer *diagnostic feedback* regarding these carefully defined attributes (Bradshaw et al., 2014).

The view of development that underlies the assumptions of CDMs “is more consistent with the nature of proportional reasoning by Lesh et al. (1988)” (Tjoe & de la Torre, 2014, p. 238). The current study uses the LCDM to investigate Spanish and Turkish middle school students’ mastery of four attributes. The LCDM was designed as an effort to extend a general diagnostic model (Von Davier, 2005) to include latent variable interaction effects (Bradshaw & Templin, 2014). We used the LCDM to compare students’ proportional reasoning because it allows for a mix of items in which there are and are not interactions between attributes. A second reason is that the LCDM provides “empirical information regarding a reasonable model” (Henson et al., 2009, p. 208) (see Table 4 for the selection process of a reasonable model). If, for instance, we consider an item measuring two attributes ( $\alpha_{e1}$  and  $\alpha_{e2}$ ), the LCDM estimates the probability of an examinee’s ( $e$ ) correct

response for an item ( $i$ ) applying the following equation:

$$P(X_{ei} = 1 | \alpha_e) = \frac{\exp(\lambda_{i,0} + \lambda_{i,1(1)}\alpha_{e1} + \lambda_{i,1(2)}\alpha_{e2} + \lambda_{i,2(1,2)}\alpha_{e1}\alpha_{e2})}{1 + \exp(\lambda_{i,0} + \lambda_{i,1(1)}\alpha_{e1} + \lambda_{i,1(2)}\alpha_{e2} + \lambda_{i,2(1,2)}\alpha_{e1}\alpha_{e2})} \quad (1)$$

In Equation 1, the parameter  $\lambda_{i,0}$  is called *intercept* and represents the predicted log-odds of a correct response for individuals who have not mastered any attribute. The parameter  $\lambda_{i,1(1)}$  is the *simple main effect* that represents the predicted log-odds of a correct response for individuals who have mastered attribute 1, but not attribute 2. Similarly,  $\lambda_{i,1(2)}$  represents the predicted log-odds of a correct response for individuals who have mastered attribute 2, but not attribute 1. Finally,  $\lambda_{i,2(1,2)}$  is the *two-way interaction effect* that represents the predicted log-odds of a correct response for individuals who have mastered both attributes 1 and 2. Equation 1 can naturally be extended to consider any number of attributes and multiple ways of interactions.

### 3. Methods

To achieve the previously introduced research objectives, we carry out a cross-sectional study with exploratory and descriptive purposes. Furthermore, as we will elaborate below, we adopt a quantitative approach both for the data collection and analysis (Bradburn & Gilford, 1990).

#### 3.1 The students' curricular background

In both countries, students are introduced to the ratio concept for the first time in sixth grade. In Turkey, the national mathematics curriculum (Turkish Ministry of National Education [TMNE], 2018) necessitates students to use the ratio concept (and its different representations) to compare quantities, the computation of ratios of two quantities (with the same or different units), and the work with part–part and part–whole ratios (see Table 1). The Spanish sixth grade national mathematics curriculum (Spanish Ministry of Education, Culture and Sports [SMECS], 2015) only introduces the use of the “rule of three” to solve real-life direct proportion problems. In Turkey, students are introduced with proportions in seventh grade. On the other hand, the Spanish curriculum provides proportion instruction in seventh and eighth grades together. The fact that the Spanish curriculum does not explicitly state which contents must be addressed in seventh or eighth grade obliged us

**Table 1.** The middle school ratio and proportion standards in Turkey and Spain (SMECS, 2015; TMNE, 2018).

Turkey	Spain
Determine the unit ratio equivalent to a given ratio.	Know the concepts of ratio and proportion, directly and inversely proportional quantities, constant of proportionality and linear function.
Determine one of the terms of a given ratio, when the other term is known.	
Decide, in real-life situations, if two quantities are directly proportional to each other.	Solve real-life problems involving direct or inverse proportion using different strategies (use of tables, use of the constant of proportionality, use of the unit ratio, etc.).
Express the relation between two proportional quantities.	
Determine and interpret the constant of proportionality defined by two proportional quantities.	Identify, in real-life situations, proportional quantities and decide whether they are directly proportional or inversely proportional to each other.
Decide, in real-life situations, if two quantities are inversely proportional to each other.	Recognize, in real-life situations, quantities that are neither directly nor inversely proportional to each other.
Solve problems involving direct and inverse proportion.	

to compare seventh grade Turkish students with eighth grade Spanish students. Therefore, we can be certain that primary instruction on ratio and proportion was complete in each country.

A quick analysis of Table 1 shows that the curricular standards of both countries are very similar. They are articulated around the same main ideas: the concept of a ratio, working with direct proportional structures, working with inverse proportional structures, understanding nonproportional relationships, and solving daily-life problems.<sup>1</sup>

### 3.2 Sample

Bradburn and Gilford (1990) point out that “studies of relatively small, localized samples in a small number of sites can also play an important role in comparative education” (p. 26). For this study, the convenience sample (Neuman, 2007) consisted of 596 students: 282 Turkish seventh grade students and 314 Spanish eighth grade students. As we stated above, in Turkey, ratio, proportion, and proportional relationships are taught in sixth and seventh grades while in Spain it takes up to eighth grade. Hence, all the participating students had already completed instruction about ratio and proportion before taking the proportional reasoning test. Finally, both the Spanish and Turkish samples consisted of students from a single district. The relative centralization of the two countries’ education systems, both of which follow a centralized national curriculum, implies that this is neither a concern nor a threat in terms of representability.

### 3.3 The development of the proportional reasoning test

The proportional reasoning test was developed by Arıcan (2019) following an LCDM perspective. The test included 22 multiple-choice items (see Appendix I) that took the key elements of proportional reasoning into consideration. The items involved understanding ratio and rate concepts, direct and inverse proportional relationships, and nonproportional relationships. Moreover, the test included items with different underlying structure (simple proportion, multiple proportions, and affine relations). In addition, the items included most of the semantic types of problems considered by Lamon (1993).

The test was specifically designed in order to measure middle school students’ mastery of the following four attributes:

- A1: Understanding the concept of a ratio and determining the value of a quantity in a given ratio.
- A2: Understanding directly proportional relationships and solving problems involving this type of relationships.
- A3: Understanding inversely proportional relationships and solving problems involving this type of relationships.
- A4: Understanding nonproportional relationships and solving problems involving this type of relationships.

The four attributes were determined by examining relevant research, Turkish curriculum, and international standards (Arıcan, 2019). These attributes reflect competence in proportional reasoning, and their development follows concepts central to proportional reasoning (see Section 2.1). In particular, A1 refers to the concept of ratio, whereas A2 and A3 are related to the understanding of proportional relationships (both direct and inverse). On the other hand, A4 accounts for understanding of nonproportional relationships such as additive and affine relationships and distinguishing directly and inversely proportional relationships from nonproportional relationships. These attributes also reflect the main concepts that underpin curriculum standards in both countries (see Table 1). Further details about the design and validation of the test can be found in Arıcan (2019).

According to Arıcan (2019), the *Q-matrix* (Tatsuoka, 1985) that describes the alignment between items and attributes was obtained as a result of four mathematics education experts independently coding the 22 items in relation to their measured attributes. The Q-matrix is presented in Table 2, where a “1” entry represents that the attribute is measured by the item.

The test items can be further classified along two dimensions: representations and task context. Regarding the representations, we combine the categories described by Bayazit (2013) and by Martínez-Juste et al. (2015). Particularly, we consider verbal (V), symbolic-algebraic (S), tabular (T), and pictorial (P) representations. On the other hand, regarding the task context, we distinguish between intramathematical (IM) and nonmathematical (NM) contexts (Organization for Economic Co-operation and Development [OECD], 2004). Table 3 provides the description of each item according to these dimensions and categories.

The frequency and distribution of the different representations and contexts among the items is very similar to that of textbooks and traditional teaching practices in two countries. These two variables were not explicitly taken into account during the development of the proportional reasoning test, but we believe they can provide useful information and insight during the detailed item-level analysis of the students’ responses.

### 3.4 Data collection and analysis

In Turkey, the data were collected during the 2018 spring semester. The students were given 40 min to complete the test in person, and five preservice teachers assisted the researcher in collecting the students’ responses. On the other hand, in Spain, the data were collected during the 2020 spring semester. Due to the COVID-19 outbreak, the students were given 50 min to complete the test online using Google Forms<sup>2</sup> (10 additional minutes were granted in case there were technical issues), and the researcher collaborated with the students’ teachers in collecting all the responses. The data collection and analysis were performed separately in the two countries for comparison purposes and to get a better insight into the students’ attribute profiles. All the required permissions and consents were obtained, in both countries, prior to the test administration. This guaranteed that all the participating students did so on a voluntary basis. In each country, the test was administered in the corresponding language, and the only difference was that the names used in test items were adapted to be Turkish and Spanish names, respectively.

The students’ responses to the proportional reasoning test were coded dichotomously (0/1 for incorrect and correct responses, respectively). Using the Q-matrix and these coded responses, the data were analyzed using the Mplus 6.12 statistical software (Muthen & Muthen, 2011). Based on the number of attributes used, the LCDM estimates intercept, simple main effects, and interaction effects (see Equation 1). In the current study, we had four attributes and, since no item measured these four attributes simultaneously, we only had to consider one-way, two-way, and three-way log-linear structural model parameterizations in our LCDM model. The one-way structural model estimated the intercept and simple main effects. On the other hand, the two-way structural model estimated the intercept, simple main effects, and two-way interaction effects (e.g.,  $\lambda_{i,2(1,2)}$ ). Finally, the three-way structural model estimated the intercept, simple main effects, two-way interaction effects, and three-way interaction effects (e.g.,  $\lambda_{i,3(1,2,3)}$ ). Using the intercept, main effects, two-way effects, and three-way effects, the LCDM calculates proportions for deciding mastery and nonmastery of attributes for each examinee. In general, proportions below .40 are considered as nonmastery and .60 and above are considered as mastery. However, proportions in between .40 and .60 are questionable. For these questionable proportions, the LCDM applies a model-internal criterion for deciding mastery status. More information regarding the mastery decision process can be found in Rupp et al. (2010).

The first step of the data analysis was to determine the best-fit model. Arıcan (2019) reported that for the Turkish sample the two-way model fitted the data better than the one-way and three-way



**Table 2.** The Q-matrix [retrieved from Arıcan (2019)].

Attribute	Items																						Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	
A1	1	1	1	1	0	1	1	1	0	0	1	0	1	0	1	1	1	1	0	1	1	0	13
A2	1	1	1	1	0	1	1	0	0	0	0	1	0	0	1	1	0	1	1	1	1	1	13
A3	0	0	0	0	0	0	0	1	1	1	1	0	0	0	1	0	1	0	1	0	1	1	9
A4	0	0	0	0	1	0	0	0	0	0	1	0	1	0	1	0	0	0	1	0	1	0	5

**Table 3.** Representations and task contexts involved in each test item.

Item	Representations	Context	Item	Representations	Context
1	V/S	IM	12	V	NM
2	V/S	NM	13	V/S	IM
3	V	IM	14	V	NM
4	V	NM	15	V/T	NM
5	V	NM	16	T	NM
6	V	NM	17	V/S	IM
7	V	NM	18	V/P	NM
8	V	NM	19	P	IM
9	V	NM	20	V/P	NM
10	V	NM	21	P	IM
11	V/S	IM	22	V	NM

Note. IM= intramathematical; NM= nonmathematical; P= pictorial; S= symbolic-algebraic; T= tabular; V= verbal.

**Table 4.** Chi-square test results and fit indices for one-way, two-way, three-way, and the final model for the Spanish sample.

Model	AIC	BIC	SSA BIC	LL	NPR	Chd	df	p
One-way	7618.18	7906.88	7662.66	-3732.089	77	-	-	-
Two-way	7605.55	7969.24	7661.58	-3705.773	97	52.63	1	.000
Three-way	7605.95	7977.14	7663.14	-3703.974	99	3.59	1	.057
Best-fit	7588.81	7877.51	7633.29	-3717.403	77	-	-	-

Note. AIC=Akaike's information criteria; BIC=Bayesian information criteria; Chd=chi-square difference; df=degrees of freedom; LL=log-likelihood; NPR=number of estimated parameters; SSA=sample size adjusted.

models. For the Spanish case, conducting a chi-square difference test using the log-likelihood values with the Maximum Likelihood Robust estimator (Satorra & Bentler, 2010), we compared one-way, two-way, and three-way models with each other (Table 4). In Table 4, a significant chi-square difference ( $p < .05$ ) suggests that the larger model with more freely estimated parameters would better fit the data than the smaller model with less freely estimated parameters (Werner & Schermelleh-Engel, 2010). However, a nonsignificant chi-square difference indicates that both models fit the data equally well. In that case, the data estimation should be carried out using the model with smaller information model fit indices (i.e., Akaike's information criteria, Bayesian information criteria [BIC], and sample size adjusted BIC). The analysis suggested that the two-way model fitted the data better than one-way and three-way models. Moreover, again following Arcan (2019), we removed one by one nonsignificant interactions and simple main effects from the two-way model by paying attention to the model fit indices at each step. Our best-fit model for the Spanish sample included 77 freely estimated parameters.

### 3.5 Item-attribute discrimination and difficulty indices

Item-attribute discrimination indices show how well an item discriminates between masters and nonmasters of the required attribute or attributes (Henson et al., 2008). An index value 0 indicates that masters and nonmasters of an attribute have the same chance for answering the item in question. On the other hand, an index value of 1 indicates that only masters of the attribute can answer that item, which shows perfect discrimination between masters and nonmasters. Although there is not

a cut-score for determining poor discrimination indices, .30 and below usually suggest low discrimination indices (Arıcan, 2019; de la Torre, 2008). Therefore, Table 5 shows that overall, the items discriminated well between masters and nonmasters of the required attributes in both countries.

Table 5 presents item difficulty indices for both samples. Item difficulty index shows the proportion of those who answered an item correctly and ranges from 0 to 1. An index close to 0 indicates that the item is difficult and an index close to 1 indicates that the item is easy. As reported by Arıcan (2019), item difficulty index for the Turkish sample ranged between .25 and .93, with a mean of .59. In the case of Spain, item difficulty index ranged between .19 and .92, with a mean of .53. While items 5 and 14 turned out to be the most difficult items for Turkish students (being answered by only 26% and 25% of the students, respectively), Spanish students had most difficulty in answering items 17, 2, and 3. On the other hand, item 7 was the easiest one for students in both countries, being correctly answered by 93% and 92% of the students, respectively. Although this information is not included in Table 5, it is worth focusing on the most unanswered items (see Appendix II). In this regard, items 10 and 22 were left unanswered by around 18% and 16% of the Turkish students, whereas 12% and 7% of the Spanish students left unanswered items 3 and 17.

We have performed a Fisher exact test to compare the difficulty index in each item (grouping together incorrect and blank answers). The  $p$  values reported in Table 5 show that there were significant differences in students' performances in two countries on 16 out of 22 items at a 95% confidence level. However, we should note that Spanish students performed better than Turkish students on only 5 of these 16 items.

**Table 5.** Item-attribute discrimination and difficulty indices for the Turkish and Spanish sample [Turkish data were retrieved from Arıcan (2019)].

Items	Item-attribute discrimination indices								Difficulty indices		
	Turkey				Spain				Turkey	Spain	$p$
	A1	A2	A3	A4	A1	A2	A3	A4			
1	.46	.49			.61	.65			.84	.71	.000
2	.85	.69			.43	.66			.72	.22	.000
3	.88	.85			.40	.54			.45	.24	.000
4	.51	.37			.61	.60			.87	.70	.000
5				.78				.87	.26	.55	.000
6	.71	.70			.63	.68			.61	.61	.933
7	.42	.32			.35	.40			.93	.92	.646
8			.773				.75		.61	.49	.003
9			.814				.89		.49	.61	.006
10			.886				.82		.45	.37	.045
11	.85		.732		.67		.67		.60	.31	.000
12				.84				.60	.66	.63	.440
13	.59	.52			.80	.66			.82	.62	.000
14				.84				.77	.25	.55	.000
15	.85		.788		.83		.91		.58	.44	.000
16	.78	.77			.42	.35			.73	.85	.000
17	.83		.688		.37		.42		.39	.19	.000
18	.81	.77			.64	.65			.63	.72	.014
19		.59	.712	.97		.52	.48	.73	.37	.36	.608
20	.71	.80			.78	.80			.76	.70	.096
21		.77	.775	.91		.46	.52	.52	.55	.37	.000
22		.76	.730			.55	.65		.56	.64	.053

## 4. Results

The findings are presented in two sections: first, we compared the students in two countries in terms of attribute profiles that they belonged and attribute mastery status; second, we present an item-level analysis for selected items that provide valuable information about the two samples.

### 4.1 Attribute profile and attribute mastery comparison

In Table 6, the estimated number of students belonging to each attribute profile and their percentages were presented for both countries. Since there were four attributes,  $16 (=2^4)$  attribute profiles obtained as a result of the LCDM analysis in which 0 and 1 represent mastery and nonmastery of attributes, respectively. Table 6 shows that Spanish students belonged to more attribute profiles, which indicated variety in their profiles, than Turkish students. Spanish students mostly belonged to attribute profiles 0001, 0101, 0111, and 1111. On the other hand, Turkish students mostly belonged to attribute profiles 0000, 0100, and 1111. Hence, the Turkish sample mostly consisted of students who either mastered or nonmastered the four attributes and students who mastered attribute two alone. However, there were less Spanish students who mastered all four attributes and did not master any of these four attributes.

Table 6 shows that the estimated distribution of students among the different profiles varies in each country. A chi-square test was used in testing the association between the variables “Country” and “Attribute profile” that showed a significance association ( $p < .01$ ). In fact, there was a very strong level of association because we obtained a value of .765 for Cramer’s V coefficient (Blaikie, 2003). This means that the students are distributed among the profiles quite differently in each country, leading to a (so to speak) different country profile regarding proportionality. More particularly, a post hoc z-test on the adjusted residuals with Bonferroni correction reveals that there are significant differences ( $p < .05$ ) between two countries regarding the attribute profiles 0000, 0001, 0100, 0101, 0110, 0111, 1011, 1101, and 1111. This implies that the profiles in which there are significant differences between both countries are mostly low-performance (mastering at most one attribute) and high-performance profiles (mastering three or more attributes).

The LCDM analysis provides estimations for the intercept, main, and interaction effect parameters. For instance, in Equation 1, the intercept (i.e.,  $\lambda_{i,0}$ ) parameter represents the predicted

**Table 6.** Estimated number of Spanish and Turkish students belonging to each attribute profile and their proportions [Turkish data were retrieved from Arican (2019)].

Attribute profile	Spain	Turkey
0000	11 (3.6%)	71 (25%)
0001	85 (26.9%)	0
0010	2 (0.7%)	0
0100	6 (2%)	43 (15.3%)
1000	5 (1.7%)	5 (1.6%)
0011	3 (1.1%)	0
0101	63 (20.1%)	0
0110	23 (7.3%)	0
1001	16 (5%)	21 (7.6%)
1010	4 (1.4%)	2 (0.8%)
1100	0	0
0111	41 (13.1%)	0
1011	0	8 (2.9%)
1101	0	7 (2.4%)
1110	16 (5%)	14 (4.9%)
1111	38 (12.3%)	111 (39.2%)

log-odds of a correct response for individuals who have not mastered any attribute. Equation 1 takes the form of  $\frac{\exp(\lambda_{i,0})}{1+\exp(\lambda_{i,0})}$  when only intercept parameters are replaced in the equation for each item. Replacing  $\lambda_{i,0}$  parameter estimations for each item in the formula, we calculated the estimated proportions of intercepts for both countries (Table 7). The calculations showed that in average 37.4% of nonmasters in the Turkish sample and 34.1% of nonmasters in the Spanish sample answered items correctly. Hence, there was a similarity between the two samples when it came to nonmasters answering items correctly. According to Table 7, nonmasters found items 5 and 14 to be the most challenging in the Spanish sample, while items 9 and 19 were the most challenging in the Turkish sample. For nonmasters in the Spanish sample, items 7 and 12 were the easiest, while in the Turkish sample, items 4 and 7 were the easiest.

It is also important to compare the percentages of low-performing and high-performing students in both countries. As we have just mentioned, we consider that low-performing students are those mastering at most one attribute, and high-performing students are those mastering three or more attributes. Table 8 shows that the Turkish sample mostly consisted of either low-performing or high-performing students. The Spanish sample, on the other hand, had a more dispersed distribution of students.

**Table 7.** Intercept parameter estimates, standard errors, and estimated proportions for the Turkish and Spanish sample [Turkish data were retrieved from Arican (2019)].

Spain				Turkey		
Items	$\lambda_{i,0}$	Standard errors	Proportion	$\lambda_{i,0}$	Standard errors	Proportion
1	0.103	.23	.526	0.086	.27	.521
2	-1.684	.19	.157	-0.037	.26	.491
3	-1.895	.34	.131	-1.363	.23	.204
4	-0.022	.23	.495	1.006	.22	.732
5	-21.724	1	.000	-0.294	.21	.427
6	-0.160	.21	.460	-0.652	.25	.343
7	1.410	.32	.804	1.002	.29	.731
8	-1.160	.20	.239	-0.597	.19	.355
9	-0.609	.21	.352	-1.462	.24	.188
10	-1.546	.22	.176	-1.106	.23	.249
11	-1.516	.23	.180	-0.880	.21	.293
12	0.570	.12	.639	-0.579	.21	.359
13	-0.927	.28	.284	0.405	.20	.600
14	-27.251	1	.000	-0.675	.20	.337
15	-1.535	.29	.177	-0.957	.22	.277
16	0.642	.28	.655	-0.376	.28	.407
17	-1.843	.23	.137	-0.913	.22	.286
18	-0.021	.28	.495	-0.541	.19	.368
19	-0.697	.24	.332	-1.417	.32	.195
20	-0.837	.36	.302	-1.286	.39	.217
21	-0.458	.28	.387	-0.895	.27	.290
22	0.293	.16	.573	-0.586	.27	.358

**Table 8.** Estimated percentages of low-performing and high-performing students in Spain and Turkey.

	Low performance	High performance
Spain	34.9%	30.4%
Turkey	41.9%	49.4%

Proportions of Spanish and Turkish students' attribute mastery are calculated by summing up corresponding proportions in Table 6. Table 9 shows the proportions of Spanish and Turkish students' mastery of each attribute and reliability indices.

The corresponding two proportions  $z$ -test shows that there are significant differences in attribute mastery between both countries only regarding A1 and A4. This finding shows that while Turkish students were better at "understanding the concept of a ratio and determining the value of a quantity in a given ratio," Spanish students were better at "understanding nonproportional relationships and solving daily-life problems involving this type of relationships." In Table 9, the reliability indices were slightly lower for the Spanish sample than for the Turkish sample. However, this result is understandable since the proportional reasoning test was originally designed by following the Turkish middle school mathematics curriculum. On the other hand, the reliability indices for the Spanish sample also indicated highly reliable attribute mastery estimations.

#### 4.2 Detailed analysis of the answers to some items

As can be seen in Table 5, there were some items in which the performance of Spanish and Turkish students differ. In those cases, it is interesting not only to compare the item difficulty indices (i.e., the proportions of students that correctly answered items) but also to analyze the different incorrect answers given by the students (the specific distribution of answers can be found in Appendix II). Particularly, we focus only on those items for which the difference between the difficulty indices was greater than 20%. Namely, we focus on items 2, 3, 5, 11, 13, 14, and 17.

**Item 2:** There are some number of red and blue balls in a container. The ratio of the number of red balls to the number of blue balls is  $\frac{3}{4}$ . If there are a total of 42 balls in this container, how many of them are blue?

This item is presented in a NM context using a verbal-symbolic representation. The difference between the difficulty indices for this item was the greatest (50%) and in Turkish students' favor. Almost one-half of the Spanish students chose 12 as their answer (possibly due to considering the data  $\frac{3}{4}$  as a part-whole ratio and then rounding the result of computing  $\frac{1}{4}$  of 42), while less than 9% of the Turkish students chose that option.

**Item 3:** The measures of outer angles of a triangle are proportional to the numbers 9, 12, and 15. What is the measure of the greatest internal angle of this triangle?

This item involved a proportional situation in a geometrical context using a verbal representation. The Turkish students performed better in this item, even if only 45% of them answered correctly. This item was left unanswered by about 12% of Spanish students (being the most unanswered item of the test). Moreover, the students who provided an answer, opted almost uniformly between the four options.

**Item 5:** Two different but identical candles, A and B, are burning at the same rate but they are lit at different times. We know that when candle A has burned 24 mm, candle B has burned 18 mm. When 36 mm of candle B has burned, how many millimeters will candle A have burned?

**Table 9.** Spanish and Turkish students' attribute mastery proportions and reliability indices [Turkish data were retrieved from Arıcan (2019)].

	A1		A2		A3		A4	
	Spain	Turkey	Spain	Turkey	Spain	Turkey	Spain	Turkey
Mastery	.254	.595	.596	.619	.407	.479	.784	.522
Reliability	.91	.99	.88	.97	.96	.99	.93	.96

This item has the same structure as a missing-value problem, but in an additive situation. It was presented in a NM context using a verbal representation. The Spanish students performed better on this item than their Turkish counterparts. In fact, this was one of the items in which Turkish students performed poorly (only 26% of them provided correct answers). In both cases, the most commonly chosen wrong answer was 48, which implies treating the situation as if it was proportional.

**Item 11:**  $x$  and  $y$  are two inversely proportional quantities. We know that when  $x$  is 12,  $y$  is 4. When  $x$  is 6, what is the value of  $y$ ?

This item is a missing-value problem in an inversely proportional situation. It was presented in a mathematical context using a verbal-symbolic representation involving algebraic notation. The Turkish students performed much better on this item than the Spanish students (only 31% of the Spanish students provided correct answers). In both countries, the most popular incorrect answer was 2 (about 30% of the Turkish students and about 55% of the Spanish students selected this answer). A possible explanation is that it would be the right choice if the situation was directly proportional.

**Item 13:**  $a$  and  $b$  are two numbers that vary in a directly proportional relationship. We know that when  $a$  is 5,  $b$  is 15. So, when  $b$  is 24, what is the value of  $a$ ?

This item is a missing-value problem in a directly proportional situation. It was presented in a mathematical context using a verbal-symbolic representation involving algebraic notation. Both countries performed quite well, but the results were particularly good for the Turkish students (82% of correct answers). The most common wrong answer in both cases was 14, which implied the use of additive reasoning.

**Item 14:** Ayse and Esra, who are running at the same speed, started running around an 800 m circular track at different times. It is known that when Ayse ran 300 m, Esra ran 200 m. According to this information, when Esra ran 400 m, how many meters can Ayse run?

Similar to item 5, this item was written as a missing-value problem, but in an additive situation. It was presented in a NM context using a verbal representation. Again, the Spanish students performed better than Turkish ones. Moreover, this was the item that Turkish students had the most difficulty answering correctly (only 25% answered correctly). Also as in item 5, the most commonly chosen wrong answer (in both countries) was 600, which implies treating the situation as if it was proportional.

**Item 17:**  $(m + 3)$  and  $(n + 1)$  are two numbers that vary in an inversely proportional relationship. We know that when  $m$  is 1,  $n$  is 2. When  $m$  is 3, what is  $n$ ?

This item was written as a missing-value problem in an inversely proportional situation. It was presented in a mathematical context using a verbal-symbolic representation involving algebraic notations, and the solution of this item required algebraic manipulations. This item obtained the worst combined result from the two samples (with only 19% of correct answers in Spain and 39% in Turkey). As for the most preferred wrong answer, Spanish students mostly chose 4 (about 50%) as their answers. This incorrect selection seems to imply omitting some of the required algebraic manipulations. On the other hand, the Turkish students equally selected 2 and 4 (about 25% and 21%, respectively) as their answers.

## 5. Discussion and conclusions

CDM are widely used statistical tools that are very useful in investigating examinees' mastery status of certain skills (Arıcan & Kuzu, 2020). Furthermore, these models are also fruitfully used to perform large-scale detailed comparisons between different student populations (e.g., Dogan & Tatsuoka, 2008; Şen & Arıcan, 2015). In the current study, we applied the LCDM in a small-scale study to compare Spanish and Turkish middle school students' mastery of four core cognitive skills related to proportional reasoning. Curriculum standards of both countries place great emphasis on these

skills because of their critical role in developing students' proportional reasoning. The application of the LCDM, which is not common in small-scale comparative research, turned out to be very useful and allowed us to identify some significant differences between the two samples and to obtain different attribute profiles and estimated distributions of the students belonging to these profiles. As Hilton et al. (2013) point out, the use of these type of instruments in small-scale samples can also be useful because they provide teachers with information about their own students that can even be used in designing professional development workshops.

The data analysis indicated significant differences in Spanish and Turkish students' mastery of two attributes. The first attribute in which we determined a significant difference was attribute 1, defined as "understanding the concept of a ratio and determining the value of a quantity in a given ratio." In average, 25.4% of Spanish students and 59.5% of Turkish students mastered this attribute. This attribute involves understanding the ratio concept, which is regarded as a core concept of proportional reasoning. This finding is consistent with the finding provided by Ponte and Marques (2011) who stated that in comparison to Brazilian, Portuguese, and American middle school mathematics textbooks, less attention is given to the notion of ratio in Spanish textbooks. While the ratio concept is covered both in Spanish and Turkish curricula, Martínez-Juste et al. (2017) point out that Spanish textbooks do not pay much attention to conceptual aspects of proportionality.

The difference in the mastery proportion of attribute 1 is particularly clear in the light of items 1, 2, and 4. These three items are missing-value problems in which a ratio and an "absolute" quantity were given. In all of these three items, Spanish students performed significantly worse than their Turkish counterparts (see Table 5) but the difference was particularly remarkable in item 2, which was correctly answered only by 22% of Spanish students versus 72% of Turkish students. This item involves a part-part-whole situation (Lamon, 1993) in which the part-part ratio and whole were given, the students were asked to determine one of the missing parts. Some researchers have pointed out that these so-called *proportional division problems* may be difficult for students (Alatorre & Figueras, 2004; Wright, 2014). This difficulty seems to be supported by our results, and the poor performance of Spanish students might be partly explained by the certain lack of this type of problems in Spanish textbooks, just like it happens in the US setting (I et al., 2020).

The second attribute with a significant difference between both countries was attribute 4, defined as "understanding nonproportional relationships and solving daily-life problems involving this type of relationships." While an average of 78.4% of Spanish students mastered this attribute, 52.2% of Turkish students mastered it. This attribute is especially important when dealing with real-life problems since there seems to exist a "students' tendency to overgeneralize the range of applicability of the proportional model" (Van Dooren et al., 2005, p. 59).

The difference in the mastery of attribute 4 is particularly clear in the light of items 5 and 14, in which Spanish students performed significantly better than their Turkish counterparts (see Table 5). These two problems were mathematically identical, having the same structure as a missing-value problem, but in an additive situation. Around 55% of Spanish students answered both items correctly, while less than 26% Turkish students provided correct responses. In both countries, the most popular wrong answer (36%–38% in Spain and 65%–66% in Turkey, respectively) consisted in reasoning in a proportional way illustrating the well-known phenomenon of "illusion of linearity" (Van Dooren et al., 2008). Hence, as reported by Atabaş and Öner (2016), Turkish students' difficulties appeared to be mostly related with their lack of experience with this type of additive situations. However, it must be noted that, in item 12, which involved a missing-value problem in an affine situation, the students from both countries performed similarly well (63% of correct answers in Spain and 66% in Turkey). This item did not involve the comparison of two quantities, but rather required reasoning with a single quantity (i.e., univariate reasoning). Therefore, the students in two countries might have solved this item more easily than items 5 and 14.

Turkish students' overuse of directly proportional strategies is also clear especially in item 9 (a missing-value problem with an inversely proportional relationship). Almost 40% of Turkish students



answered this problem as if it was a direct proportion situation, while only 13% of Spanish students did so. It must be noted, however, that the performance in this item was not significantly different (see Table 5). In this regard, Fernández et al. (2012) noted that a progressive transition from overusing additive methods to overusing multiplicative methods takes place along primary and secondary education and that it occurs at different ages depending on the country. Although Spanish students (eighth grade) were 1 year ahead of Turkish students (seventh grade), the transition from overuse of additive methods to overuse of multiplicative methods was more apparent for Turkish students. Thus, the real difficulty appears to be the lack of understanding about which model (multiplicative and additive) is the correct one to use, and not the lack of skills to solve certain problems.

When we focus on the different attribute profiles (Table 6), it is noteworthy that approximately 80% of the Turkish students are distributed across three profiles (0000, 0100, and 1111) with seven profiles containing no students. Hence, the Turkish sample mostly consisted of either low-performing or high-performing students. On the other hand, we find only three empty profiles for the Spain sample. Therefore, Spanish students had a more dispersed distribution of students among different attribute profiles. Our findings are somewhat consistent with the Programme for International Student Assessment (PISA) 2018 results (OECD, 2019), in comparison to Spanish students, we found a greater difference in Turkish students' performance.

The results obtained from some items show significant differences between the Turkish and Spanish students. These differences could be interpreted in terms of the problem contexts and external representations used in presenting problems. Particularly, we observe that Spanish students had a lower performance than their Turkish peers in all the items with IM contexts (items 1, 3, 11, 13, 17, 19, and 21). A general explanation for this result could be that, in Spain, proportionality is usually presented in an arithmetic context involving contextualized situations (Martínez-Juste et al., 2015). This type of presentation yields a disconnection between the ratio and proportion and Geometry, Algebra, and Functions in which proportionality and linear models also play a critical role (García et al., 2006).

In item 3, one additional reason for the Spanish students' poor performance could be explained by the fact that the problem involved the knowledge of geometry (the exterior angle of a triangle) which is not explicitly included in the Spanish curriculum. On the other hand, regarding the remaining items presented in an IM context, both those involving algebraic language (items 11, 13, and 17) and those involving pictorial representations using function graphs (items 19 and 21), Spanish students' poor performance might also be related to the different representations used in problem statements. In fact, works such as those of Van Dooren et al. (2008) or De Bock et al. (2017) point out the influence of different external representations over students' successes of solving tasks involving proportionality. In items 19 and 21, we should point out that, the graphs of directly inversely proportional relationships are not introduced in the Turkish middle school curriculum, and in Spain, only the graph of directly proportional relationships is introduced in seventh grade.

To conclude, as Cai et al. (2016, p. 9) point out:

Given the fact that simple comparisons of international rankings provide little guidance for understanding and improving students' mathematics learning, it is important to understand international performance differences by other means. [...] Therefore, in order to provide the education community with a deeper understanding of the teaching and learning of mathematics, it is essential for some international comparative studies to provide evidence of students' thinking and reasoning beyond correctness of answers to mathematical problems. The evidence of students' thinking includes the qualitative analysis of solution strategies, mathematical errors, mathematical justifications, and representations.

A limitation of the current study is that it included a relatively small sample of Spanish and Turkish middle school students. Moreover, the students' proportional reasoning was compared

using quantitative techniques alone. Hence, we suggest future studies to use this study as the basis for designing a wider comparative study that should combine quantitative data with qualitative data to better understand students' proportional reasoning. Future studies should also examine students' reasoning and performances in different topics and subject areas.

### Acknowledgements

This work was carried out within the research group "Investigación en Educación Matemática" (S60\_20R) officially recognized by Gobierno de Aragón (SM-J, JMM-E, and AMO-M). Parts of this study were presented in the first International Online Conference on Mathematics Education, Istanbul, Turkey in 2021.

### Contributorship

The proportional reasoning test was designed by Muhammet Arıcan and applied to the Turkish sample. Sergio Martínez-Juste translated the same test to Spanish and applied to the Spanish sample. The LCDM analysis was performed by Muhammet Arıcan and item-level analysis was performed by Sergio Martínez-Juste, José M. Muñoz-Escolano, and Antonio M. Oller-Marcén. The first draft of the manuscript was written by all the authors. All authors read and approved the final manuscript.

### Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

### Informed consent

Informed consent was obtained from all individual participants included in the study. Ethical approval was obtained from Kırşehir Ahi Evran University with the registration number: 201803072484825651

### Declaration of Conflicting Interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


### Funding


The author(s) received no financial support for the research, authorship, and/or publication of this article.

### ORCID iDs

Sergio Martínez-Juste  <https://orcid.org/0000-0002-2607-6789>

Muhammet Arıcan  <https://orcid.org/0000-0002-0496-9148>

José M. Muñoz-Escolano  <https://orcid.org/0000-0002-8713-4591>

Antonio M. Oller-Marcén  <https://orcid.org/0000-0002-8191-3199>

### Notes

1. We must point out that since the moment this research was conducted, a new curriculum was introduced in Spanish middle school. However, there are no significant changes in the new curriculum regarding ratio and proportion standards.
2. See <https://forms.gle/MrLMfzVf21NoJPaw6>.

### References

- Alatorre, S., & Figueras, O. (2004). Proportional reasoning of quasi-literate adults. In M. J. Høines, & A. B. Fuglestad (Eds.), *28th Annual Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 9–16). Bergen, Norway: PME.

- Arıcan, M. (2018). Preservice middle and high school mathematics teachers' strategies when solving proportion problems. *International Journal of Science and Mathematics Education*, 16(2), 315–335. <https://doi.org/10.1007/s10763-016-9775-1>
- Arıcan, M. (2019). A diagnostic assessment to middle school students' proportional reasoning. *Turkish Journal of Education*, 8(4), 237–257. <https://doi.org/10.19128/turje.522839>
- Arıcan, M. (2021). The development and application of an interview structure on determining preservice mathematics teachers' competence in proportional reasoning. *Mathematics Education Research Journal*. Advance online publication. <https://doi.org/10.1007/s13394-021-00388-5>
- Arıcan, M., & Kuzu, O. (2020). Diagnosing preservice teachers' understanding of statistics and probability: Developing a test for cognitive assessment. *International Journal of Science and Mathematics Education*, 18(4), 771–790. <https://doi.org/10.1007/s10763-019-09985-0>
- Atabaş, S., & Öner, D. (2016). An examination of Turkish middle school students' proportional reasoning. *Boğaziçi University Journal of Education*, 33(1), 63–85. <https://dergipark.org.tr/en/pub/buje/issue/29688/319298>
- Bayazıt, I. (2013). Quality of the tasks in the new Turkish elementary mathematics textbooks: The case of proportional reasoning. *International Journal of Science and Mathematics Education*, 11(3), 651–682. <https://doi.org/10.1007/s10763-012-9358-8>
- Blaikie, N. (2003). *Analyzing quantitative data: From description to explanation*. London, UK: Sage.
- Bokhove, C. (2022). Are instructional practices different between east and west? An analysis of grade 8 TIMSS 2019 data. *Asian Journal for Mathematics Education*, 1(2), 221–241. <https://doi.org/10.1177/27527263221109752>
- Bradburn, M. B., & Gilford, D. M. (1990). *A framework and principles for international comparative studies in education*. Washington, DC: National Academic Press.
- Bradshaw, L., Izsak, A., Templin, J., & Jacobson, E. (2014). Diagnosing teachers' understandings of rational numbers: Building a multidimensional test within the diagnostic classification framework. *Educational Measurement: Issues and Practice*, 33(1), 2–14. <https://doi.org/10.1111/emip.12020>
- Bradshaw, L., & Templin, J. (2014). Combining item response theory and diagnostic classification models: A psychometric model for scaling ability and diagnosing misconceptions. *Psychometrika*, 79(3), 403–425. <https://doi.org/10.1007/s11336-013-9350-4>
- Cai, J. (1995). A cognitive analysis of U.S. and Chinese students' mathematical performance on tasks involving computation, simple problem solving, and complex problem solving. In *Journal for Research in Mathematics Education Monograph Series* (Vol. 7, pp. i–151). Reston, VA: National Council of Teachers of Mathematics.
- Cai, J., Mok, I., Reedy, V., & Stacey, K. (2016). *International comparative studies in mathematics: Lessons for improving students' learning*. New York, NY: Springer.
- Chin, E., & Lin, F. (2009). A comparative study on junior high school students' proof conceptions in algebra between Taiwan and the UK. *Journal of Mathematics Education*, 2(2), 52–67.
- Cramer, K., & Post, T. (1993). Making connections: A case for proportionality. *Arithmetic Teacher*, 60(6), 342–346. <https://doi.org/10.5951/AT.40.6.0342>
- de Bock, D., Neyens, D., & Van Dooren, W. (2017). Students' ability to connect function properties to different types of elementary functions: An empirical study on the role of external representations. *International Journal of Science and Mathematics Education*, 15(5), 939–955. <https://doi.org/10.1007/s10763-016-9724-z>
- de la Torre, J. (2008). An empirically based method of Q-matrix validation for the DINA model: Development and applications. *Journal of Educational Measurement*, 45(4), 343–362. <https://doi.org/10.1111/j.1745-3984.2008.00069.x>
- Dogan, E., & Tatsuoka, K. (2008). An international comparison using a diagnostic testing model: Turkish students' profile of mathematical skills on TIMSS-R. *Educational Studies in Mathematics*, 68(3), 263–272. <https://doi.org/10.1007/s10649-007-9099-8>

- Fernández, C., Llinares, S., Van Dooren, W., De Bock, D., & Verschaffel, L. (2012). The development of students' use of additive and proportional methods along primary and secondary school. *European Journal of Psychology of Education, 27*(3), 421–438. <https://doi.org/10.1007/s10212-011-0087-0>
- Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Dordrecht, The Netherlands: D. Reidel Publishing Company.
- García, F. J., Gascón, J., Ruiz-Higueras, L., & Bosch, M. (2006). Mathematical modelling as a tool for the connection of school mathematics. *ZDM—The International Journal on Mathematics Education, 38*(3), 226–246. <https://doi.org/10.1007/BF02652807>
- Henson, R., Roussos, L., Douglas, J., & He, X. (2008). Cognitive diagnostic attribute-level discrimination indices. *Applied Psychological Measurement, 32*(4), 275–288. <https://doi.org/10.1177/0146621607302478>
- Henson, R., Templin, J., & Willse, J. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika, 74*(2), 191–210. <https://doi.org/10.1007/s11336-008-9089-5>
- Hilton, A., Hilton, G., Dole, S., & Goos, M. (2013). Development and application of a two-tier diagnostic instrument to assess middle-years students' proportional reasoning. *Mathematics Education Research Journal, 25*(4), 523–545. <https://doi.org/10.1007/s13394-013-0083-6>
- Hino, K., & Kato, H. (2019). Teaching whole-number multiplication to promote children's proportional reasoning: A practice-based perspective from Japan. *ZDM—Mathematics Education, 51*(1), 125–137. <https://doi.org/10.1007/s11858-018-0993-6>
- I, J. Y., Martinez, R., & Dougherty, B. (2020). Misconceptions on part-part-whole proportional relationships using proportional division problems. *Investigations in Mathematics Learning, 12*(2), 67–81. <https://doi.org/10.1080/19477503.2018.1548222>
- Incikabi, L., & Tjoe, H. (2013). A comparative analysis of ratio and proportion problems in Turkish and the US middle school mathematics textbooks. *Journal of Kırsehir Education Faculty, 14*(1), 1–15. <https://dergipark.org.tr/en/pub/kefad/issue/59473/854629>
- Jiang, R., Li, X., Fernández, C., & Fu, X. (2017). Students' performance on missing-value word problems: A cross-national developmental study. *European Journal of Psychology of Education, 32*(4), 551–570. <https://doi.org/10.1007/s10212-016-0322-9>
- Karplus, R., Pulos, S., & Stage, E. K. (1983). Proportional reasoning of early adolescents. In R. Lesh, & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 45–90). New York, NY: Academic Press.
- Lamon, S. J. (1993). Ratio and proportion: Children's cognitive and metacognitive processes. In T. P. Carpenter, E. Fennema, & T. A. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 131–156). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Lamon, S. J. (2007). Rational numbers and proportional reasoning: Toward a theoretical framework. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 629–668). Charlotte, NC: National Council of Teachers of Mathematics-Information Age Publishing.
- Lesh, R., Post, T., & Behr, M. (1988). Proportional reasoning. In J. Hiebert, & M. Behr (Eds.), *Number concepts and operations for the middle grades* (pp. 93–118). Reston, VA: National Council of Teachers of Mathematics.
- Levain, J. P., & Vergnaud, G. (1995). Proportionnalité simple, proportionnalité multiple [Simple proportionality, multiple proportionality]. *Grand N, 56*, 55–67.
- Lim, K. (2009). Burning the candle at just one end: Using nonproportional examples helps students determine when proportional strategies apply. *Mathematics Teaching in the Middle School, 14*(8), 492–500. <https://doi.org/10.5951/MTMS.14.8.0492>
- Lo, J. J., Cai, J., & Watanabe, T. (2001). A comparative study of the selected textbooks from China, Japan, Taiwan, and the United States on the teaching of ratio and proportion concepts. In R. Speiser, C. A. Maher, & C. N. Walter (Eds.), *Proceedings of the Twenty-Third Annual Meeting of the North American Chapter of the International Group of the Psychology of Mathematics Education* (pp. 509–520). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

- Lobato, J., Ellis, A., & Zbiek, R. M. (2010). *Developing essential understanding of ratios, proportions, and proportional reasoning for teaching mathematics: Grades 6-8*. Reston, VA: National Council of Teachers of Mathematics.
- Lundberg, A. (2011). Proportion in mathematics textbooks in upper secondary school. In M. Pytlak, T. Rowland, & E. Swoboda (Eds.), *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education* (pp. 336–345). Rzeszów, Polonia: University of Rzeszów.
- Manzon, M. (2011). *Comparative education: The construction of a field*. Hong Kong SAR, China: CERC/Springer.
- Martínez-Juste, S., Muñoz-Escolano, J. M., & Oller-Marcén, A. M. (2015). Un estudio comparativo sobre la proporcionalidad compuesta en libros de texto españoles de Educación Secundaria Obligatoria durante la LOGSE-LOE-LOMCE [A comparative study about compound proportionality in Spanish secondary textbooks under LOGSE-LOE-LOMCE]. *Avances de Investigación en Educación Matemática*, 8, 95–115. <https://doi.org/10.35763/aiem.v1i8.119>
- Martínez-Juste, S., Muñoz-Escolano, J. M., & Oller-Marcén, A. M. (2019). Una experiencia de Investigación-Acción para la enseñanza de la proporcionalidad compuesta [An action research experience to teach compound proportional situations]. *Enseñanza de las Ciencias*, 37(2), 85–106. <https://doi.org/10.5565/rev/ensciencias.2603>
- Martínez-Juste, S., Muñoz-Escolano, J. M., Oller-Marcén, A. M., & Ortega del Rincón, T. (2017). Análisis de problemas de proporcionalidad compuesta en libros de texto de 2º de ESO [Analysis of compound proportion problems in 8th grade textbooks]. *Revista Latinoamericana de Investigación en Matemática Educativa*, 20(1), 95–122. <https://doi.org/10.12802/relime.17.2014>
- Modestou, M., Elia, I., Gagatsis, A., & Spanoudis, G. (2008). Behind the scenes of pseudo-proportionality. *International Journal of Mathematical Education in Science and Technology*, 39(3), 313–324. <https://doi.org/10.1080/00207390701691541>
- Modestou, M., & Gagatsis, A. (2007). Students' improper proportional reasoning: A result of the epistemological obstacle of "linearity". *Educational Psychology*, 27(1), 75–92. <https://doi.org/10.1080/01443410601061462>
- Muthen, L. K., & Muthen, B. O. (2011). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthen & Muthen.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- Neuman, W. L. (2007). *Basics of social research*. London, UK: Pearson.
- OECD (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris, France: OECD. <https://doi.org/10.1787/9789264006416-en>
- OECD (2019). *PISA 2018 results (Volume I): What students know and can do*. Paris, France: OECD. <https://doi.org/10.1787/5f07c754-en>
- Oller-Marcén, A. M., & Gairín, J. M. (2015). Proportionality problems in some mathematical texts prior to fourteenth century. In K. Krainer, & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (pp. 1859–1865). Prague, Czech Republic: Charles University and ERME.
- Orrill, C. H., Brown, R. E., Burke, J. P., Millett, J., Nagar, G. G., Park, J., & Weiland, T. (2017). *Extending appropriateness: Further exploration of teachers' knowledge resources for proportional reasoning* [Paper presentation]. Thirty-Ninth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (PME-NA), Indianapolis, Indiana, United States.
- Ponte, J. P., & Marques, S. (2011). Proportion in school mathematics textbooks: A comparative study. *RIPEM - International Journal for Research in Mathematics Education*, 1(1), 36–53. <http://hdl.handle.net/10451/4222>
- Postlethwaite, N. (1988). *The encyclopedia of comparative education and national systems of education*. Oxford, UK: Pergamon.
- Rupp, A., Templin, J., & Henson, R. (2010). *Diagnostic measurement: Theory, methods, and applications*. New York, NY: Guilford.

- Rupp, A. A., & Templin, J. L. (2008). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement*, 6(4), 219–262. <https://doi.org/10.1080/15366360802490866>
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75(2), 243–248. <https://doi.org/10.1007/s11336-009-9135-y>
- Schubring, G. (1987). On the methodology of analysing historical textbooks: Lacroix as textbook author. *For the Learning of Mathematics*, 7(3), 41–51. <https://www.jstor.org/stable/40247906>
- Singh, P. (2000a). Understanding the concepts of proportion and ratio among grade nine students in Malaysia. *International Journal of Mathematical Education in Science and Technology*, 31(4), 579–599. <https://doi.org/10.1080/002073900412688>
- Singh, P. (2000b). Understanding the concepts of proportion and ratio constructed by two grade six students. *Educational Studies in Mathematics*, 43(3), 271–292. <https://doi.org/10.1023/A:1011976904850>
- Spanish Ministry of Education, Culture and Sports. (2015). *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato* [Royal decree 1105/2014, from December 26, which establishes the basic curriculum for Compulsory Secondary Education and High School]. [www.boe.es/eli/es/rd/2014/12/26/1105/con](http://www.boe.es/eli/es/rd/2014/12/26/1105/con)
- Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist*, 35(2), 87–100. [https://doi.org/10.1207/S15326985EP3502\\_3](https://doi.org/10.1207/S15326985EP3502_3)
- Şen, S., & Arıcan, M. (2015). A diagnostic comparison of Turkish and Korean students' mathematics performances on the TIMSS 2011 assessment. *Journal of Measurement and Evaluation in Education and Psychology*, 6(2), 238–253.
- Tatsuoka, C., Clements, D. H., Sarama, J., Izsák, A., Orrill, C. H., de la Torre, J., & Tatsuoka, K. K. (2016). Developing workable attributes for psychometric models based on the Q-Matrix. In A. Izsák, J. T. Remillard, & J. Templin (Eds.), *Psychometric methods in mathematics education: Opportunities, challenges, and interdisciplinary collaborations (Monograph #15)* (pp. 73–96). Reston, VA: National Council of Teachers of Mathematics.
- Tatsuoka, K. (1985). A probabilistic model for diagnosing misconceptions by the pattern classification approach. *Journal of Educational Statistics*, 10(1), 55–73. <https://doi.org/10.3102/10769986010001055>
- Tjoe, H., & de la Torre, J. (2014). The identification and validation process of proportional reasoning attributes: An application of a cognitive diagnosis modeling framework. *Mathematics Education Research Journal*, 26(2), 237–255. <https://doi.org/10.1007/s13394-013-0090-7>
- Turkish Ministry of National Education. (2018). *Matematik dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar)* [Mathematics curriculum (primary and secondary school, grades 1, 2, 3, 4, 5, 6, 7 and 8)]. [mufredat.meb.gov.tr/ProgramDetay.aspx?PID=329](http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=329)
- Valverde Soto, A. G., & Castro Martínez, E. (2009). Actuaciones de maestros en formación en la resolución de problemas de proporcionalidad directa [Behaviour of pre-service primary education teachers when solving proportionality problems]. In M. J. González, M. T. González, & J. Murillo (Eds.), *Investigación en educación matemática XIII* (pp. 523–531). Santander, Spain: SEIEM.
- Van Dooren, W., De Bock, D., Evers, M., & Verschaffel, L. (2009). Students' overuse of proportionality on missing-value problems: How numbers may change solutions. *Journal for Research in Mathematics Education*, 40(2), 187–211. <https://doi.org/10.2307/40539331>
- Van Dooren, W., De Bock, D., Hessels, A., Janssens, D., & Verschaffel, L. (2005). Not everything is proportional: Effects of age and problem type on propensities for overgeneralization. *Cognition and Instruction*, 23(1), 57–86. [https://doi.org/10.1207/s1532690xci2301\\_3](https://doi.org/10.1207/s1532690xci2301_3)
- Van Dooren, W., De Bock, D., Janssens, D., & Verschaffel, L. (2008). The linear imperative: An inventory and conceptual analysis of students' over-use of linearity. *Journal for Research in Mathematics Education*, 39(3), 311–342. <https://doi.org/10.2307/30034972>
- Vergnaud, G. (1983). Multiplicative structures. In R. Lesh, & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 127–174). New York, NY: Academic Press.

- Von Davier, M. (2005). A general diagnostic model applied to language testing data. In *Educational Testing Service, Research Report, RR-05-16* (pp. 1–39). Princeton, NJ: Educational Testing Service.
- Werner, C., & Schermelleh-Engel, K. (2010). Deciding between competing models: Chi-square difference tests. In *Introduction to Structural Equation Modeling with LISREL* (pp. 1–3). Frankfurt, Germany: Goethe University.
- Wright, V. (2014). Frequencies as proportions: Using a teaching model based on Pirie and Kieren's model of mathematical understanding. *Mathematics Education Research Journal*, 26(1), 101–128. <https://doi.org/10.1007/s13394-014-0118-7>

### Author biographies

**Sergio Martínez-Juste** is an assistant professor of mathematics education at University of Zaragoza, Spain. He has worked as a middle and high school teacher. His main research interest is middle school students' proportional reasoning.

**Muhammet Arıcan** is an associate professor of mathematics education at the Boğaziçi University, Turkey. He conducts studies on students' and preservice teachers' proportional reasoning and diagnostic classification models.

**José M. Muñoz-Escolano** is an associate professor of mathematics education at University of Zaragoza, Spain. His main research interests are the history of mathematics education and middle school students' proportional reasoning.

**Antonio M. Oller-Marcén** is an associate professor at Centro Universitario de la Defensa de Zaragoza, Spain. He has a PhD in mathematics education from Universidad de Valladolid, Spain. His main research interests are the history of mathematics education and its use as a tool, and several aspects of proportional reasoning.

### Appendix I. Proportional reasoning test [Retrieved from Arıcan (2019), pp. 253–254].

- (1) Two numbers are in a ratio of  $\frac{3}{5}$ . If small number is 45, what is the value of large number?  
(A) 27      (B) 55      (C) 75      (D) 85
- (2) There are some numbers of red and blue balls in a container. The ratio of the number of red balls to the number of blue balls is  $\frac{3}{4}$ . If there are a total of 42 balls in this container, how many of them are blue?  
(A) 6      (B) 12      (C) 18      (D) 24
- (3) The measures of outer angles of a triangle are proportional to the numbers 9, 12, and 15. What is the measure of the greatest internal angle of this triangle?  
(A) 45      (B) 60      (C) 75      (D) 90
- (4) A teacher gives her students 2 candies for every 5 questions they answer correctly. If a student received a total of 14 candies, how many correct answers is this student had?  
(A) 28      (B) 35      (C) 45      (D) 70

(5) Two different but identical candles, A and B, are burning at the same constant rate but they are lit at different times. We know that when candle A has burned 24 mm, candle B has burned 18 mm. When 36 mm of candle B has burned, how many millimeters will candle A have burned?

- (A) 48      (B) 44      (C) 42      (D) 40

(6) Lemonade A is made by mixing 4 cups of water with 2 tablespoons of sugar, Lemonade B is made by mixing 7 cups of water with 5 tablespoons of sugar, and Lemonade C is made by mixing 3 cups of water with 1 tablespoon of sugar. Which one of the following information is true about the flavors of these lemonades?

- (A) The least taste of sugar is taken from Lemonade A.  
(B) The most taste of sugar is taken from Lemonade B.  
(C) The same taste of sugar is taken from Lemonade A and Lemonade C.  
(D) The most taste of sugar is taken from Lemonade C.

(7) An athlete, who runs at a constant speed, consumes 150 calories in 2 h. If this athlete keeps running at this same constant speed, how many calories does he consume after 3 h?

- (A) 225      (B) 200      (C) 175      (D) 120

(8) Three painters working at the same pace can paint a building in 12 days. In how many days 4 painters can paint the same building?

- (A) 16      (B) 8      (C) 9      (D) 10

(9) A car travels the distance between two cities in 6 h driving at a constant speed of 60 km/h. If this car increases its speed to 90 km/h, how many hours can it take between the two cities?

- (A) 3      (B) 4      (C) 6      (D) 9

(10) Twelve workers working at the same pace can finish a certain job in 15 days working 7 h each day. How many days are needed for 14 workers to finish the same job working 5 h each day?

- (A) 10      (B) 12      (C) 16      (D) 18

(11)  $x$  and  $y$  are two inversely proportional quantities. We know that when  $x$  is 12,  $y$  is 4. When  $x$  is 6, what is the value of  $y$ ?

- (A) 8      (B) 6      (C) 4      (D) 2

(12) Mehmet bought a 90 cm long plant from the market. This plant grows 30 cm each year. After 3 years, the height of the plant was measured as 180 cm. According to this information, how many centimeters should be the height of the plant after 6 years?

- (A) 360      (B) 300      (C) 270      (D) 240

(13)  $a$  and  $b$  two numbers that vary in a directly proportional relationship. We know that when  $a$  is 5,  $b$  is 15. So, when  $b$  is 24, what is  $a$ ?

- (A) 8      (B) 12      (C) 14      (D) 16



(14) Ayse and Esra, who are running at the same speed, started running around an 800 m circular track at different times. It is known that when Ayse ran 300 m, Esra ran 200 m. According to this information, when Esra ran 400 m, how many meters can Ayse run?

- (A) 500      (B) 600      (C) 700      (D) 800

(15) Three painters working at the same pace can paint a building in 8 days. Which one of the following tables depicts the relationship between the number of painters and painting time correctly?

(A)

Number of painters	1	2	3	4
Painting time	2	4	8	16

(B)

Number of painters	1	2	3	4
Painting time	24	16	8	4

(C)

Number of painters	1	2	3	4
Painting time	16	12	8	6

(D)

Number of painters	1	2	3	4
Painting time	24	12	8	6

(16) The table below shows the relationship between the number of chickens on a farm and the average number of eggs they produce per week. According to this table, how many chickens are required to have 240 eggs per week?

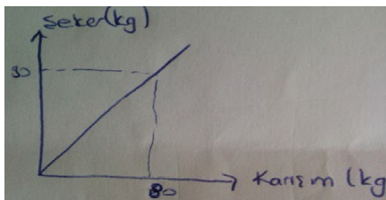
- (A) 120      (B) 100  
(C) 80      (D) 60

Number of chickens	Number of eggs per week
4	12
6	18
9	27
12	36

(17)  $(m + 3)$  and  $(n + 1)$  are two numbers that vary in an inversely proportional relationship. We know that when  $m$  is 1,  $n$  is 2. When  $m$  is 3, what is  $n$ ?

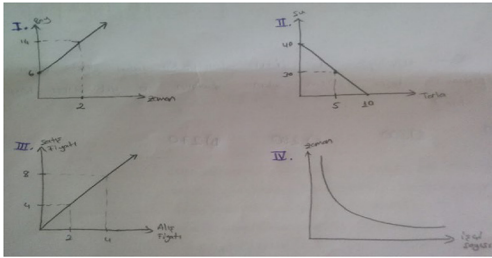
- (A) 1      (B) 2      (C) 3      (D) 4

(18) The graph below shows that 30 kg sugar is present in an 80 kg mixture. How many kilograms sugar are there in a 120 kg mixture?



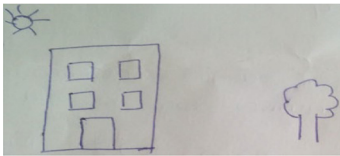
- (A) 45      (B) 40      (C) 35      (D) 20

(19) Which one(s) of the above graphs do not depict a directly proportional relationship?



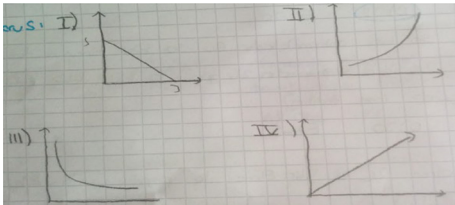
- (A) Only IV      (B) Only II      (C) II and IV      (D) I, II, and IV

(20) At a certain time of a day, the length of the shadow of a 60 m long apartment is measured as 12 m. At the same certain time, what is the length of the shadow of a 5 m long tree?



- (A) 4      (B) 3      (C) 2      (D) 1

(21) Which one(s) of the above graphs depict an inversely proportional relationship?



- (A) Only I      (B) Only III      (C) I and III      (D) I, II, and IV

(22) In a bakery, 3 people working at the same constant rate can frost 12 cakes in 4 h. In this bakery, how many hours are needed by 4 people to frost 8 cakes?

- (A) 1      (B) 2      (C) 3      (D) 4

## Appendix II. The percentages of Turkish (N = 282) and Spanish (N = 314) students' responses to the proportional reasoning test.

Item	Turkish students' responses					Spanish students' responses				
	A	B	C	D	NA	A	B	C	D	NA
1	13.83%	0.71%	84.04%	0.71%	0.71%	9.55%	8.60%	71.34%	7.32%	3.18%
2	10.99%	8.51%	7.09%	71.99%	1.42%	6.05%	49.04%	16.88%	21.97%	6.05%
3	17.02%	8.16%	22.70%	45.39%	6.74%	23.57%	19.43%	21.34%	23.57%	12.10%
4	2.48%	86.88%	0.71%	9.22%	0.71%	6.05%	70.06%	4.46%	18.79%	0.64%
5	65.25%	2.13%	25.89%	2.84%	3.90%	36.62%	3.82%	55.41%	3.18%	0.96%
6	8.16%	60.99%	9.22%	19.86%	1.77%	4.14%	60.51%	19.43%	13.69%	2.23%
7	92.91%	1.77%	2.84%	2.48%	0.00%	91.72%	2.55%	2.23%	1.59%	1.91%
8	27.30%	10.64%	60.99%	0.71%	0.35%	17.83%	24.84%	49.04%	7.01%	1.27%
9	8.87%	48.94%	1.42%	39.72%	1.06%	20.38%	60.19%	4.14%	13.06%	2.23%
10	7.09%	23.40%	6.74%	45.04%	17.73%	10.19%	29.62%	18.47%	36.94%	4.78%
11	60.28%	3.19%	6.03%	29.79%	0.71%	30.57%	5.10%	4.14%	55.73%	4.46%
12	19.15%	5.67%	66.31%	7.80%	1.06%	27.71%	4.46%	63.06%	3.50%	1.27%
13	81.56%	2.48%	11.35%	3.55%	1.06%	62.42%	6.05%	25.80%	3.50%	2.23%
14	24.47%	65.96%	3.19%	4.61%	1.77%	55.41%	38.22%	2.23%	2.55%	1.59%
15	11.35%	17.73%	10.28%	57.80%	2.84%	10.19%	30.25%	14.97%	43.63%	0.96%
16	7.45%	5.32%	73.40%	7.45%	6.38%	4.46%	4.14%	84.71%	5.10%	1.59%
17	39.01%	25.53%	7.09%	21.63%	6.74%	19.11%	13.06%	9.87%	50.96%	7.01%
18	62.77%	11.70%	4.96%	15.60%	4.96%	72.29%	16.24%	2.87%	4.46%	4.14%
19	17.38%	11.70%	31.21%	37.23%	2.48%	26.43%	7.64%	28.03%	35.99%	2.87%
20	6.38%	8.16%	6.03%	76.24%	3.19%	6.69%	8.92%	11.46%	70.06%	2.87%
21	14.54%	55.32%	17.73%	7.80%	4.61%	17.20%	36.94%	33.76%	9.24%	2.87%
22	10.28%	56.03%	10.28%	7.45%	15.96%	6.37%	64.01%	19.43%	7.64%	2.55%

Notes. NA=no answer; The boxes in gray indicate the correct option in each item.