

# TEACHERS' CONCEPTIONS AND PROFESSIONAL KNOWLEDGE OF VARIABILITY FROM THEIR INTERPRETATION OF HISTOGRAMS: THE CASE OF VENEZUELAN IN-SERVICE SECONDARY MATHEMATICS TEACHERS

ORLANDO GONZÁLEZ  
Assumption University, Thailand  
ogonzalez@au.edu

## ABSTRACT

*Many studies have reported on the influence of teachers' conceptions of variability on different aspects of their professional knowledge for teaching statistics and their classroom practices. However, research on conceptions of variability is scarce, particularly in Latin American countries like Venezuela. In an effort to help fill this gap, a qualitative study was conducted that aimed to characterize the different ways in which Venezuelan in-service secondary school mathematics teachers conceptualize variability. For that purpose, a survey instrument was developed and administered to 27 teachers working at the metropolitan area of Caracas. This paper focuses on the participants' answers to two items in which interpretation of histograms was necessary. It was found that about a third of the participants exhibited a sophisticated recognition of variability (e.g., gave answers connecting both middles and extremes), whereas about half of them exhibited misconceptions of variability, such as acknowledging variability from the viewpoint of idiosyncratic ideas, or the degree of symmetry (or lack thereof) of a histogram. Moreover, it was also found that about two-thirds of the participants were unable to correctly match real-life contexts to their corresponding histograms, while about two-fifths were unable to correctly determine the accuracy or inaccuracy of descriptions of the variability in a histogram. The author discusses possible reasons for the results obtained, in order to identify relevant implications for teacher education in the area of statistics.*

**Keywords:** *Statistics education research; Statistical literacy; Specialized content knowledge; Teachers' professional competencies; Variability*

## 1. INTRODUCTION

The teaching of statistics at school level is fundamental to promote statistical literacy during the stages of childhood and adolescence, especially in secondary education. For many students, exposure to statistics at secondary school could be the last and only formal program of this discipline they will take to prepare themselves as future users of statistics. Moreover, for those students wishing to move on to university education, prior exposure to statistics at secondary school would prepare them to understand advanced versions of the content studied at that level.

Curricular reforms recently carried out in different countries have increased the number of topics related to statistics to be studied at secondary level (for detailed reviews on this subject see Batanero et al., 2011, pp. 1–45; Pfannkuch, 2018, pp. 387–413). In the particular case of Venezuela, topics on statistics and probability were introduced in the first (basic) cycle of secondary education (i.e., from seventh to ninth grade) in 1985, while their introduction in the upper (diversified) cycle of secondary education (i.e., from tenth to eleventh grade) occurred in 1972 (Ministry of Education [ME], 1997; National Center for the Improvement of the Teaching of Science and Mathematics [CENAMEC], 1990).

In Venezuela, as well as the rest of the world, teachers responsible for teaching statistics at the secondary level of schooling must teach various content related to variability (i.e., the propensity of a statistical object to vary or change, considered the *raison d'être* of statistics as a discipline). Such work demands from them specific professional competencies, which basically depend on their professional

knowledge and conceptions of the curricular contents (Ball et al., 2008; Döhrmann et al., 2012; Gal, 2004; González, 2014a).

For any mathematics teacher, the specialized pedagogical content knowledge (i.e., a type of professional knowledge specific to the work of teaching, difficult to be articulated by other mathematically trained professionals who do not dedicate to teaching) is what distinguishes him or her from any other user of mathematics in their professional work, especially in the ability to determine the degree of correctness of all types of answers given by students (Ball et al., 2008). There is evidence worldwide (although still scarce in Latin American countries) that specialized content knowledge for teaching statistics is influenced by teachers' conceptions of variability, particularly in the context of evaluation of histograms (e.g., Bruno & Espinel, 2009; Peters, 2009; González, 2014a, 2014b, 2014c; Isoda et al., 2018).

The research scarcity on this issue in Latin American countries is particularly true in Venezuela, where the few research studies in statistics education reported prior to the publication of this article focused on the statistical content included in the school curriculum (e.g., Salcedo, 2006) or in students' knowledge about statistics and probability (e.g., León, 2011; Tapia, 2011), with no reported studies on teachers' professional competencies for teaching statistics at any school level, with the exception of the doctoral dissertation and a previous article by the author of this article (González, 2014a, 2014b).

Consequently, the following research questions were posed, in order to clarify aspects of the professional knowledge and conceptions of variability held by Venezuelan secondary school mathematics teachers:

1. How do Venezuelan senior high school mathematics teachers conceptualize variability, and how appropriate or inappropriate are those conceptions?
2. What specific knowledge base do Venezuelan senior high school mathematics teachers need to ensure they are able to appropriately conceptualize variability?

## 2. LITERATURE REVIEW

Considering "knowledge" as the set of cognitive skills and mental representations of reality which, in tandem, comprise the implicit structure behind the actions of a person (Sfard, 1991), two ideas related to the discussion of teachers' professional knowledge for the teaching of variability are reviewed in this section: conceptions of variability and teacher competencies for teaching statistics.

### 2.1. CONCEPTIONS OF VARIABILITY

Conceptions can be understood as lenses through which people perceive and interpret phenomena (Pratt, 1992; Sfard, 1991). In particular, it has been reported that teachers' conceptions of the subject matter influence their approaches to teaching and, consequently, their students' approaches to learning (Kuntze & Friesen, 2018; Sivunen & Pehkonen, 2009).

The conceptions of variability held by students at any school level have been exhaustively explored in the literature (e.g., Shaughnessy, 2007). Some research studies on the conceptions of variability held by in-service teachers have been also conducted, although to a lesser extent (e.g., Biehler et al., 2018; González, 2014a, 2014c; Isoda et al., 2018; Makar & Confrey, 2004; Peters, 2009).

González (2014a) identified the following nine ways in which teachers conceptualize variability, based on his work with 53 Venezuelan in-service secondary school mathematics teachers in different statistical contexts, such as interpretation of histograms, comparison of distributions, and probabilistic and statistical experiments. Such work was used as evidence to validate and refine the conceptual framework developed by Shaughnessy (2007, pp. 984–985), in which eight ways of conceptualizing variability were identified. As a result, the following classification was obtained (the interested reader can refer to González [2014a] for additional examples and a more detailed explanation of this classification).

***Variability in particular values.*** This conception is characterized by the use of particular data values in a data set (such as centers, outliers or extremes) as variability indicators.

**Variability as change over time.** In this conception, limited to graphs in which time is the independent variable, people focus their attention on finding an overall pattern in the data, as well as on looking for causes why the data might be varying in the way they are.

**Variability as whole range.** This conception involves the spread of all possible values of a distribution or an entire data set, and it is closely related to the concept of sample space in probability.

**Variability as the likely range of a sample.** This conception is evidenced when people consider the most likely range of values as the outcome of a probabilistic or statistical experiment. This conception is grounded on the concept of relative frequency, and hence relies heavily on proportional reasoning.

**Variability as distance or difference from some fixed point.** This conception is evidenced when an individual quantifies variability by either visually estimating or quantitatively measuring the distance from one or more values in a data set to some measure of central tendency or endpoint value.

**Variability as the sum of residuals.** This conception is evidenced when variability is measured through deviation-based metrics, such as the mean absolute deviation, the sum of residuals, or the averages of the absolute value differences from a measure of central tendency or some fixed value.

**Variation as covariation or association.** This conception involves consideration of the interaction of several variables, and how changes in one variable may correspond to changes in one or more other variables.

**Variation as distribution.** This conception is evidenced by the use of theoretical properties of a distribution or data set (e.g., central tendency or skewness), or theoretical probability distributions, to estimate or quantify variability.

**Variability as visual cues in the graph.** This conception is evidenced when an individual thinks of variability in terms of symmetry or degree of fit (or lack thereof) to a normal distribution.

## 2.2. TEACHER COMPETENCIES FOR TEACHING STATISTICS

It is common to find in literature conceptual models of teacher competencies (i.e., the core set of a teacher's skills, abilities and traits, which determines both the approach and outcomes of his or her pedagogical interventions) in which these are considered to be comprised of a cognitive and an affective-motivational facet (e.g., Blömeke & Delaney, 2012; Döhrmann et al., 2012; Tatto et al., 2012). The present study also takes the same standpoint on the conceptual modeling of teacher competencies for teaching statistics; particularly, the model proposed by González (2014a). This framework shows the convergence of different ideas, as a result of adapting, for the particular case of statistics education, some previous research efforts to model teacher competencies for teaching mathematics (e.g., Döhrmann et al., 2012; Gal, 2004; Tatto et al., 2012).

According to González's (2014a) framework (Figure 1), the cognitive facet of teachers' professional competencies for teaching statistics (i.e., their professional knowledge regarding statistics instruction, or statistical knowledge for teaching, hereafter SKT) is comprised of the two domains of mathematical knowledge for teaching (MKT), as proposed by Ball et al. (2008): *subject matter knowledge* (SMK) and *pedagogical content knowledge* (PCK). Each of these domains is divided into three different subdomains. SMK is comprised of *common content knowledge* (CCK, a solid knowledge about the statistical ideas included in the mathematics curriculum and the ability to perform tasks related to statistical literacy over such ideas), *specialized content knowledge* (SCK, the ability to determine the degree of correctness or incorrectness of common and/or non-standard solutions that could be given by students when solving statistical problems), and *horizon content knowledge* (HCK, the teacher's understanding of the broader set of statistical ideas to which a particular concept connects). PCK is comprised of *knowledge of content and students* (KCS, the teacher's knowledge about how students think of the statistical ideas being taught in the mathematics curriculum), *knowledge of content and teaching* (KCT, the teacher's capacity to plan and execute meaningful teaching of statistical ideas in

the light of students' previous cognitive structures and traits), and *knowledge of content and curriculum* (KCC, the teacher's knowledge about how the statistical ideas that must be taught at a particular grade level are developed throughout the entire mathematics curriculum).

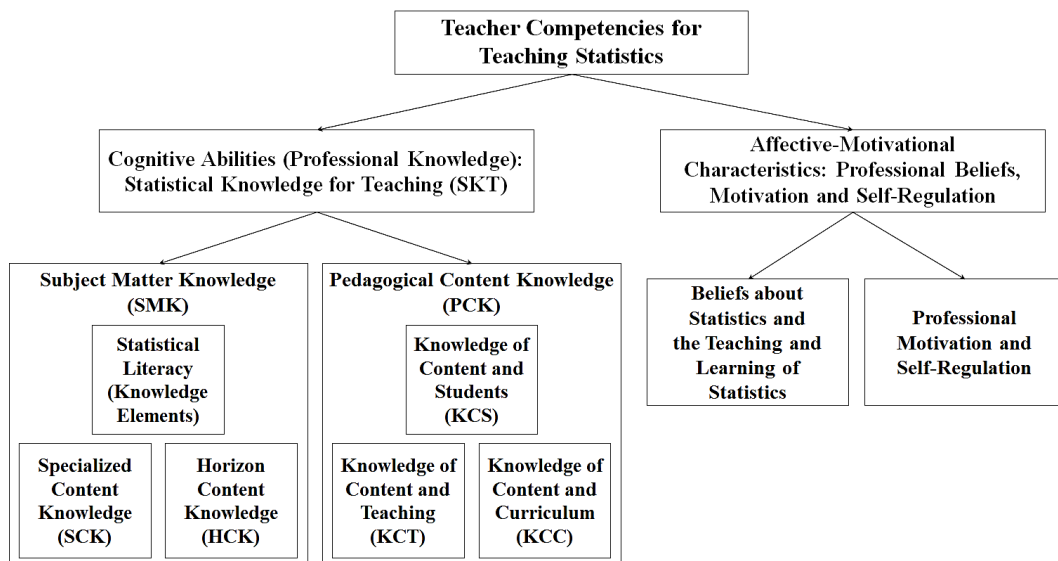


Figure 1. Conceptual model of the teachers' professional competencies for teaching statistics, according to González (2014a)

In the conceptual framework adopted by this study, CCK is understood as the knowledge elements of statistical literacy, in the sense of Gal (2004). This is because statistical literacy is associated with the statistics-related cognitive skills expected from any individual after completing compulsory schooling. The remaining knowledge components in this framework are defined in the same way as in Ball et al. (2008).

The affective-motivational facet of the model developed by González (2014a) is comprised of two components: (a) teachers' beliefs about statistics, its teaching and learning; and (b) professional motivation and self-regulation. The teachers' beliefs about statistics, its teaching and learning are the psychologically held understandings that teachers feel to be true about the nature of statistics and how it should be taught and learned, which are key for both their perception of classroom situations and their decisions on how to act. Professional motivation is defined as a persisting tendency or inclination, triggered by specific driving forces, to perform the duties associated with the teaching profession and work toward the continued fulfillment of them. The performance of such professional duties is guided by the professional self-regulation, which is the execution of volitional cognitive and behavioral processes by the teachers, that are conducive to positive decision-making, adjustment and adaptation in various situations related to the teaching profession. The conceptions of variability are within the teachers' beliefs about statistics, its teaching and learning, based on the premise adopted in this work that conceptions are "conscious beliefs" (Sivunen & Pehkonen, 2009).

### 3. METHODOLOGY

#### 3.1. PARTICIPANTS, DESIGN AND PROCEDURE

A sample comprised of nine schools (2 public, 7 private), located in the metropolitan area of Caracas, Venezuela, was selected for this study, using purposeful sampling (Merriam & Tisdell, 2016). In those schools, 27 mathematics teachers from secondary level voluntarily agreed to participate in this qualitative, descriptive-interpretative study. All participants were career teachers, with a teaching experience ranging from 1 to 44 years ( $M = 12.6$  years,  $Mdn = 8$  years).

### 3.2. DATA COLLECTION

From May to June 2018, a 10-item anonymous questionnaire was administered to and completed by all the participants. The time given to complete the questionnaire was 60 minutes. The questionnaire was an adaptation to Spanish of the instrument originally used by Isoda et al. (2018) with secondary school mathematics teachers from Japan and Thailand. Item 1 (Figure 2) was modified to prompt respondents to provide the justification for the answers given, and based on the comments from the original developers, Item 2 (Figure 3) was added, which involves evaluating student responses to questions about graphical representations. This study reports on the questionnaire results obtained from the participants for Items 1 and 2, which were chosen for this article because they require the evaluation of histograms.

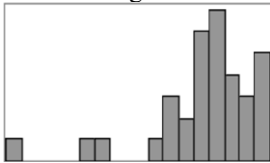
**Item 1: Matching histograms to variables.** The purpose of this item (Figure 2) was to assess surveyed teachers' ability to establish a correspondence between the varying behavior of phenomena framed in a real-life context, and their respective graphical representations, or vice versa. Item 1 was posed in order to elicit teachers' answers and explanations from which a comprehensive picture of their conceptions of variability could emerge.

Match each of the following variables to its corresponding histogram. For each case, please briefly justify your choice.

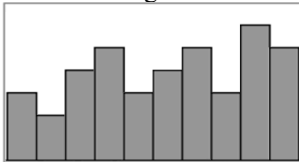
Variable	Corresponding histogram	Justification
Ⓐ Age at death of a random sample of 34 persons		
Ⓑ The last digit in the ID card of each of the 40 students from a randomly chosen group		
Ⓒ Scores on a fairly easy mathematics test		
Ⓓ Height of a random group of adults		
Ⓔ Number of gold medals won by medal-winning countries in the 2008 Beijing Olympic Games		


Histogram 1



Histogram 2

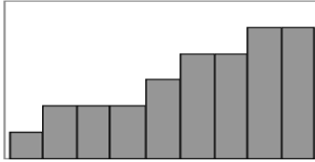


Histogram 3



Histogram 4



Histogram 5

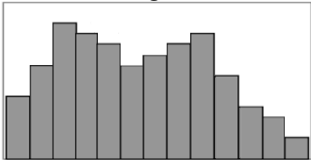


Figure 2. Item 1: Matching histograms to variables.

**Item 2: Choosing the distribution with more variability.** The purpose of this item (González, 2014a) was to identify and assess teachers' understanding of the interpretation of variability in histograms (Figure 3). By answering this item, it was anticipated that the teachers surveyed would show

evidence of their specialized content knowledge (SCK), in a context of assessing the variability of data sets represented by histograms.

Look at the histograms of the following two distributions:

**A**

**B**

Suppose you pose the following question to your students: Which distribution (A or B) do you think has more variability (dispersion), and why do you think so?

Suppose that, after posing this question to your students, three of them come up with the following answers:

Student 1:	“Distribution A has more variability because it’s not symmetrical.”
Student 2:	“Distribution A ranges from 3 to 14, while Distribution B ranges from 1 to 14. Then, Distribution B has more variability.”
Student 3:	“The bars in Distribution A are clumped closer to the central bar than they are in Distribution B. Then, Distribution B has more variability.”

Dealing with each student separately, please comment briefly on each of these answers, focusing on whether the answer is correct or not, why you think so, and what reasoning might have led students to produce each answer.

*Figure 3. Item 2: Choosing the distribution with more variability.*

The answers in Item 2 given by Students 1 and 2 are examples of misconceptions of variability, frequently exhibited by students when comparing histograms (González, 2014a; Meletiou & Lee, 2003). In the case of the answer given by Student 1, it reflects a misconception about the relationship of variability and symmetry, in which histograms with bars distributed in a symmetrical pattern are incorrectly understood as having less variability than those that do not. In the case of the answer provided by Student 2, it reflects a misconception in which variability is measured in terms of the largest span on the vertical axis, instead of considering the spread of data values on the horizontal axis, or measures of distance or difference from a center. The answer given by Student 3 exemplifies an appropriate visual estimation and interpretation of the variability in the given histograms. The student estimated the variability, without making explicit calculations, by assessing the degree of data clustering around the modal class.

### 3.3. DATA ANALYSIS

The author transcribed the given answers into spreadsheets for further analysis (both quantitatively and qualitatively), and then reviewed and coded the collected data. The qualitative coding process was carried out using open and axial coding as data interpretation methods (Merriam & Tisdell, 2016). The open coding stage, guided by this study’s first research question, allowed for the identification of open codes pertaining to the way participants conceptualized variability. During axial coding, open codes were grouped under the nine ways of conceptualizing variability identified by González (2014a), as relationships between open codes and these nine conceptions were determined. Quantitatively, the frequencies and percentages of all answers given to Items 1 and 2 by the participating teachers were

calculated, and the number of correct and incorrect responses for each item were also recorded by the researcher.

#### 4. RESULTS AND ANALYSIS

##### 4.1. RESULTS CORRESPONDING TO ITEM 1

Only six of the participating teachers (22.2%) were able to correctly answer this item by matching appropriately all the variables at hand with the corresponding histograms. Moreover, only 20 out of the 27 teachers in the study (74.1%) answered this item. Table 1 presents the distribution of the matches done on Item 1 by the participants.

*Table 1. Frequencies and percentages of all the matches done on Item 1 by the participating teachers*

Variable	Histogram 1		Histogram 2		Histogram 3		Histogram 4		Histogram 5	
	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
Ⓐ	12	44.4	0	0.0	1	3.7	4	14.8	3	11.1
Ⓑ	3	11.1	<b>14</b>	<b>51.9</b>	0	0.0	2	7.4	1	3.7
Ⓒ	3	11.1	1	3.7	8	29.6	7	25.9	1	3.7
Ⓓ	1	3.7	3	11.1	0	0.0	5	18.5	<b>11</b>	<b>40.7</b>
Ⓔ	1	3.7	2	7.4	<b>11</b>	<b>40.7</b>	2	7.4	4	14.8

The values in bold characters correspond to the correct match.

As it can be seen in Table 1, the best result obtained by the participants on this item was in matching Variable Ⓑ (last digit in the ID card of each of the 40 students from a randomly chosen group) to Histogram 2 (correct response rate of 51.9%). This result could be interpreted as familiarity with the uniform distribution, by recognizing that all the bars of the graph corresponding to Variable Ⓑ should have similar heights, since the probability of occurrence of each number is equiprobable. This provides evidence of holding the conception of variability identified as “variation as distribution”, as reflected by the comments given by some of the participants:

- Teacher 13: The bar heights are slightly varying.  
 Teacher 16: Probabilities are somewhat evenly distributed.  
 Teacher 25: By chance.

Also, it is possible that the participating teachers who did this match correctly were able to determine the sample space for all possible outcomes of Variable Ⓑ (evidence of holding the conception of variability identified as “variability as whole range”), or simply focused on the cardinality of the Variable Ⓑ’ s sample space (evidence of holding the conception of variability identified as “variability in particular values”):

- Teacher 4: Each bar represents one of the digits that the ID card could end with.  
 Teacher 14: There are 10 bars that identify the 10 numbers from 0 to 9 with their distribution.  
 Teacher 15: №. of columns = 10. The numbers of the ID card’s last digit.  
 Teacher 24: It depicts the number of students whose ID card ends with the digit 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.  
 Teacher 26: Because there are 10 bars (0, 9).  
 Teacher 27: Because the last number of the ID card only goes from 0 to 9, and is the only histogram with 10 bars, also it is normally distributed.

In relation to matching Variable Ⓐ (age at death of a random sample of 34 persons) to Histogram 1 (correct response rate of 44.4%), most of the surveyed teachers who did this match correctly could have noted that the histogram of Variable Ⓐ should be skewed to the left, with a heavy right tail and a very light left tail (evidence of holding the conception of variability identified as “variation as distribution”):

- Teacher 1: Within the population I would expect a few deaths in newborns and/or children, and more in adults. The tendency would be for all people to die being elderly, at an old age.
- Teacher 8: The frequencies are larger by the end of the graph.
- Teacher 27: It is natural that people of older age are the ones who die at a greater rate.

In relation to matching Variable ④ (height of a random group of adults) to Histogram 5 (correct response rate of 40.7%), most of the surveyed teachers who made the match correctly seemed to have understood that the histogram of Variable ④ should look like the histogram of a normal distribution, with a more or less bell-like and symmetrical shape, less density in the tails of the distribution, and a higher density in the central region (evidence of holding the conception of variability identified as “variation as distribution”):

- Teacher 1: I would expect a behavior similar to a normal distribution, with a higher density in intermediate heights.
- Teacher 14: The heights are concentrated in the middle of the distribution, with few tall heights.
- Teacher 27: Being a random group the heights should be normally distributed.

None of the respondents mentioned expecting to find a histogram with two peaks when matching Variable ④ to Histogram 5, since the sampling distribution of the height of a random group of male and female adults is bimodal. This is a noteworthy difference between the participants in this study and the Japanese secondary school teachers from the study conducted by Isoda et al. (2018).

In relation to matching Variable ⑤ (number of gold medals won by medal-winning countries in the 2008 Beijing Olympic Games) to Histogram 3 (correct response rate of 40.7%), most of the surveyed teachers who did this match correctly seemed to have understood that the histogram of Variable ⑤ should have its modal class in the first class interval of the frequency distribution, with the most likely range of values located on the left extreme of the distribution (evidence of holding the conception of variability identified as “variability as the likely range of a sample”), and should depict a right-skewed distribution, with a very heavy left tail and a very long and light right tail, with some outliers on the latter (evidence of holding the conception of variability identified as “variation as distribution”):

- Teacher 1: Generally, there are 2 or 3 countries winning the majority of the events in the Olympics, there will be other medal-winning countries ... without winning a gold medal though.
- Teacher 4: Usually gold medals are won by very few countries and among them always few are the ones obtaining the most part of those medals.
- Teacher 13: Because the majority of the medals are obtained by few countries.
- Teacher 14: In the histogram it is depicted one country that obtained a great share of the medals and others that obtained few or none.
- Teacher 16: Gold medals are very difficult to win in relation to the medals won by most of the countries.
- Teacher 25: Because there are countries that win a large number of medals, others very few and others nothing.
- Teacher 27: The number of gold medals for each medal-winning country cannot be that large, although there is always a country standing out, the probabilities should be distributed more or less like that.

The most difficult answer on this item was matching Variable ⑥ (scores on a fairly easy mathematics test) to Histogram 4 (correct response rate of 25.9%). In fact, a higher number of respondents incorrectly matched Variable ⑥ to Histogram 3 (8 out of the 27 participating teachers, 29.6%). This mistake could be the result of confusing histograms with bar graphs, possibly thinking that the first bar in Histogram 3 represented the group that obtained the best result:

- Teacher 2: Because the scores are very variant or very good until arriving to the lowest score.
- Teacher 7: Because a value with an extremely large frequency can be observed.
- Teacher 8: The frequency is larger at the beginning of the graph.
- Teacher 21: A high number of students passed the test.
- Teacher 24: It is shown the number of students according to their score.



In order to correctly match Variable © to Histogram 4, respondents must have understood that the distribution of scores on an easy test should be left-skewed, with most of the data heavily concentrated on the right tail, and less accumulation of data on the left tail (evidence of holding the conception of variability identified as “variation as distribution”), or simply that the most likely range of scores would be located on the right extreme of the histogram (evidence of holding the conception of variability identified as “variability as the likely range of a sample”):

- Teacher 1: The tendency would be for everybody to do well, the kurtosis of the distribution will be towards the higher scores.  
 Teacher 14: Many obtained the higher scores.  
 Teacher 15: Scores increasing more and more.  
 Teacher 16: Scores on a FAIRLY EASY test, we assume that everybody scored more than 12 points.\*  
 Teacher 27: If the test is very easy, the majority will be distributed towards the higher scores and the minority will get low scores, for that reason is not normally distributed.

*\*Note from the author:* In Venezuela, a 20-point grading scale is commonly used for both test scores and grades, with 20 points being the highest score/grade possible, and 9.5, rounded upwards to 10, the minimum passing score/grade.

## 4.2. RESULTS CORRESPONDING TO ITEM 2

All the participants (100%) answered to this item. Only seven (25.9%) of the surveyed teachers were able to appropriately assess the correctness or incorrectness of the answers provided by all three fictitious students. Within these teachers, five (18.5%) also gave a correct answer to Item 1. Table 2 shows the distribution of answers given by the participating teachers to Item 2.

*Table 2. Frequencies and percentages of all answers to Item 2 given by the participating teachers.*

Teacher assessment of the answer's degree of correctness	Student 1		Student 2		Student 3	
	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
Correct	6	22.2	12	44.4	16	59.3
Incorrect	<b>21</b>	<b>77.8</b>	<b>15</b>	<b>55.6</b>	<b>11</b>	<b>40.7</b>

The values in bold characters correspond to the correct answers.

From the results shown in Table 2, it can be seen that 21 (77.8%) of the surveyed teachers were able to appropriately assess the answer given by Student 1 as incorrect. From the justifications provided by the teachers for this assessment, it was clear that they conceptualized variability in terms of the degree of data clustering around a measure of central tendency (e.g., Teacher 26) or spreading from (e.g., Teacher 1), with both views being evidence of holding the conception of variability identified as “variability as distance or difference from some fixed point.” Thus, these teachers understood that Student 1 has a misconception regarding the interpretation of variability in histograms:

- Teacher 1: Statistically Distribution A is not symmetrical (drawing a line from the X-axis up to the highest point) but because it is less spread out on both sides, it has less dispersion.  
 Teacher 9: Symmetry in a graph does not mean having either more or less variability.  
 Teacher 19: Symmetry is related to mean = median = mode, but not with dispersion.  
 Teacher 26: It's incorrect because the bars are more clustered around the central bar, being Distribution B the one having more variability.

Regarding the assessment of the response given by Student 1, six (22.2%) of the surveyed teachers inappropriately assessed it as correct. From the justifications provided by the teachers for this assessment, it was clear that they misconceptualize variability, focusing on visual cues in the graph and not on the actual data values, their spreads, or in the relation between measures of central tendency and measures of variation. In other words, these teachers evidenced holding the conception of variability identified as “variability as visual cues in the graph”, because they mistakenly interpreted variability as how symmetrical, or how visually close to a normal distribution the histograms were, potentially disregarding the actual data represented:

- Teacher 3: The student associates the asymmetry with the data heterogeneity.  
 Teacher 10: The values don't totally repeat themselves as in case B.  
 Teacher 18: Because indeed that's what variability is all about.

Fifteen (55.6%) of the surveyed teachers were able to appropriately assess the answer given by Student 2 as incorrect. Within this group of teachers, 10 (37.0%) also appropriately assessed the answer given by Student 1 as incorrect. From the justifications provided by the teachers for their assessment of Student 2's answer, it was clear that they conceptualized variability in terms of the degree of data clustering around a measure of central tendency (e.g., Teacher 15) or as data spread, quantifiable through measures of variation such as the range (e.g., Teacher 21), and therefore they understood that Student 2 had a misconception regarding the interpretation of variability in histograms:

- Teacher 1: The fact that there are less data values ( $f$ ) in one distribution than in the other doesn't indicate that there is more or less dispersion.  
 Teacher 15: You are assessing the range not the deviation from the mean.  
 Teacher 19: The simplest measure of dispersion is the total distance covered (horizontal displacement or range) = difference between the largest and the smallest data values, not the data frequencies.  
 Teacher 21: The student is being driven by the frequencies and that doesn't indicate the range of the data.  
 Teacher 25: The range in frequency isn't what determines the dispersion, is the concentration of the data and the frequency.  
 Teacher 27: (Distribution A goes from 2 to 8 and B from 0 to 10). The student doesn't know how to read the graph.

Regarding the assessment of the response given by Student 2, 12 (44.4%) of the surveyed teachers mistakenly assessed it as correct. From the justifications provided by the teachers for this assessment, it was clear that they have a conceptual misunderstanding about range (i.e., judging the variability of the data displayed in a histogram by the largest difference in height of its bars), instead of looking at the span of X-axis values or the horizontal spread of data around a measure of central tendency. This is a common mistake made by many students, and even by some of their teachers, at any school level (Isoda et al., 2018). These teachers showed evidence of holding the conception of variability identified as "variability as whole range", but misinterpreted the concept of range:

- Teacher 3: Distribution B has more variability because the difference between the higher value and the smaller one is higher than in "A".  
 Teacher 13: The student took into account the data distribution as well as the range of the distribution.  
 Teacher 17: Difference in frequencies. The larger the difference, the greater the variability.  
 Teacher 23: It's correct because the student understands the concept of dispersion and applies it to the example.

Sixteen (59.3%) of the surveyed teachers were able to appropriately assess the answer given by Student 3 as correct. All these 16 teachers also appropriately assessed the answer given by Student 1 as incorrect, while only seven of them (25.9%) also appropriately assessed the answer given by Student 2 as incorrect. From the justifications provided by the teachers for their assessment of Student 3's answer, it appeared that they conceptualized variability as either the degree of data clustering around, or data spreading from, a measure of central tendency (evidence of holding the conception of variability identified as "variability as distance or difference from some fixed point"):

- Teacher 1: The distribution of the bars in A makes the curve steep toward the central bar while in B it widens to the sides making the dispersion to be higher.  
 Teacher 8: Because they're more clustered around the mean, mode and median, which in this case are all five.  
 Teacher 9: It's based on the fact that the closer to the value with more frequency, the lower the dispersion.  
 Teacher 15: The measures of dispersion measure the variability in relation to the mean and this reasoning meets that.  
 Teacher 22: If the bars tend more toward the mean there is less dispersion.

Within the group of teachers who appropriately assessed the answer given by Student 3 as correct, four (14.8%) provided inappropriate reasons for this assessment, showing evidence of conceptualizing variability as distance or difference from a fixed reference point or pointer in the graph (in this case, the modal class), but mistakenly focused on the vertical axis:

- Teacher 17: Concentration of the frequency values in relation to the mean.  
 Teacher 21: Looking at the heights of the bars located closer to the central bar, the student realizes that there is variation in the bar heights in B while bar heights in A do not vary as much.  
 Teacher 26: It's correct that Distribution B has more variability and it can be seen that it ranges from 1 to 14, whereas the bars in Distribution A are grouped closer to the central bar.

Regarding the assessment of the response given by Student 3, 11 (40.7%) of the surveyed teachers mistakenly assessed it as incorrect. From the justifications provided by these teachers for such assessment, it was clear that, instead of interpreting variability in histograms in terms of the degree of data clustering around, or data spreading from, a measure of central tendency or some other fixed reference or pointer in the graph, they demonstrated a misconception regarding the interpretation of variability in histograms, judging variability in terms of the number of different “bar height changes” or “jumps” in a histogram, or as the degree of evenness (or lack thereof) to which histogram bars spread out from, or in toward, the central bar. This was evidence that those teachers also held the conception of variability identified as “variability as visual cues in the graph”:

- Teacher 3: Most of the variable's data tend toward values either in the center or of higher frequency.  
 Teacher 5: There are more jumps or discontinuity.  
 Teacher 7: Possibly the student thinks that there is less variability because of the fact that the bars are more grouped together. No. The distribution must be thoroughly observed, taking into account the intervals and the frequency.  
 Teacher 18: The student analyzes variability as having the central bar as a reference point, and from there bars smaller and smaller in height symmetrically spreading away, then the student assumes that the distribution “varies”.  
 Teacher 24: The clustering of bars has nothing to do with it.

## 5. CONCLUSIONS AND DISCUSSION

It is concluded that consideration and evaluation of histograms, in the sense of Items 1 and 2, provided the teachers surveyed with clear opportunities to demonstrate the different ways that they conceptualized variability, as well as to exhibit elements of their professional teaching competencies for teaching statistics.

The data analyzed offer answers to the research questions posed at the beginning of this study. In relation to the first research question, different ways in which the surveyed teachers conceptualized variability were identified, and such ways were described through the analytical framework developed by González (2014a) and illustrated through the justifications made by the teachers. In this way, some limitations were discovered in the teachers' professional knowledge for teaching statistics, such as conceptual errors regarding statistical ideas and objects (e.g., range, symmetry, bar graphs and kurtosis), or the inability of about four-fifths of the respondents to correctly match variables describing real-life situations to their respective histograms. Some cognitive and affective-motivational strengths were also identified. For example, the appropriate use of theoretical properties of a distribution or data set by those teachers who correctly answered to the items discussed in this article. Based on these strengths and limitations, it was possible to answer the second research question; it is evident there is a need for Venezuelan secondary school mathematics teachers to address and eradicate the misconceptions of variability and conceptual errors regarding statistical objects identified in this study, even more so when teachers exhibit misconceptions of variability common to students (Biehler et al., 2018; Isoda et al., 2018; Shaughnessy, 2007), in order to develop and strengthen their competencies for teaching statistics (González, 2014a, 2014b).

There is strong evidence that mathematics teachers from other countries, have similar challenges and problems to those described here as being experienced by Venezuelan teachers (e.g., Biehler et al., 2018, Bruno & Espinel, 2009; Isoda et al., 2018). Therefore, more information is needed on the specific

aspects of the interaction of the different elements comprising the model of teacher competencies for teaching statistics developed by González (2014a), adopted in this study. A future extension of this work may carry out similar studies on a larger sample and use a comprehensive mixed-method approach, combining for that purpose quantitative and qualitative techniques, in order to study, in greater depth, the interaction between teachers' conceptions of variability and other cognitive and affective-motivational elements of their teacher competencies. In this way, we could continue increasing our understanding of the professional competencies for teaching statistics of our in-service teachers, in order to offer recommendations for their improvement and further development.

## REFERENCES

- Ball, D. L., Thames, M., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.
- Batanero, C., Burrill, G., & Reading, C. (Eds.). (2011). *Teaching statistics in school mathematics – Challenges for teaching and teacher education: A joint ICMI/IASE study. The 18th ICMI study*. Springer.
- Biehler, R., Frischmeier, D., Reading, C., & Shaughnessy, J. M. (2018). Reasoning about data. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 139–192). Springer.
- Blömeke, S., & Delaney, S. (2012). Assessment of teacher knowledge across countries: A review of the state of research. *ZDM*, 44(3), 223–247.
- Bruno, A., & Espinel, M. C. (2009). Construction and evaluation of histograms in teacher training. *International Journal of Mathematical Education in Science and Technology*, 40(4), 473–493.
- Centro Nacional para el Mejoramiento de la Enseñanza de la Ciencia y la Matemática. (1990). *Programa de articulación: Contenidos de matemática para la educación media, diversificada y profesional. primer y segundo año (Ciencias)*. [Articulation program: Mathematics content for secondary, diversified and vocational education. First and second year (Science)] CENAMEC.
- Döhrmann, M., Kaiser, G., & Blömeke, S. (2012). The conceptualisation of mathematics competencies in the international teacher education study TEDS-M. *ZDM*, 44(3), 325–340.
- Gal, I. (2004). Statistical literacy. Meanings, components, responsibilities. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47–78). Kluwer.
- González, O. (2014a). *Examining Venezuelan secondary school mathematics teachers' statistical knowledge for teaching: Focusing on the instruction of variability-related concepts*. [Doctoral dissertation, Hiroshima University]
- González, O. (2014b). Survey on secondary school teachers' statistical knowledge for teaching: The need of developing Venezuelan teachers' competence to teach statistics. *International Journal of Curriculum Development and Practice*, 16(1), 27–44.
- González, O. (2014c). Secondary mathematics teachers' professional competencies for effective teaching of variability-related ideas: A Japanese case study. *Statistique et Enseignement*, 5(1), 31–51.
- Isoda, M., Chitmun, S., & González, O. (2018). Japanese and Thai senior high school mathematics teachers' knowledge of variability. *Statistics Education Research Journal*, 17(2), 196–215. <https://doi.org/10.52041/serj.v17i2.166>
- Kuntze, S., & Friesen, M. (2018). The role of mathematics teachers' views for their competence of analysing classroom situations. In B. Rott, G. Törner, J. Peters-Dasdemir, A. Möller, & Safrudiannur (Eds.), *Views and beliefs in mathematics education: The role of beliefs in the classroom* (pp. 183–194). Springer.
- León, N. (2011). Explorando el conocimiento probabilístico informal en niños de edad temprana. [Exploring informal probabilistic knowledge in young children] Paper presented at the XIII Interamerican Conference on Mathematics Education (IACME), Universidade Federal de Pernambuco, Recife, Brazil, June 26–30.
- Makar, K. M., & Confrey, J. (2004). Secondary teachers' reasoning about comparing two groups. In D. Ben-Zvi & J. Garfield (Eds.), *The challenges of developing statistical literacy, reasoning, and thinking* (pp. 327–352). Kluwer.

- Meletiou, M., & Lee, C. (2003). Studying the evolution of students' conceptions of variation using the transformative and conjecture-driven research design. In C. Lee (Ed.), *Reasoning about variability* (pp. 1–39). Central Michigan University.
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.
- Ministerio de Educación. (1997). *Programas de Estudio de 7° a 9° Grado de Educación Básica*. [Study Programs from 7th to 9th grade of Basic Education] ME.
- Peters, S. A. (2009). *Developing an understanding of dispersion: AP statistics teachers' perceptions and recollections of critical moments*. [Doctoral dissertation, Pennsylvania State University]
- Pfannkuch, M. (2018). Reimagining curriculum approaches. In D. Ben-Zvi, K. Makar & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 387–413). Springer.
- Pratt, D. D. (1992). Conceptions of teaching. *Adult Education Quarterly*, 42(4), 203–220.
- Salcedo, A. (2006). Statistics education in Venezuela: The case of elementary and middle school. In A. Rossman & B. Chance (Eds.), *Working Cooperatively in Statistics Education. Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS-7)*, Salvador, Brazil, July 2–7. IASE. <https://iase-web.org/documents/papers/icots7/C125.pdf?1402524966>
- Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, 22(1), 1–36.
- Shaughnessy, J. M. (2007). Research on statistics' reasoning and learning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957–1009). Information Age Publications & NCTM.
- Sivunen, M., & Pehkonen, E. (2009). Finnish elementary teachers' conceptions on problem solving in mathematics education. In J. Maass & W. Schölglmann (Eds.), *Beliefs and attitudes in mathematics education* (pp. 75–86). Sense Publishers.
- Tapia, J. M. (2011). Propuesta didáctica para la enseñanza de la estadística basada en R Commander [Didactic proposal for the teaching of statistics based on R Commander]. Paper presented at the *XIII Interamerican Conference on Mathematics Education (IACME)*, Universidade Federal de Pernambuco, Recife, Brazil, June 26–30.
- Tatto, M. T., Schwille, J., Senk, S., Ingvarsson, L., Rowley, G., Peck, R., & Reckase, M. (2012). *Policy, practice and readiness to teach primary and secondary mathematics in 17 countries*. IAE.

ORLANDO GONZÁLEZ  
 Graduate School of Human Sciences  
 Assumption University  
 592/3 Ramkhamhaeng Rd., Soi 24, Hua Mak, Bangkok  
 Bangkok 10240  
 Thailand