8TH GRADE STUDENTS' STATISTICAL LITERACY OF AVERAGE AND VARIATION CONCEPTS

Ayşe Yolcu, Çiğdem Haser

Hacettepe University, Middle East Technical University

The purpose of this study was to investigate 8th grade students statistical literacy in average and variation concepts through Watson's (1997) three tiered framework. A total of 1074 eight grade students were surveyed with an instrument developed by researchers about statistics content in the elementary mathematics. Students' responses were examined through nine major questions in this instrument for this study. Descriptive analysis of correct and incorrect responses indicated that majority of students understood average and variation concepts though measures of central tendency and spread. Reasons of results and educational implications were discussed.

Keywords: Statistical Literacy, Average, Variation, Elementary Students

INTRODUCTION

Statistical messages have extensively been in the media through several types of arguments, advertisements, or suggestions (Ben-Zvi & Garfield, 2004). Therefore, the ability to understand, interpret, and critically evaluate statistical messages in daily lives of individuals which have been addressed as "statistical literacy" (Watson, 1997), has become important in information societies. Gal (2004) provides a statistical literacy conceptualization and its elements in a model for adults or "future adults", in his term. In this model, communication with statistics, interpretation and judging of statistical claims are treated as the possible skills of statistically literate individuals. In addition statistical literacy plays a crucial role in public and private decision making of individuals where their daily life is full of statistics (Wallman, 1993).

Statistical literacy has also been a part of the school mathematics curriculum to prepare students to encounter the needs of society when they complete their compulsory education (Watson & Callingham, 2003). Understanding of average and variation is fundamental for statistical literacy as the words "average", "variable", or "vary" are a part of everyday language (Watson, 2006). Although there has been studies in the literature examining statistical literacy from different aspects such as sampling (Watson & Moritz, 2000) or graphing (Aoyama & Stephens, 2003) in terms of grade level, research considering statistical literacy of average and variation concepts is scarce specifically in middle school context in Turkey. Therefore, this study investigated Turkish 8th grade students' statistical literacy of average and variation concepts towards the end of their compulsory education.

THEORETICAL FRAMEWORK

Watson's (1997) three tiered framework is at the core of present study and used as a main analysis of students' statistical literacy. Statistical literacy has been addressed by Watson (1997) as the ability of understanding, interpreting, and evaluating statistical messages in various contexts. She presented statistical literacy in a three tiered framework:

> **Tier 1:** Familiarity with and understanding of terminology used in statistical messages.
>
> **Tier 2:** Interpretations of these statistical terms where they are contextualized in statistical claims which appear in the media or elsewhere.
>
> **Tier 3:** The ability to question others' statistical reports critically; in other words, the critical evaluation of biased statistical information and posing possible critical questions to this statistical information.

More specifically, the first tier refers to the familiarity with terminology used in statistical messages in media. To illustrate, understanding the term "average" in context or defining "average" is a feature of Tier 1. For variation concept, the ability for Tier 1 includes expressing ideas of variation in daily life of individuals. The second tier includes the interpretations of these terms where they are contextualized in statistical claims. For example, interpreting or applying ideas of average in a variety of context is a characteristic of Tier 2. The last tier is the ability to question others' statistical reports critically; in other words, the critical evaluation of biased statistical information and posing possible critical questions to this information constitute the third tier of statistical literacy. For instance, examining whether mean or median is an appropriate average in a given statistical report or recognizing extreme values in distributions are basic characteristics of Tier 3 for statistical literacy of average and variation concepts (Watson, 1997, 2006; Watson & Moritz, 2000).

The compulsory education in Turkey addresses the elementary school period which comprises grades 1 to 8. It aims at developing informed citizens who possessed knowledge of statistics with an appreciation of the importance regarding the position of statistics in society (MoNE, 2005) through the National Elementary Mathematics Education curriculum. The elementary mathematics curriculum in Turkey is in spiral in nature and statistical topics including average and variation presented through measures of central tendency and spread across grades 6 to 8. At the end of the elementary school, students are expected to learn meanings of these concepts, how to measure them and where to use them (MoNE, 2005). Therefore, investigating the level of statistical literacy of average and variation concepts that students have developed at the end of the compulsory education becomes important in order to understand the effect of elementary education for citizens. The present study

investigated eighth grade students' statistical literacy in the average and variation concepts in terms of Tier 1, Tier 2, and Tier 3 which Watson (1997) have addressed.

METHODOLOGY

The study aimed at describing aspects and characteristics of students' ability in understanding, interpreting and evaluating certain types of knowledge; therefore, a survey research design was employed (Fraenkel & Wallen, 2006).

Participants

A total of 1074 eighth grade students from 48 classes in 9 randomly selected public schools in a district of Ankara participated in the study. Data were collected by the first author (except 7 classes who were surveyed by their own teachers) in participating students' classrooms during regular class hours.

Data Collection and Analysis

The data collection tool used in this study was prepared to investigate 8^{th} grade students' statistical literacy in sample, average, graph, inference, chance and variation in terms of three tiers. The instrument was developed by the researchers, piloted with 292 8^{th} grade students, and revised through mathematics education researchers', mathematics teachers', and students' comments. The Cronbach's alpha reliability measure in the pilot study was .72 and in the implementation was .75. The analysis of four items related to average concept (A) and five items related to variation concept (V) are presented in this paper. There were both multiple choice and open ended questions for these two concepts. A holistic rubric was prepared in order to classify students' responses in open-ended items and eliminating subjectivity based on their responses in the pilot study and related literature on statistical literacy. Students' responses were coded as non-statistical/incorrect, pre-statistical, and statistical for open-ended items. In addition, explanations of terminology with arithmetic procedures were coded separately.

The responses of Statistical Literacy Test were classified according to the codes in the rubric. These codes were summarized as frequencies and percentages. Then, for descriptive statistics; mean, standard deviation, percentages, and frequencies were calculated.

RESULTS

Statistical Literacy of Average Concept

The first tier of statistical literacy of average concept gives indication of students' understanding of average concept. The answers of students were classified through four categories which were blank or incorrect responses, pre-statistical responses, responses through measures of central tendency and statistical responses. The frequencies and percentages regarding this classification regarding first tier of statistical literacy of average concept are represented in Table 1.

Table 1 indicated that majority of students either explained the term "average" through pre-statistical words (48.7%) or described through measures of central tendency (29.6%). The most notable response in pre-statistical responses was "almost" (34.8%) while "arithmetic mean" or "add them up and divide" algorithm were the most frequent descriptions (26.2%) for those who explained average through measures of central tendency. However, statistically correct responses constituted only 5.2% percent of total responses.

Classification of Responses	**Students' Responses**	**f**	**p**
Blank/Wrong or unrelated responses		170	15.8%
Pre-Statistical		523	48.7%
	Almost	374	34.8%
	Approximately	56	5.2%
	More or less	94	8.8%
Descriptions via Measures of Central Tendency		318	29.6%
	Arithmetic Mean	281	26.2%
	Median	14	1.3%
	Mod	23	2.1%
Statistical		56	5.2%
	Balance point	23	2.1%
	Representative value of data set	33	3.1%

Table 1: Descriptive Statistics for Item A1

Students' familiarity with methods for finding average or central tendency as a characteristic of Tier 1 is analyzed in addition to the explanations regarding average. The results indicated that 44% of students labeled "range" which was not a method for finding average. Yet, 36% of them labeled either "median" (18.1%) or "mod" (18.5%) as if these were not a method for finding average. This finding indicated that almost one third of the participants did not count median and mod as a measure of central tendency.

The second tier of statistical literacy of average concept requires students apply ideas related to average in context. The results showed that 40.2% of students correctly interpreted average in context whereas others (59.1%) chose the incorrect

interpretations. It could be inferred that for average content in the second tier, only less than half of the participants had performed properly.

Evaluation of statistical claims which involves average concept constitute third tier group. There were two items related to this group. The first one was an evaluation of a statistical claim which involved calculating arithmetic mean with an outlier in true/false format where mean was .17 and standard deviation was .37. Majority of students gave incorrect responses or left this item blank (82.7%) which indicated that they were not able to critique a statistical claim in average context. The rest (16.4%) could correctly evaluate the appropriateness of this claim. The explanation for this evaluation was asked through another item where the mean for this item was .05 and standard deviation was .20. Since this item was open-ended, responses of students were classified as incorrect, pre-statistical and statistical. The explanations regarding this item including classifications of these explanations are presented in Table 2.

Classification of Responses	**Students' Responses**	**f**	**p**
Blank/Wrong or unrelated responses		1000	93.1%
	Justification with arithmetic mean	563	52.4%
	Wrong explanations related to context	26	2.4%
	Other blank/wrong explanations	411	38.3%
Pre-Statistical		24	2.2%
	Notice the difference between numbers	14	1.3%
	Notice the outlier/extreme value	10	0.9%
Statistical		40	3.7%

Table 2: Descriptive Statistics for Item A4

Table 2 indicated that most of the participants provided wrong or unrelated responses (93.1%). These students accepted the statistical claim in average context without criticizing either providing wrong explanations related to context (2.4%) such as "Five questions can be solved in a class period" or justifying the results with arithmetic mean (52.4%). The rest of the participants gave pre-statistical (2.2%) or statistical (3.7%) responses. The statistical responses included either recognizing outlier in the data set or stating that getting average with median or mode is more appropriate. The difference between these statistical and pre-statistical responses was the appreciation of variability in the data set occurred in statistical explanations whereas recognizing outlier appeared in pre-statistical responses.

The detailed analysis of items revealed that majority of students had inadequate knowledge regarding average concept for statistical literacy. The most notable finding was that several students understood average which was a characteristics of the first tier behavior as "add them up and divide" algorithm which referred to the arithmetic mean and they did not consider median and mode as a way of finding average. In addition, only less than half of the participants were able to interpret average in context as a characteristic of second tier of statistical literacy. The majority of participants had failed to evaluate a statistical claim which was contextualized as third tier where they could not recognize outlier or justified this claim by providing evidence through arithmetic mean.

Statistical Literacy of Variation Concept

In the first tier of statistical literacy of variation concept, students were asked to select the data set which had more variability among others without context. Results revealed that majority of students (60.8%) were able to choose the data set with more variability. Students were additionally asked to provide explanations for their selections. The frequencies and percentages regarding the classification of responses provided by participants are presented in the following table.

Classification of Responses	**Codes based on Students' Responses**	**f**	**p**
Blank/Wrong or unrelated responses		625	58.2%
	All numbers are same	70	6.5%
	Other blank/wrong explanations	555	51.8%
Pre-Statistical		130	12.1%
	Numbers are increasing	51	4.7%
	Numbers are different	79	7.4%
Descriptions via Measures of Spread		292	27.2%
	Range	248	23.1%
	Inter quartile range	1	0.1%
	Standard deviation	43	4%
Statistical		19	1.8%
	Larger variability	13	1.2%
	Away from average	6	0.6%

Table 3: Descriptive Statistics for Item V2

Table 3 indicated that majority of students (58.2%) either gave wrong responses or did not explain anything related to their answers in the first part. Of these, those who selected the data set which had the same numbers explained their responses through stating "all numbers are the same" (6.5%). Some of the participants (12.1%) provided pre-statistical explanations either stating that "numbers are increasing" (4.7%) or "numbers are different" (7.4%). A considerable percentage of students (27.2%) explained their responses through measures of spread. The most notable response in this category was "range" (23.1%) while "standard deviation" response was quite frequent (4%). Yet, very small percentage of participants (0.1%) explained their response through "inter quartile range". Statistically correct responses constituted only 1.8% percent of total responses where they either indicated the large variability in data set (1.2%) or distance from the average value (0.6%).

The second tier of statistical literacy of variation concept required students to interpret statistical claims involving variability. Results related to this ability indicated that majority of students (74%) were able to interpret statistical claims involving variability. It could be inferred that variability in the second tier was accomplished by most of the participants.

The third tier of statistical literacy of variation concept demanded students to evaluate the data sets and chose the one had more appropriate variability among others. In the context of third tier, the results revealed that majority of students gave incorrect response where only 23.6% of the participants did choose the data set with more appropriate variability which was spreading around center (where the data set is 16, 15, 14, 26, 8, 17). Of the incorrect responses, 20.7% of students did choose "Seda" which had greater variability (where data set is 16, 35, 1, 5, 29, 10), whereas almost one third of the students labeled "Zeynep" which consisted of the same numbers. The classification of the explanations regarding their answers is given in Table 4 below.

Classification of Responses	**Codes based on Students' Responses**	**f**	**p**
Blank/Wrong or unrelated responses		785	73.1%
	Same numbers in the data set	151	14.1%
	Equal to the average	96	8.9%
	Other blank/wrong responses	539	50.2%
Pre-Statistical		125	11.6%
	More difference between numbers	111	10.3%
	Different numbers	14	1.3%

Statistical		149	13.9%
	Appropriate variation	38	3.5%
	Different numbers but closer	38	3.5%
	Around average value	73	6.8%

Table 4: Descriptive Statistics for Item V5

Table 4 indicated that a high percentage of students (73.1%) either gave wrong and unrelated responses or left the explanation part blank. Those who picked "Zeynep" as data set which had the most appropriate variation explained their answers either as "the numbers were equal to the average" (8.9%) or "numbers were the same" (14.1%). The pre-statistical explanations included either more difference between numbers (10.3%) or different numbers (1.3%). Still, there were statistical explanations (13.9%) which consisted of responses such as "appropriate variation" (3.5%), "different but closer numbers" (3.5%), and "around average value" (6.8%).

The detailed analysis of statistical literacy of variation concept indicated that students obviously performed differently in different tiers. For instance, although it was possible to say that there were inadequate knowledge in understanding and evaluating variability, almost 75% of participants correctly interpreted variation in context. One of the interesting finding was that some (6.5% and 14.1% for separate items) of students, indicated that more variation was involved where the data set consisted of same numbers. In addition, very small percentage of students (1.8%) gave statistically correct explanation regarding understanding of variation whereas most of them (27.2%) described variation through measures of spread.

DISCUSSION

The present study strengthened the previous findings that Turkish students tended to consider the average as the arithmetic mean or "add them up and divide" algorithm (Toluk-Uçar & Akdoğan, 2009) as a characteristics of the first tier of statistical literacy of average concept. Most of the students did not consider median and mode as other ways of finding average of a given data set. Only less than half of the participants were able to interpret average in context, which is a second tier ability. This might be derived from students' poor understanding of average concept. Their performance in evaluation of a statistical claim involving average as a representative value, which was a Tier 3 skill, was poor as they could not recognize extreme values, or they explained this claim by providing evidence through arithmetic mean. These results confirm previous findings in which students did not consider average as a representative value for the given data set (Mokros & Russell, 1995).

Students' understanding of average as a summarizing or representative value in this study might be related to Turkish elementary mathematics curriculum. Although

Turkish curriculum has addressed average concept through measures of central tendency (mean, median, and mode), students have started to learn average through arithmetic mean, which may result in understanding average as "add them up and divide" algorithm. In addition, while teaching average concept, teachers may not focus on its characteristics of representative value of a data set; instead they may devote majority of instructional time for computational skills.

Students' performance in the second tier of statistical literacy of variation concept was relatively higher than the first and third tier, which could be attributed to objectives in the curriculum and statistics instruction in Turkish schools. Almost one third of the 8^{th} grade students explained the meaning of variation through measures of spread, particularly range. These responses might be due to the emphasis on the computational skills in statistics content. Turkish elementary curriculum has represented variation concepts through measures of spread (standard deviation, range, and interquartile range) and students might have conceptualized variation concepts through range because it was easier to calculate. Although majority of participants were able to interpret variation concept in various contexts, their responses in other tiers indicated that they considered that there were more variation where the data set consisted of the same numbers. These kinds of responses might be regarded as a sign of possible misconception about variation concept of 8^{th} grade students.

The analysis of students' statistical literacy of average and variation concepts indicated that students conceptualize these contents through arithmetic mean and range which are measures of central tendency and spread. Since these two concepts are fundamental for statistical literacy and further statistics outcomes, understanding, interpretation and critical evaluation of them in various contexts should be emphasized both in curriculum and instruction.

The findings of this study revealed that Turkish 8^{th} grade students had performed lower in first and third tier of statistical literacy of average and variation concepts compared to Tier 2 which was interpretation of statistical claims. Since statistical literacy is an important feature for building active and critical citizens, elementary mathematics curriculum should aim at developing statistical literacy within statistics and probability content area in each grade level. Furthermore, objectives might be modified in relation to support for statistical literacy. There was only one objective regarding evaluation statistical messages in the context of graph concept. Therefore, curriculum makers or planners should identify and include objectives regarding critical evaluation and questioning of statistical claims to promote the development of statistical literacy within elementary school students. For instance, evaluation of arithmetic mean as a representative value or understanding variance within or between data sets should take place as objectives in the curriculum so that there would be the possibility of instructing those objectives.

REFERENCES

Aoyama, K., & Stephens, M. (2003). Graph Interpretation Aspect of Statistical Literacy: A Japanese Perspective. *Mathematics Education Reseacrh Journal , 15* (3), 207-225.

Ben-Zvi, D., & Garfield, J. (2004). Statistical Literacy, Reasoning and Thinking: Goals, Definitions and Challenges. in D. Ben-Zvi, & J. Garfield, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp.3-15). Dordrecht: Kluwer Academic Publishers.

Fraenkel, J., & Wallen, N. (2006). *How to Design and Evaluate Research in Education.* New York: McGraw-Hill.

Gal, I. (2004). Statistical Literacy: Meanings, Components, Responsibilities. In D. Ben-Zvi, & J. Garfield, *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47-78). Dordrecht: Kluwer Academic Publishers.

Mokros, J. & Russell, S. (1995). Children's Concepts of Average and Representativeness, *Journal for Research in Mathematics Education,* 26, 20-39.

Ministry of National Education [MoNE]. (2005), *İlköğretim Okulu Ders Programları: Matematik Programı 6-7-8 (Elementary Curricula Programs: Mathematics Curricula Program for Middle Grades)*. Ankara: MEB.

Uçar, Z. & Akdoğan, E.N. (2009). 6.-8. sınıf öğrencilerinin ortalama kavramına yüklediği anlamlar, *İlköğretim-Online, 8*(2), pp.391-400.

Wallman, K. K. (1993). Enhancing Statistical Literacy: Enriching Our Society. *Journal of the American Statistical Association*, 88 (421), 1-8.

Watson, J. (1997). Assessing Statistical Thinking Using Media. In I. Gal, & J. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 107-122). Amsterdam: IOS Press.

Watson, J., & Callingham, R. (2003). Statistical Literacy: A Complex Hierarchical Construct. *Statistics Education Research Journal , 2* (2), 3-46.

Watson, J., & Moritz, J. (2000). Development of Understanding of Sampling for Statistical Literacy. *Journal of Mathematical Behavior , 19*, 109-136.

Watson, J. (2006) *Statistical Literacy at School: Growth and Goals*, Mahwah, NJ: Lawrence Erlbaum Associates.