ACTA DIDACTICA NAPOCENSIA

Volume 16, Number 2, 2023 - DOI: 10.24193/adn.16.2.8

INFORMAL STATISTICAL INFERENCES OF PRIMARY SCHOOL STUDENTS ON THE CONCEPT OF VARIABILITY

Nadide YILMAZ, Sümeyye AKTAŞ

Abstract: It was aimed to reveal the informal statistical inferences of primary school students on the concept of variability. A phenomenographic research approach was adopted. A task was adapted and used as a data collection tool. Clinical interviews were conducted with 20 primary school students from each grade level. Data were analyzed within the context of the analysis framework hierarchy of development for variability. The results showed that the highest number of students were at Level 1, followed by those who are at Level 0. It can also be seen that the numbers of the students who are at Level 3 and Level 4 are close to each other and smaller than the number of students at Level 1. In addition, it is seen that there is no 1st grade student at Level 4 and that the majority of the students at Level 3 and Level 4 are 3rd and 4th graders. One out of four students had difficulty in informal statistical inferences on the concept of variability. Informal statistical inferences on the concept of variability of three out of every four students, albeit at different levels, drew attention.

Key words: primary school students, variability, informal statistical infereneces, task, clinical interviews

1. Introduction

The fact that individuals encounter graphs, tables and probabilistic situations in all areas of life (Cobb & Moore, 1997) makes it necessary for them to understand and interpret these situations correctly (Watson, 2006). This has made it necessary to include probability and statistics in the teaching process and for students to acquire the related knowledge and skills. However, the inclusion of statistics and probability in school mathematics has a relatively recent history (Watson, 2018). National and international institutions (e.g., Franklin, et al., 2005; Ministry of National Education [MoNE], 2018; National Council of Mathematics [NCTM], 2000) aim to impart knowledge and skills related to statistics and probability concepts to students from an early age by enabling them "to understand basic concepts related to probability, to formulate questions that include data collection and analysis, to create, organize and explain the representations of data and to read and interpret data". Researchers have argued over the relevance of various statistical concepts to early-age learners and they have mostly concluded that emerging teaching approaches often limit students' opportunities to learn or that students are not ready to learn these subjects (English & Mulligan, 2013; Makar, 2018; Perry & Dockett, 2008). When the curriculum in Turkey is examined, it is seen that it aimed to enable the primary school 1st-4th grade students "to read simple tables with at most two data groups, collect data for a given research question, represent and interpret the data with tables and object graphs, prepare a frequency table and tree diagram and read a figure graph, read and interpret simple tables with up to three data groups and organize the data obtained from the table and examine and create a bar graph". In addition, it is aimed to make them use different representations to present the data they have obtained and to solve and pose problems related to daily life by using the information shown in tree diagrams, bar charts, tables and other graphs (MoNE, 2018; p. 14). It can be said that, it is formally aimed to equip students with knowledge and skills related to various graphs and tables. However, the concepts of statistics and probability that students encounter in daily life are much more than these because it is possible to engage in statistical concepts more than it is thought in daily life (Paparistodemou & Meletiou-Mavrotheris, 2018). One of these concepts is the concept of variability (Wild & Pfannkuch, 1999). In this study, it is aimed to reveal the informal statistical inferences of primary school students regarding the concept of variability.

Received March 2023.

Cite as: Yilmaz, N. & Aktaş, S. (2023). Informal statistical inferences of primary school students on the concept of variability. *Acta Didactica Napocensia,* 16(2), 98-112, https://doi.org/10.24193/adn.16.2.8

2. Theoretical Framework

2. 1. The concept of variability

Although researchers emphasize that students should encounter statistical concepts at an early age, many approaches to the teaching processes of early age students overlook this emphasis, arguing that students cannot learn these concepts because they will have difficulties (English & Mulligan, 2013; Makar, 2018; Perry & Dockett, 2008). The importance of the concept of variability in the statistical research process is noted as "variability is at the heart of statistics" (Garfield & Ben-Zvi, 2005, p. 29) because if there was no variability in data sets, there would be no need for statistics (Watson & Kelly, 2002). However, in school mathematics, variability draws attention as a concept that students have difficulty with, and most of the time, students do not understand the logic behind the formulas they use (Watson & Kelly, 2002). For this reason, this concept is not included in the curriculum especially for younger age groups and students do not encounter this concept in the formal education process (Groth, 2018; Watson & Kelly, 2002). However, the concept of variability is an essential element for effective functioning in "real world" and need not be based on an understanding of complex formulas to be taught effectively (Watson & Kelly, 2002; Watson et al., 2007). Shaughnessy (2007) defined eight different aspects of variability focusing on the aspects of the concept of variability that emerge in various statistical contexts. These are i) variation in particular values such as outliers, (ii) variation over time, (iii) variation over an entire range, (iv) variation within a likely range, (v) variation from a fixed value such as a mean, (vi) variation in sums of residuals, (vii) variation in co-variation or association, (viii) variation as distribution. Students often encounter the concept of variability in their daily lives and make informal statistical inferences (Garfield et al., 2007).

2. 2. Students' informal statistical inference about variability

Informal understanding (knowledge) is viewed as either a type of everyday real world knowledge that students bring to their classes based on out-of-school experiences, or a less formalized knowledge of topics resulting from prior formal instruction (Zieffler et al., 2008 p.42). Informal statistical inference can be expressed as what students know about statistical concepts that they have not learned yet and how they express it (Makar, 2014; Makar, et al., 2011). Informal statistical inference prepares students to work on statistical concepts formally (Garfield & Ben-Zvi, 2008; Makar, 2014; Zieffler et al., 2008). Emphasis is placed on the importance of informal knowledge in students' exploring formal concepts and expanding their understanding of what concepts mean in the real world (Cobb & McClain, 2004; Gravemeijer & Doorman, 1999). Studies draw attention to the positive contribution of students' informal knowledge to the teaching of statistical concepts (Bakker & Derry, 2011; Makar & Rubin, 2009; Makar et al., 2011; Zieffler et al., 2008). In addition, if teachers are aware of students' informal knowledge, formal learning environments can be arranged and shaped accordingly (Garfield et al., 2007; Schwarz et al., 2007; Zieffler et al., 2008). Garfield et al. (2007) revealed that university students' informal knowledge contributes to improving their formal knowledge of the concept of variability. However, it can be said that informal statistical inferences of students in younger age groups are less focused on (Doerr et al., 2017; Makar & Rubin, 2009). These studies emphasize that experiences that will allow younger students to make informal statistical inferences using various statistical concepts (e.g., variability) should be included (English, 2018). Studies have reached results showing that students in the younger age group can make comments about the concept of variability that they have not learned formally (English & Watson, 2013; Leavy & Hourigan, 2018; Makar, 2014, 2018; McGatha et al., 2002). For example, English and Watson (2013) in their study with 4th grade students wanted students to experience the contexts they encounter in daily life. Students made various measurements (e.g., arm span). The results showed that the students were able to comment on the concept of variability, including outliers. Makar (2014) revealed that when 8-year-old students had to estimate their typical height, they could make comments on the variability of the height data in their classes. Leavy and Hourigan (2018) worked on 5-6 year-old students and stated that this age-group students were able to record the variability in the data values. These results show that students' working on the tasks they encounter/are likely to encounter in their daily lives has the potential to reveal their experiences regarding the concept of variability (Braham & Ben-Zvi, 2019). Observing how students think while working on such tasks can guide the teacher in students' thinking processes.

One of the tools that allow students to reveal their thinking processes is clinical interviews (Hunting, 1997).

2. 3. Clinical interviews

The increasing awareness that it is not enough to measure the cognitive development of students only with standardized tests has led educators to use clinical interviews which allow observing in detail how students think (Piaget, 1952). Clinical interviews in education allow researchers to have in-depth knowledge of how students think and their mental processes (Goldin, 2000). In this process, the aim of the interviewer is not to direct the student but to get detailed information about his/her way of thinking (Hunting, 1997). In particular, attention is drawn to the potential effect of clinical interviews in revealing how younger students perceive mathematical concepts (Dunphy, 2010; Guimarães & Izabella Oliveira, 2018; Heng & Sudarshan, 2013; Hunting, 1997; Zapata-Cardona, 2018). It is thought that this effect will provide important clues to understand how the teaching process can be improved. If the teacher is aware of how students think and what they know informally, teacher can organize formal teaching processes accordingly (Heng & Sudarshan, 2013). Since it is aimed to reveal the informal statistical inferences of primary school students regarding the concept of variability in this study, it is thought that clinical interview is a tool that will allow to reveal this.

3. Research Question of the Study

The concept of variability is one of the most basic concepts of both statistics and statistics education (Lopes & Cox, 2018). In fact, the statistical problem solving process is based on variability (Franklin et al., 2007). Variability in data requires using many statistical concepts (e.g., measures of shape, centre and dispersion) (Lopes & Cox, 2018). It is emphasized that one of the main purposes should be to deal with the way the concept of variability is handled in statistics teaching (Groth, 2018) because there is variability in all areas of life (Cobb & Moore, 1997). However, the concept of variability is more or less implicitly addressed, especially in the early age group curricula and documents (e.g., NCTM, 2000) (Groth, 2018). At this point, students' informal statistical inferences on the concept of variability come into play. Studies have pointed out that students' making informal statistical inferences can provide richer statistical understandings for students, as well as supporting their meaningful learning when they encounter these concepts formally (English & Watson, 2013; Makar et al., 2011; Paparistodemou & Meletiou-Mavrotheris, 2008; Watson & Kelly, 2002). For example, English and Watson (2013) emphasize the importance of students having discussions on the concept of variability in primary school years before encountering formal statistical concepts such as standard deviation in middle school years. However, it is seen that the studies examining the knowledge of young children about the concept of variability are limited (English & Watson, 2013). In these studies (e.g., English & Watson, 2013; Leavy & Hourigan, 2018; Makar, 2014), it has been observed that younger students have various informal statistical inferences regarding the concept of variability. In addition, it is pointed out that detailed studies are needed to examine these statistical inferences in more depth (English & Watson, 2013). Based on all these reasons, it was aimed to reveal the informal statistical inferences of primary school students regarding the concept of variability and an answer was sought for the following research question:

How are the informal statistical inferences of primary school students regarding the concept of variability?

4. Method

4.1. Research design and participants

A phenomenographic research approach was adopted as it was aimed to reveal the experiences of primary school students regarding the concept of variability and to observe the similarities and differences of these experiences (Trigwell, 2006). The purposive sampling method was used to select the participants. Primary school students from each grade level (1st, 2nd, 3rd and 4th grades) were selected by taking the opinions of teachers in the schools. Moreover, since it was aimed to reveal ideas about the concepts focused on, attention was paid to select the students from among the students who expressed themselves well and were willing to talk. In addition to the opinions of the teachers, the level of achievement in the course and in-class participation were taken into account. According to the primary Education Institutions regulation in Turkey, there are no written or oral exams in the 1st, 2nd and 3rd grades of Primary School (MoNE, 2014). The results of the written exam are required to evaluate the achievement of 4th grade students, but since the written exam was not applied to any of the students due to the Covid-19 epidemic when the current application was performed, the achievement of each student was determined by taking the teacher's opinion. Information on student performance is shown in Table 1.

Table 1. *Students' grade levels, achievement levels and in-class participation*

Students	Grade Levels	Achievement Level	In-Class Participation
S1A	1st Grade	Needs Improvement	Active
S1B		Very Good	Active
S1C		Very Good	Passive
S1 E		Good	Passive
S1 D		Good	Active
S2E	2nd Grade	Good	Active
S2D		Very Good	Active
S2A		Very Good	Active
S2B		Good	Passive
S2C		Good	Active
S3C	3rd Grade	Very Good	Active
S3E		Very Good	Active
S3B		Needs Improvement	Passive
S3D		Good	Active
S3A		Good	Active
S4D	4th Grade	Very Good	Active
S4E		Very Good	Active
S4B		Very Good	Active
S4C		Good	Passive
S4A		Good	Active

The students' real names were not used, code names were used, instead. For example, while S4D represents a 4th grade student, S1B represents a 1st grade student. As seen in Table 1, the study was conducted on five students from each grade level. It can be said that students differed in terms of both in-class participation and achievement level.

4.2. Data collection tool and data collection process

A task used in the study of McGatha et al. (2002) was adapted and used as a data collection tool (Figure 1). The original of the task was named "Basketball All-Star". The students were given a list of the points scored by two basketball players in each of the eight games. They were then asked to decide which player should be chosen to play in the all-star tournament based on these scores. This task was translated into Turkish and submitted to expert review. Here it was suggested that if the task was presented within a scenario, it might be more meaningful for primary school students. Based on this suggestion, a context that students would be familiar with was created. The Karaman Primary School announced that students would be selected from the school to play in an inter-provincial basketball tournament. Students were expected to decide on the appropriate student selection based on the points scored by the students before the tournament. This task is presented in Figure 1.

WHO SHOULD BE IN THE BASKETBALL TOURNAMENT?

The Karaman primary school will participate in the inter-provincial basketball tournament. For this purpose, the points scored by two students were noted and shown below.

Alperen 11 31 16 28 27 14 26 15

Mehmet 21 17 22 19 18 21 22 20

Based on the points these students scored, which one do you think best represents the school for the basketball tournament? Why?

Figure 1. *Task used in the clinical interviews*

In this task the arithmetic means of the points scored by the two students are very close to each other (arithmetic mean Alperen: 21, arithmetic mean Mehmet: 20). The students are expected to determine the student who consistently scored high points. In other words, they are expected to prefer the data set with a low standard deviation. However, the students do not have formal knowledge of the concepts of arithmetic mean and standard deviation. When evaluated in this sense, the students are expected to think about all the values in the data set.

The data collection tool was made ready for a pilot study after the expert review. For this purpose, a pilot study was conducted with one student from each grade level and the process was video recorded. Since no problems were observed in the pilot study, the actual study was carried out. The data of the study was collected in the fall semester of the 2021-2022 school year. Interviews were held with each student and the process was video-recorded. In addition, the researcher took notes about the situations that would allow her to better describe the research data that attracted her attention throughout the process. Each interview lasted between 25 and 30 minutes.

4.3. Data analysis

Descriptive analysis was used to analyze the data. Reading and Shaughnessy (2004) evaluated students' opinions about variability under two main headings in their study. These headings are description and causation hierarchy about variability. Since the current study focused on students' informal statistical inferences regarding variability, the framework in which Reading and Shaughnessy (2004) described variability was adopted as the analysis framework. However, while the findings were analyzed, it was observed that some answers could not be explained within the framework of the analysis. Therefore, a level was added, and this level was described as 0. This framework consists of four levels. Each level is explained below.

4.3.1. Level 0. Instead of evaluating the data set, the student makes comments by considering various components (e.g., personal experiences, the sum of the numbers of data) and gives wrong answers. For example, the statement "*Alperen should participate in the basketball tournament because the total number of points scored by him is higher than Mehmet*" is evaluated at this level in the current study.

4.3.2. Level 1. Students at this level focus on the median or outliers. The student describes the data set he/she is considering in terms of either medians or outliers. They cannot focus on both measures. For example, as can be seen from the following statement in the current study; *"The maximum number of points scored by Alperen was 31, but he also scored 11 points. Mehmet, on the other hand, scored a maximum of 22 points and scored a minimum of 17 points. Choosing Alperen makes sense at first glance, but what if he scores 11 points in the tournament? That's why I'm sending Mehmet to the* *tournament.”*, the student made evaluations only on the basis of outliers and thus this answer of the student was considered to be at level 1.

4.3.3. Level 2. Students at this level focus on both medians and outliers. Students describe the dataset by mentioning possible outliers (e.g., maximum and minimum) and making associations between them. For example, as seen from the following statement: *"Let’s look at Alperen, he made a pretty good total score of 31 points, but he also made a bad total score of 11 points. Also, in the middle there are 16 and 26. Let’s look at Mehmet, his all scores are close to 20 points. His maximum score is 22 points, and minimum score is 17 points. That's why I chose Mehmet for the tournament.”*, the student made his/her evaluations on the basis of the values in the middle as well as outliers. Thus, this answer of the student was considered to be at Level 2.

4.3.4. Level 3. Students at this level began to consider some deviations in the data set. However, these deviations do not necessarily have to be the central idea. For example, as seen from the following statement: *“Alperen scored 11 14 15 Mehmet scored 17 18 19. I choose Mehmet for the tournament ….. because when we look at these values, we can see that Mehmet scored more …. Moreover, Alperen scored between 31 and 11 while Mehmet scored between 22 and 17. Therefore, I choose Mehmet.”*, it can be seen that the student made his/her evaluations considering some deviations in the data set. Thus, this answer of the student was considered to be at Level 3.

4.3.5. Level 4. Students at this level define the data set by taking into account the changes in the centre. They can define values that deviate from the centre. For example, as seen from the following statement: *“Mehmet always scored about 20 points like 20, 21. His points below 20 are also very close to 20. That is, he scored not higher or not lower than 20. Yet, Alperen even scored more than 30 points and also scored points very close to 10. He also scored points around 20 but also highly lower and higher scores than 20. Therefore, it would be more suitable for Mehmet to join the tournament.”*, the student made evaluations focused on the centre of the data set. Thus, this answer of the student was considered to be at Level 4.

The video recordings taken during the clinical interviews with each student were watched and transcribed. The answers given by the students individually (drawings, expressions and explanations) were determined as the unit of analysis. According to the analysis framework explained above, informal statistical inferences regarding the concept of variability were coded for each student.

In the current study, a number of measures were taken to ensure validity and reliability. First, how the research was carried out was explained in detail. In addition to how the data was collected, how the codes were created and the compatibility between the coders were examined. The rate of agreement between the coders was found to be 90%. This rate of agreement can be said to be enough to ensure reliability of the study (Miles & Huberman, 1994). The points of disagreement were discussed until reaching a consensus. In addition, video interview recordings and notes taken from the observations were used to ensure the verifiability of the data. Moreover, the process was explained to the participants in detail and instead of their real names, codes were used to represent the participants.

5. Findings

Results related to the informal statistical inferences of the primary school students about the concept of variability are presented in Table 2.

Table 2. *Levels of the primary school students regarding the concept of variability*

Level	1st Grade	2nd Grade	3rd Grade	4th Grade	Total
Level 0	S1A, S1B	S2A	S3A	S4A	5
Level 1	S1C, S1D	S2B, S2C	S3B	S4B, S4C	7
Level 2		S2D			1
Level 3	S1E		S3C, S3D	S4D	4
Level 4		S2E	S3E	S4E	3

It is seen that the highest number of students were at Level 1, followed by those who are at Level 0. It can also be seen that the numbers of the students who are at Level 3 and Level 4 are close to each other and smaller than the number of students at Level 1. In addition, it is seen that there is no 1st grade student at Level 4 and that the majority of the students at Level 3 and Level 4 are 3rd and 4th graders. Examples of interviews with students for each level are given below.

5.1. Level 0

It was observed that students at this level made comments and gave incorrect answers by considering various factors (e.g., personal experiences, the sum of the numbers of data) instead of evaluating the data set. It was noted that the students evaluated the given task with three different ways of thinking. The first of these is that one of the students made evaluations on the basis of the total points scored. Below is given a section of the interview conducted with S2A, a 2nd grade student.

S2A: (He reads the activity)

Researcher: Yes, what do you think, should Mehmet or Alperen represent the school?

S2A: (Thinks for a while). I think Alperen should be chosen (performing some operations)

Researcher: Why?

S2A: I summed up the points they scored. Alperen scored a total of 168 points while Mehmet scored a total of 160 points; therefore, Alperen should be chosen.

Researcher: Then, the number of points scored was decisive in your selection.

S2A: Yes, Alperen can score higher in the tournament as he previously scored more.

S2A could not evaluate the data set and he decided which student to be selected on the basis of the total number of points scored. Another way of thinking noticed was that students made their choices by considering outliers. This way of thinking was observed in two students (S3A and S4A). A section from the interview with S3A regarding this situation is given below.

Based on the points scored by these students, which one do you think best represents the school for the basketball tournament?

S3A: (Reads the text in the activity) I think Alperen

Researcher: Why?

S3A: Alperen scored more.

Researcher: Which numbers are more? According to what did you decide?

S3A: 31, 28, 27, 26

Researcher: According to what did you choose the numbers?

S3A: I chose the higher points he scored.

Researcher: Should Alperen go as these numbers are high?

S3A: Yes.

It was observed that S3A focused on the higher points scored by Alperen and that she decided to choose Alperen by taking into consideration the outliers. Another way of thinking noted was that students reached wrong conclusions based on both outliers and values in the centre of the data set (S1A and S1B). A section from the interview conducted with S1B regarding this situation is given below.

Researcher: Yes, let's read activity S1B. (The student reads the activity). Yes, what do you say?

S1B: I think Alperen should be chosen.

Researcher: Why?

S1B: Because he scored 31 points, so he scored more.

Researcher: Is there anything else that influenced your choice?

S1B: Mehmet has never scored 30 points. He scored points around 20 points. But Alperen both scored 31 points and scored points close to 30 like 26, 27, 28. That's why I think Alperen should be chosen.

When the section from the interview with S1B was evaluated, it was noticed that while S1B was making her decision, she took the outliers into account and also took into account the values in the middle of the data set and reached wrong results.

5.2. Level 1

It was observed that students at this level made comments by focusing on either the values in the middle or the outliers in the data set. They did not take both into account. Below is the section from the interview with S1C, one of the students who made evaluations by focusing on outliers.

Researcher: Who would you send to the tournament based on these scores?

S1C: Now I look at the points scored by Mehmet and Alperen. Alperen scored 31 points, but also 11 points. When I look at Mehmet, I see that he has scored 22 points, and his lowest score is 17. I think Mehmet should be chosen.

Researcher: Why?

S1C: Because Alperen scored 31 points but also 11 points. Mehmet, on the other hand, scored 17 points and 22 points. Choosing Alperen makes sense at first glance, but what if he scores 11 points in the tournament? That's why I would send Mehmet to the tournament.

It was observed that S1C made his choice according to the maximum and minimum values in the data set. It was revealed that there were also students who made their evaluations on the basis of the values in the middle. Below is given a section of the interview with S4B, one of these students.

Researcher: Who would you send to the tournament based on these scores?

S4B: Now I look at the points scored by Mehmet and Alperen (He sorted points the students scored from largest to smallest.)

Alperen: 31 28 27 26 16 15 14 11

Mehmet: 22 22 21 21 20 19 18 17

Now I think we should choose Mehmet.

Researcher: Why?

S4B: When I look at the points scored by Mehmet and Alperen, the middle values are 26 and 16 in Alperen while they are 21 and 20 in Mehmet. That is, Alperen scored 26 points but also 16 points. But Mehmet scored 21 points and 20 points. Therefore, I would choose Mehmet. Because Alperen can score 16 points in the tournament.

It was observed that S4B evaluated and decided based on only two values in the middle of the data set.

5.3. Level 2

At this level, the students evaluated the data set by focusing on both the values in the middle and outliers. Only one student was found at this level. Below is given a section of the interview with S2D, a 2nd grade student at this level.

Researcher: Based on the points scored by these students, which one do you think best represents the school for the basketball tournament?

S2D: (Sorted the data)

Alperen: 11 14 15 16 26 27 28 31

Mehmet: 17 18 19 20 21 21 22 22

Now let's look at Alperen, he has a pretty good score of 31, but there is also a bad score of 11. Also, there are 16 and 26 in the middle. Mehmet on the other hand has scores around 20. The highest score is 22 and the lowest score is 17. Therefore, I would choose Mehmet for the tournament.

Researcher: Why?

S2D: Because when you look at it, it makes more sense to choose Mehmet.

It was detected that S2D made a decision by evaluating the values in the middle of the data as well as the outliers.

5.4. Level 3

At this level, the students evaluated the data set in terms of deviations from an anchor value that is not a central value. A section from the interview with S3C regarding this situation is given below.

Researcher: Based on the points scored by these students, which one do you think best represents the school for the basketball tournament?

S3C: (Students marks certain points) (See Figure 2)

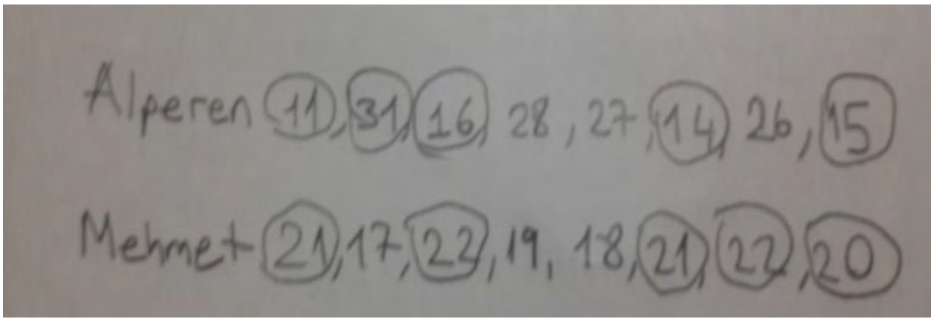

Figure 2. *Student's answer*

Researcher: Why did you mark these values?

S3C: These values caught my attention because.

Researcher: Who do you think should be chosen for the tournament?

S3C: Mehmet should go.

Researcher: Why?

S3C: Because when I look at the values I chose, it is seen that Mehmet scored points around 20, but Alperen scored more in the 10s. Therefore, Mehmet should go.

Researcher: Well, didn't the other values (that you didn't mark) affect your choice for the tournament?

S3C: No, it didn't.

Researcher: Why?

S3C: Few because (meaning that the number of the remaining three values is less than the number of the other value).

It can be said that S3C made an evaluation by taking into account the points mostly scored by the students, that is, he tried to take into account some deviations in the data set.

5.5. Level 4

It was observed that students at this level made evaluations by taking into account the changes in the centre of the data set. A section from the interview conducted with S4E regarding this situation is given below.

Researcher: Based on the points scored by these students, which one do you think best represents the school for the basketball tournament?

S4E: (She performs some operations for a while) If I were you, I would send Mehmet.

Researcher: Why?

S4E: Now I have just looked; I sorted the data first. (Sorted the data)

Researcher: Why?

S4E: In this way, I can evaluate the points scored better.

Researcher: What did you do then?

S4E: Now, when I look at it, I can see that Alperen scored different points, but Mehmet scored points close to each other. First, I picked a number.

Researcher: Why did you pick such a number?

S4E: Because I can better determine how many points they scored in this way.

Researcher: Which number did you pick? How did you decide on this number?

S4E: Now this number should be common in both of the points scores by these students. First, I picked 10, but since Mehmet did not have 10, I decided to pick 20.

Researcher: What did you do then?

S4E:

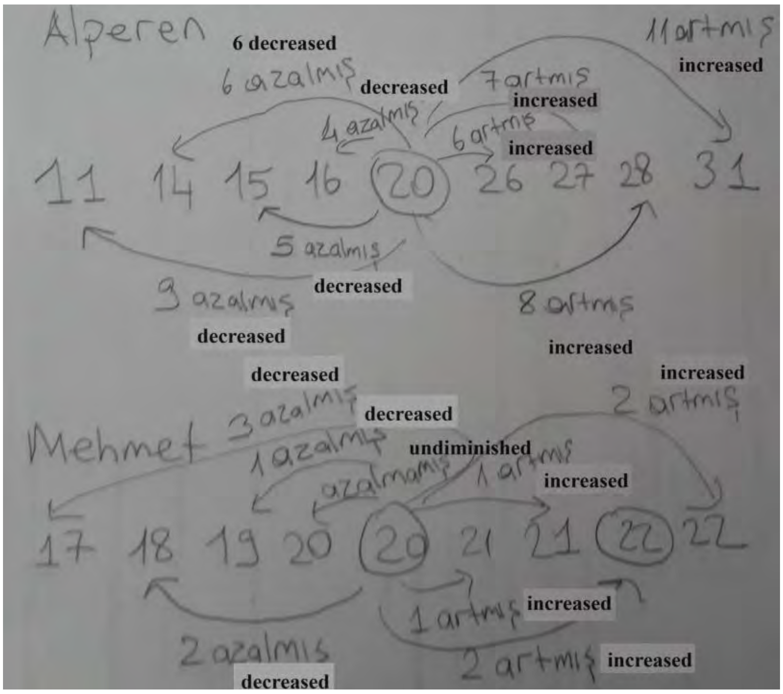

Figure 3. *Student's answer*

(See Figure 3) *Now, I looked at the points scored by Alperen according to 20 points. First, they increased by 6, then they decreased by 4, increased by 7, decreased by 5, increased by 8, decreased by 6, increased by 11 and decreased by 9; when I looked at them, I saw that there were large increases and decreases in the points scored. But when I looked at the points scored by Mehmet, I saw that there was not much fluctuation; first increased by 1, then increased by 1, decreased by 1, increased by 2, decreased by 2, increased by 2 and decreased by 3. Thus, I saw that there are small decreases and increases. Therefore, it makes more sense to choose the one with smaller decreases and increases, so I would send Mehmet.*

Researcher: What does it mean to prefer the one with smaller decreases and increases? I don't understand.

S4E: I actually tried to find the more balanced performance. For example, Alperen scored 31 but also scored 11. What if he would score fewer points in the tournament. But this is not the case with Mehmet. That's why I would choose Mehmet.

It was observed that S4E interpreted the data in the data set in a way that took into account the deviations from the centre.

6. Discussion and Conclusion

In the current study, it was aimed to reveal the informal statistical inferences of primary school students regarding the concept of variability. The results obtained showed that one out of four students had difficulties in informal statistical inferences regarding variability. Some of these students made wrong evaluations by considering only the sum of the numbers of data and some students made wrong evaluations by considering only outliers and the values in the middle. There are studies reporting similar results (Watson & Moritz, 2001). The difficulties encountered by both younger age groups (e.g., Watson et al., 2007) and the secondary and high school students (e.g., Watson, 2009; 2018) regarding the concept of variability have been pointed out in the literature. The difficulties experienced by students can be evaluated under several headings. The first of these is the extent of their familiarity with the given context. It has been stated that the content of the context may affect students' understanding of the concept of variability (Torok & Watson, 2000; Watson et al., 2007; Watson & Kelly, 2005). In some studies (e.g., Watson, 2018), it has been observed that students do not take the data set into account at all but make evaluations based on their own experiences. This reasoning is named imaginary reasoning (Kinnear, 2013). In the current study, however, such a result was not found. In this sense, it can be said that the results obtained differ from the literature. In addition, the fact that the students did not learn this concept formally may have caused the students to have difficulties with this concept.

Informal statistical inferences on the concept of variability of three out of every four students, albeit at different levels, drew attention. It was observed that the students were mostly at Level 1, and there were few students at Levels 2, 3 and 4. In other words, students mainly made evaluations according to the values in the middle or outliers in the data set. Students who took both into account constituted only 5 percent of all the students. One out of every five students interpreted the concept of variability at Level 3, taking into account some deviations in the data set. Fifteen percent of the students made evaluations by taking into account the changes in the centre of the data set. Informal statistical inferences about the concept of variability, albeit at different levels, can be considered as an indication that students are trying to engage in beginning to inferential reasoning (Ben-Zvi et al., 2012; Makar & Rubin, 2009). Working on a situation that students are likely to encounter in their daily lives may have caused them to produce accurate informal statistical inferences regarding the concept of variability. In the studies conducted, attention is drawn to the content of the given context in cases where the data need to be interpreted (Langrall et al., 2006). For example, it is emphasized that informal statistical inferences of children can be revealed better when the contexts in the tasks are motivating and meaningful (Falk et al., 2012). In different studies, it has been observed that students can make interpretations about the concept of variability even if they do not know it formally (English & Watson, 2013; Leavy & Hourigan, 2018; Makar, 2014; 2018; McGatha et al., 2002). However, at this point, it should be noted that students' thinking levels should not be evaluated only by their familiarity/unfamiliarity with the given context because emergence of differences from student to student even when commenting on the given context makes it possible to interpret that students have difficulties in establishing connections between real-world knowledge and the given situation. When it was evaluated which aspect of the concept of variability defined by Shaughnessy (2007) was included in their interpretations by the students, it was seen that they included the following aspects; i) variation in particular values such as outliers, (ii) variation over time, (iii) variation over an entire range and (v) variation from a fixed value such as a mean.

Another result found is that none of the 1st grade students were at Level 4 and that the students at Level 3 and 4 were mainly 3rd and 4th grade students. When evaluated in this context, it can be said that as the grade level increases, the informal statistical inferences of the students regarding the concept of variability increase. The possibility of increasing informal experiences with increasing grade level may have triggered such a result. When the literature is examined, it is seen that some studies show a parallel development between the grade level and the development level of students (e.g., Shaughnessy et al., 1999; Watson & Moritz, 2001), while in some studies, it has been revealed that there is no correlation between the development level and the grade level (Watson & Moritz, 2001).

7. Limitations and Suggestions

This study focused on revealing the informal statistical inferences of primary school students at different grade levels regarding the concept of variability through a single context and did not aim to generalize the results to the informal statistical inferences of other primary school students. Working on a single task can be considered a limitation. In future studies, informal statistical inferences about the concept of variability can be revealed by studying different contexts. The similarities/differences of how they employ the concept of variability in different contexts can be evaluated. The results showed that most of the students were able to make informal statistical inferences on the concept of variability, albeit at different levels. In light of these results, it can be suggested that the concept of variability should be included in the curriculum (Watson & Kelly, 2002). It is even stated that with appropriate learning experiences (e.g., Groth, 2018; Watson & Kelly, 2002), younger students show improvement in the concept of variability. Thus, informal statistical inferences about the concept of variability of younger age group students can be developed by arranging instructional designs in further studies.

References

Bakker, A. & Derry, J. (2011). Lessons from inferentialism for statistics education. *Mathematical Thinking and Learning, 13*(1–2), 5–26. https://doi.org/10.1080/10986065.2011.538293

Ben-Zvi, D.; Aridor, K.; Makar, K., & Bakker, A. (2012) Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM Mathematics Education, 44*(7), 913–925. https://doi.org/10.1007/s11858-012-0420-3

Braham, H. M.& Ben-Zvi, D. (2019). Design for Reasoning with Uncertainty. In Burrill, G., Ben-Zvi, D. (eds) *Topics and Trends in Current Statistics Education Research* (pp. 97-121). Springer.

Cobb, P., & McClain, K. (2004). Proposed design principles for the teaching and learning of elementary statistics. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 375-396). Kluwer.

Cobb, G., & Moore, D. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly, 104*(9), 801–823.https://doi.org/10.2307/2975286.

Doerr, H. M., Delmas, R., & Makar, K. (2017). A modeling approach to the development of students' informal inferential reasoning. *Statistics Education Research Journal, 16*(1), 86-115.

Dunphy, E. (2010). Exploring young children's (mathematical) thinking: Preservice teachers reflect on the use of the one-to-one interview. *International Journal of Early Years Education, 18*(4), 331-347. https://doi.org/10.1080/09669760.2010.531610.

English, L. D. (2018). Young children's statistical literacy in modelling with data and chance. Leavy, A., Meletiou-Mavrotheris, M., Paparistodemou, E. (eds) *Statistics in Early Childhood and Primary Education. Early Mathematics Learning and Development* (pp 295–313). Springer.

English, L. D., & Mulligan, J. T. (2013). Perspectives on reconceptualizing early mathematics learning. In L. English & J. Mulligan (Eds.), *Reconceptualizing early mathematics learning* (pp. 1–4). Springer.

English, L.D. & Watson, J. (2013). Beginning inference in fourth grade: Exploring variation in measurement. In V Steinle, L Ball, & C Bardini (Eds.), *Mathematics education: yesterday, today and tomorrow (Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia)* (pp. 274–281). MERGA.

Falk, R., Yudilevich-Assouline, P., & Elstein, A. (2012). Children's concept of probability as inferred from their binary choices-Revisited. *Educational Studies in Mathematics, 81*(2), 207–233. https://doi.org/10.1007/s10649-012-9402-1.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M. ve Scheaffer, R. (2005) *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K–12 Curriculum Framework*. American Statistical Association.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M. ve Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A preK-12 curriculum framework.* American Statistical Association.

Garfield, J. & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal, 4*(1), 92–99.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning. Connecting research and teaching practice.* Springer.

Garfield, J., delMas, R. & Chance, B. (2007). Using students' informal notions of variability todevelop an understanding of formal measures of variability. In M. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 117–147). NY: Lawrence Erlbaum.

Goldin, G. (2000) A scientific perspective on structures, task-based interviews in mathematics education research. In Lesh R, Kelly, A.E. (eds) *Research design in mathematics and science education*. (pp 517–545). Erlbaum.

Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational Studies in Mathematics, 39,* 111-129. https://doi.org/10.1023/A:1003749919816.

Groth, R. E. (2018). Unpacking implicit disagreements among early childhood standards for statistics and probability, In Leavy, A., Meletiou-Mavrotheris, M., Paparistodemou, E. (eds) *Statistics in Early Childhood and Primary Education. Early Mathematics Learning and Development*. (pp 149–162). Springer.

Guimarães, G. & Oliveira, I. (2018). How kindergarten and elementary school students understand the concept of classification, In Leavy, A., Meletiou-Mavrotheris, M., Paparistodemou, E. (eds) *Statistics in Early Childhood and Primary Education. Early Mathematics Learning and Development*. (pp 129–146). Springer.

Heng, M. A., & Sudarshan, A. (2013). "Bigger number means you plus!" – Teachers' learning to use clinical interviews to understand students' mathematical thinking. *Educational Studies in Mathematics, 83*(3), 471-485. https://doi.org/10.1007/s10649-013-9469-3.

Hunting, R.P. (1997). Clinical interview methods in mathematics education research and practice. *Journal of Mathematical Behavior, 16,* 145-165. https://doi.org/10.1016/S0732-3123(97)90023-7.

Kinnear, V. A. (2013). Young children's statistical reasoning: a tale of two contexts. [Unpublished Doctoral Thesis]. Queensland University of Technology.

Langrall, C. W., Nisbet, S., & Mooney, E. S. (2006). The interplay between students' statistical knowledge and context knowledge in analysing data. In A. Rossman & B. Chance (Eds.), *Proceedings of the 6th International Conference on Teaching Statistics (ICOTS6, Salvador, Brazil).* Voorburg, The Netherlands: International Statistics Institute.

Leavy, A. & Hourigan, M. (2018). Inscriptional capacities and representations of young children engaged in data collection during a statistical investigation, In Leavy, A., Meletiou-Mavrotheris, M., Paparistodemou, E. (eds) *Statistics in Early Childhood and Primary Education. Early Mathematics Learning and Development*. (pp. 89–107). Springer.

Lopes, C. E. & Cox, D. (2018). The Impact of Culturally Responsive Teaching on Statistical and Probabilistic Learning of Elementary Children, In Leavy, A., Meletiou-Mavrotheris, M., Paparistodemou, E. (eds) *Statistics in Early Childhood and Primary Education. Early Mathematics Learning and Development*. (pp. 77-88). Springer.

Makar, K. (2014). Young children's explorations of average through informal inferential reasoning. *Educational Studies in Mathematics, 86*(1), 61-78. https://doi.org/10.1007/s10649-013-9526-y.

Makar, K. (2018). Theorising links between context and structure to introduce powerful statistical ideas in the early years. In Leavy, A., Meletiou-Mavrotheris, M., Paparistodemou, E. (eds) *Statistics in Early Childhood and Primary Education. Early Mathematics Learning and Development*. (pp. 3-20). Springer.

Makar, K., Bakker, A. & Ben-Zvi, D. (2011). The reasoning behind informal statistical inference. *Mathematical Thinking and Learning, 13,* 152-173, https://doi.org/10.1080/10986065.2011.538301

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal, 8*(1), 82–105.

McGatha, M., Cobb, P., & McClain, K. (2002). An analysis of student's initial statistical understandings: Developing a conjectured learning trajectory. *The Journal of Mathematical Behavior, 21*(3), 339–355. https://doi.org/10.1016/S0732-3123(02)00133-5.

Miles, M. B. & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook.* Sage.

Ministry of National Education [MoNE]. (2014). *Millî Eğitim Bakanlığı Okul Öncesi Eğitim ve İlköğretim Kurumları Yönetmeliği. [Regulation on Pre-School Education and Primary Education Institutions of the Ministry of National Education].* Ankara, Turkey.

Ministry of National Education [MoNE]. (2018). *Matematik dersi (İlkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. Sınıflar) öğretim programı [Primary and middle school mathematics curricula for grades 1, 2, 3, 4, 5, 6, 7, and 8]*. Ankara, Turkey

National Council of Mathematics [NCTM]. (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author.

Paparistodemou E. & Meletiou-Mavrotheris, M. M. (2018). Teachers' Reflection on Challenges for Teaching Probability in the Early Years. In Leavy, A., Meletiou-Mavrotheris, M., Paparistodemou, E. (eds) *Statistics in Early Childhood and Primary Education. Early Mathematics Learning and Development*. (pp. 201-215). Springer.

Perry, B., & Dockett, S. (2008). Young children's access to powerful mathematical ideas: A Review of Current Challenges and New Developments in the Early Years. In L. English (Ed.), *Handbook of International Research in Mathematics Education* (pp. 153-190). Routledge.

Piaget, J. (1952). *The Child's Conception of Number*. Routledge & Kegan Paul Ltd.

Reading, C., & Shaughnessy, J. M. (2004). Reasoning about variation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 201–226). Kluwer.

Schwartz, D. L., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 319–344). Lawrence Erlbaum Associates Publishers.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957–1009). NC: NCTM.

Shaughnessy, J. M., Watson, J., Moritz, J., & Reading, C. (1999, April). School mathematics students' acknowledgment of statistical variation. In C. Maher (Chair), *There's more to life than centers*. Presession Research Symposium, 77th Annual National Council of Teachers of Mathematics Conference, San Francisco, CA

Torok, R. & Watson, J. (2000). Development of the concept of statistical variation: An exploratory study. *Mathematics Education Research Journal, 12*(2), 147-169. https://doi.org/10.1007/BF03217081

Trigwell, K. (2006). Phenomenography: An approach to research into geography education. *Journal of Geography in Higher Education, 30*(2), 367-372. https://doi.org/10.1080/03098260600717489

Watson, J. (2006). *Statistical literacy at school*. Lawrence Erlbaum Associates

Watson, J. M. (2009). The influence of variation and expectation on the developing awareness of distribution. *Statistics Education Research Journal, 8*(1), 32–61.

Watson, J. (2018). Variation and expectation for six-year-olds. In Leavy, A., Meletiou-Mavrotheris, M., Paparistodemou, E. (eds) *Statistics in Early Childhood and Primary Education. Early Mathematics Learning and Development*. (pp. 55-73). Springer.

Watson, J. M., Callingham, R. A. & Kelly, B. A. (2007). Students' appreciation of expectation and variation as a foundation for statistical understanding. *Mathematical Thinking and Learning, 9*(2), 83–130. https://doi.org/10.1080/10986060709336812.

Watson, J.M. & Kelly, B. A. (2002). Can Grade 3 students learn about variation?, In B. Phillips (ed.), *Proceedings of the Sixth International Conference on Teaching Statistics: Developing a Statistically Literate Society, Cape Town, South Afric*a, International Statistical Institute, Voorburg, The Netherlands.

Watson, J. M. & Kelly, B. A. (2005). The winds are variable: Students' intuitions about the weather. *School Science and Mathematics, 105*, 252–269.https://doi.org/10.1111/j.1949-8594.2005.tb18165.x.

Watson, J. M. & Moritz, J. B. (2001). Development of reasoning associated with pictographs: Representing, interpreting, and predicting. *Educational Studies in Mathematics, 48*(1), 47-81. https://doi.org/10.1023/A:1015594414565.

Wild, C., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223–265. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x.

Zapata-Cardona, L. (2018). Supporting young children to develop combinatorial reasoning, In Leavy, A., Meletiou-Mavrotheris, M., Paparistodemou, E. (eds) *Statistics in Early Childhood and Primary Education. Early Mathematics Learning and Development*. (pp. 257-272). Springer.

Zieffler, A., Garfield, J., Delmas, R. & Reading, C. (2008). A Framework to support research on informal inferential reasoning. *Statistics Education Research Journal, 7*(2), 40-58.

Authors

Nadide YILMAZ, Karamanoğlu Mehmetbey University, Karaman (Turkey). E-mail: nadideylmz70@gmail.com

Sümeyye AKTAŞ, Karamanoğlu Mehmetbey University, Karaman (Turkey). E-mail: saktas664@gmail.com