YOUNG CHILDREN'S PROBABILISTIC AND STATISTICAL REASONING IN THE CONTEXT OF INFORMAL INFERENTIAL REASONING

GAMZE KURT
Mersin University
gamzekurt@mersin.edu.tr

ABSTRACT

This paper reports the statistical and probabilistic reasoning of young children in terms of randomness, variability, and data representations in the context of informal inferential reasoning (IIR). Using the IIR approach, a task was designed and conducted one-on-one with 28 children aged 5 to 6 years old, in a case study setting. The researcher used a voice recorder during interviews, took photos, and recorded field notes. The data were analyzed according to the principles of informal inferential reasoning, which are generalizing beyond the data, using data as evidence for generalizing, and using probabilistic language to acknowledge the uncertainty embedded in the data. The findings indicate that young children are capable of making informal inferences from a sample space, describing event probability, and constructing bar graph and pie chart data representations.

Keywords: *Statistics education research; Probabilistic reasoning; Statistical reasoning; Informal inferential reasoning; IIR; Bar graph; Pie chart; Randomness; Sample*

1. INTRODUCTION

When I was discussing this study with Professor Katie Makar at the International Conference on Teaching Statistics in 2018, she asked me a question. "Why do we need a bar graph representation (she was showing bars and counts for categories)?" "We expect a bar graph to make data more comprehensible and appreciable, as well as provide information about the data," I replied. The ensuing conversation made me wonder about other possibilities and motivated me to develop the research reported in this paper. Makar's perspective added a really nice influence on this study.

The very earliest years of a child's life are filled with experiences that teach them what is possible, what is not possible, and the degree to which something is possible. During this time, children begin to learn about data handling and analysis, and they develop an understanding of certain statistical and probabilistic concepts such as uncertainty, randomness, chance, or sample space (Nikiforidou, 2018). Children's ability to handle data, however, is largely undervalued and ignored (McPhee & Makar, 2018). Children are faced with working with data at all levels of their education, which means they need to develop data sense early on. In the same way that they must develop number sense (Perry & Docket, 2008), children need to advance their ideas about data from an early age (e.g., Leavy, 2008; Shaughnessy, 2010; Watson & Moritz, 2000). While reading, forming, and interpreting data, they must also construct and interpret data representations in the early stages (Estrella, 2018). Friel et al. (2001) defined an essential idea of graph comprehension as the ability to read graphs to draw meaning from them in relation to data represented. There are few studies, however, that have explored children's interaction with data representations and it is unclear how the development of graph comprehension in a statistical context is related to the creation of a data representation (Estrella, 2018).

Learning and teaching of data representations in school settings might be hindered because of the lack of focus on how to work with data representations (diSessa, 2004). There are many curricular guidelines to introduce graphs in the early years of education in many countries (Batanero et al., 2018). For the most part, tasks used currently in statistics classes only expose children to calculations based

Statistics Education Research Journal, 22(2), Article 4. https://doi.org/10.52041/serj.v22i2.434

on the data and ignore why we need data representations. Therefore, it is important to analyze how young children interact with data representations.

One of the most important steps in data analysis is making inferences, which means making judgments about the data set, within a statistical context. We need data representations for making those inferences. An answer to a statistical problem can be arrived at through the process of inference, which involves the interaction and coordination of real-world information with data structures and representations (Lehrer et al., 2007). However, repeated observations reveal an unpredictable pattern for the outcomes of uncertain events, and probability provides a mathematical basis for the representation of uncertainty (Moore, 1990). While there is "certainty" in mathematical arguments, children must deal with "uncertainty" in statistical arguments to reason meaningfully (Ben-Zvi, 2018; Shaughnessy, 2010). Young children use probabilistic reasoning to decide whether certain events are likely or unlikely (Nikiforidou, 2018). Children must experience the uncertain nature of data and must be exposed to tasks that make them identify the variability in data while making inferences that forecast the future or draw conclusions from the past (Borovcnik, 2011). Therefore, children must develop their probability sense, which can be explained as developing a probabilistic language (Perry & Docket, 2008). Although, there is no international agreement on when to start teaching probability (Groth et al., 2021), it is problematic to put off teaching probability concepts (Greer, 2014). Children's probability sense should be promoted while helping them realize the variability in a data set and interrogate that variation (Reading & Shaughnessy, 2004) to make inferences about the data set, while constructing data representations.

One strategy for improving young learners' statistical and probabilistic reasoning is to provide rich learning tasks that help them see the connection between variation, chance, and uncertainty. This can be accomplished by providing appropriate contexts for emphasizing fundamental stochastic concepts, such as randomness and chance. Children's intuitions should also be developed since probability requires more than the application of procedures to find solutions, "[r]ather, it requires a way of thinking that is genuinely different from that required by most school mathematics. In learning probability, students must create new intuitions" (Fischbein & Schnarch, 1997, p. 104), which means establishing "a feeling of obviousness, of intrinsic certainty" (p. 96). As a strategy to enhance children's statistical and probabilistic reasoning, research with students engaged in statistical and probability tasks has been proposed. (Shaughnessy, 1992). These suggestions pave the way for research on children's informal inferential reasoning within a context.

Recent research also suggested analysis of data visualizations focus "not just *beneath* the data, but also *beyond* the data, towards thinking and reasoning inferentially with data" (Makar & Rubin, 2009, p. 83). As well. there is a need to analyze children's probabilistic and statistical reasoning in contextual settings (Langrall et al., 2006). According to Makar (2018), young children may be exposed to statistical and probabilistic ideas before they are formally taught. However, instruction often focuses exclusively on procedural comprehension using computational skills, rather than conceptual comprehension (Sorto, 2006). Additionally, the activities to which children are exposed encompass a greater number of directions than expected, hindering children's ability to think creatively, even when they are capable (McPhee & Makar, 2018). Overall, one could argue that this method of instruction reduces statistics and statistical reasoning to a limited mode of comprehension, which is underpinned by the notion that "any phenomenon can be captured by a bar chart" (Ben-Zvi & Sharett-Amir, 2005, p. 1).

The purpose of the study reported in this paper was to examine young learners' statistical and probabilistic reasoning using a task that was designed to foster their data and probability sense while forming data representations. The aim was to provide the opportunity for the children to connect randomness and uncertainty, with a particular emphasis on informal inferential reasoning. The research questions for this study are:

1) How do young children make informal inferential reasoning through a task designed to improve their data and probability sense while creating a bar graph and a pie chart representation?
2) How do young children build evidence through the task?
3) How do young children reason about sample space and probability of an event, while forming a bar graph and a pie chart representation?

1.1. THEORETICAL FRAMEWORK

Informal inferential reasoning (IIR) was defined as "the way in which students use their informal statistical knowledge to make arguments to support inferences about unknown populations based on observed samples" (Zieffler et al., 2008, p. 44). That is, IIR does not use formal hypothesis testing to evaluate arguments based on collected data; rather, it allows students to reason intuitively about the data. As a result, IIR aims to provide children with a basic understanding of fundamental statistical concepts without looking deep into their formal applications (Makar & Rubin, 2009). Ziefller et al. (2008) defined the IIR framework as consisting of three major components: (1) making inferences without using formal statistical procedures, (2) utilizing prior knowledge (formal knowledge of descriptive statistics, but informal knowledge of making inferences about different samples, and use of statistical language), and (3) making arguments, which include claims about populations while utilizing evidence gathered from samples belonging to those populations (Zieffler et al., 2008, p. 45). Makar and Rubin (2009) used the IIR framework to "broaden accessibility to inferential reasoning with data" by emphasizing three fundamental principles: (1) generalization that extends beyond describing the given data through predictions, estimations, and the like, (2) the use of data as evidence for making a generalization, and (3) use of probabilistic language in describing the generalization drawn and articulating the level of certainty (p. 85). The authors viewed these three components as observable abilities that comprise young children's informal inferential reasoning.

Together with IIR, randomness, variation, distribution, and expectation are the big ideas in elementary statistics and probability education (Watson et al., 2018). Researchers try to connect these important concepts to give purpose to learning about statistics at the elementary level of the curriculum. The second critical big idea is randomness, which is defined as "a phenomenon in which the outcome of a single repetition is uncertain, but a regular distribution of relative frequencies exists over a large number of repetitions" (Watson & Fitzallen, 2019). Watson and Fitzallen (2019) extended the connection between randomness and variation (defined as observed differences) and expectation (defined as summarizing data while incorporating variation, for example, when determining the chance of an event), observing that children frequently ignore the expectation while overthinking over the variation when determining the chance of an event. As a result, students must first grasp the relationship between variation and expectation before diving into randomness. Finally, distribution serves as the lens through which variation is identified in data representations. Watson and Fitzallen (2019) suggested that these are the overarching concepts on which statistics and probability education should center learning efforts.

Most of the research on children's understanding of statistics has focused on their everyday probabilistic judgments rather than on fundamental statistical concepts, such as data handling or distribution (Leavy, 2008), or on the relationship between some statistical and probabilistic concepts. Because the literature undervalues uncertainty, variability, and distribution as important concepts, particularly at the elementary level, and underestimates the value of their teaching, probability concepts have been removed from some curricula (Groth et al., 2021). For instance, Turkey's elementary mathematics curriculum has neglected key probability concepts in the first seven grades of schooling for nearly nine years (Ministry of National Education.[MoNE], 2004; 2005; 2013). Although Turkey was one of the first countries to incorporate statistics and probability into elementary mathematics curricula, the concepts of likely and unlikely events are now introduced in eighth grade, in contrast to countries such as Australia or the United States (Argün et al., 2010; Batanero et al., 2018; Makar, 2018).

The most frequently used data representations are bar graphs and pie charts, which can be presented in a variety of ways (Spence & Krizel, 1994). Additionally, these graphs are one of the first data representations taught in data handling. However, their inclusion is at the level of descriptive statistics, focusing exclusively on the frequencies and categories visible on them. Even so, the courses in Turkey addressed the differences in frequency distributions, which should not be the primary focus of statistics instruction. Rather, students should be introduced to probability using probabilistic language while collecting data with inherent variability, focusing on expectation beyond the data, and then organizing and inferring about the data collected.

The purpose of this study was to make the connection between probabilistic and statistical reasoning visible through experimenting with different colored balls drawn from a bag. The task was intended to demonstrate the competencies of the young children about the concepts of sample space and probability of an event in terms of probabilistic reasoning, as well as to focus on children's data sense about describing data and organizing and reducing data in terms of informal inferential reasoning. One of the parts of the tasks entailed drawing a ball from a box twenty times and recording the results on a sheet of paper using the colored cards provided. Additionally, the task included several questions for children, such as, "What color ball do you expect drawing?", "How can we record the outcomes (while drawing the balls)?", and "How can we organize the data records?"

2. METHOD

This study employed qualitative methods based on the paradigms of an exploratory case study design, as the primary objective was to investigate children's experiences while completing the task, which was developed from an IIR perspective (Yin, 2014). The study's main objective was to explore those experiences within the framed theoretical basis summarized above. Case study entails the analysis of in-depth data gathered through specific cases (Patton, 2014). This method allowed for an in-depth analysis of the rich data collected to determine how young children reason statistically and probabilistically concerning IIR. The study described here used the IIR theoretical framework to descriptively analyze young children's statistical and probabilistic reasoning. A goal of the study was to generate novel approaches to young children's statistics and probability reasoning, resulting in the development of learning tasks that effectively foster young children's IIR.

Participants. Twenty-eight young children (25 children aged 6-years-old, and three 5-years-old) were enrolled in this study from four different classes at a public kindergarten school on my university campus. Academic and administrative staff at the university send their children to this kindergarten school. Most parents whose children were enrolled were public servants, such as teachers, doctors, and police officers. Because I, as a researcher, was unfamiliar with the children, the class teachers selected the children who would likely be interested in participating in the study.

Data collection procedure. The research task completed by the children was carried out in a small room that was used for chess classes. The school's directors offered the chess classroom in response to my request for a quiet place for voice recording. I introduced myself and asked for each child's name in the first minutes of the interview, while wearing much more colorful clothing than I normally do, to appear friendly and to make the participants feel comfortable. The participants were interviewed one-to-one. Each participant was interviewed across a table while holding the box of balls and using a mat and colored cards. The interviews were audiotaped. It took about 20 minutes for each participant to complete the tasks. By using a pre-prepared interview protocol, I was able to ask the same questions of each of the participants during the interviews. Each participant was asked the questions, and his or her oral responses were audio-recorded. Throughout the interview, the researcher took several photographs to document each participant's progress through the activities. The children were asked a variety of questions, such as, "What color ball do you expect to draw?", "How can we record the outcomes (while drawing the balls)?", and "How can our records (the colored cards) be organized?"

Data collection tool and data analysis. The children's data collection tool was a task that involved drawing balls from a box. Given that one of the study's objectives was to investigate probabilistic and statistical reasoning with an IIR approach, the task required the children to create a random sample of drawings from a box of colored balls. It was comprised of two parts with questions, which were organized with the assistance of IIR to establish a connection among the concepts of sample space, probability of an event, and bar graph and pie chart representations, in terms of probabilistic and statistical reasoning. The design of the task was based on the premise that children can list the possible outcomes of a one-stage experiment (e.g., Borovcnik & Bentz, 1991; English, 1993; Piaget & Inhelder, 1975). Also, according to the literature on probability of an event, children's understanding of probability is based largely on their understanding of part-whole relationships. (e.g., Piaget & Inhelder, 1975).

I began by showing the children a box containing six balls (three yellow, two blue, and one red). I then asked the children to draw a ball from the box twenty times, with replacement. Following their viewing of the color, the researcher asked them to record the colors drawn using the similarly colored rectangular cards on a mat. I then inquired as to whether there was another way to arrange the cards to see the least or greatest number of the colors drawn from the box. The second part of the task involved drawing eight balls from the box, which contained only red and yellow balls. At this point, cards resembled sectors of a circle, eight of which formed a circle). At the end, I inquired whether there was a need to organize the cards to better reason about the data, that is, create a data representation to distinguish data drawn. Throughout this task, I examined young children's informal inferential reasoning about sample space by asking them, "Which colors could appear while drawing the balls from the box?" I asked some questions regarding the probability of an event which were, "Which color is most likely to appear after twenty drawings?", and "Which color is least likely to appear?"

Interview data were transcribed verbatim from audio recordings and data were analyzed descriptively. With the help of several photographs taken during the interview, the responses of each participant to the interview questions were classified and organized. The gathered data analyzed holistically to examine the participants' statistical and probabilistic reasoning with an IIR perspective. The researcher classified the data into four categories: (1) how sample space was perceived, (2) how probability of an event was explained, (3) how the data obtained in the task was organized, and (4) how and to what extent the data representation looked (to a bar graph or a pie chart). The participants' responses were analyzed to determine how many of them represented the data as expected and how many responded appropriately to interview questions.

The task dealt with the statistical reasoning of young children in terms of collecting, describing, and organizing data, with the question "How do we organize the cards to understand what the data tells us?" Additionally, the task provided for the representation and interpretation of data following the completion of the drawing balls and the collection of cards on the mat with the matching colors. Finally, each child was expected to create a bar graph (using colored cards to represent 20 balls) and a pie chart (using colored circle sectors to represent eight balls). While interpreting the data, participants were expected to recognize the frequency of the colors. As a result, the task aided in the exploration of young children's statistical reasoning.

3. FINDINGS

This study's findings are divided into two parts. First, the data were quantitatively summarized in Table 1. Second, the data were summarized according to the task-related issues, with examples provided in the form of photographs. For reporting purposes, each child was assigned an alphanumeric identifier to maintain privacy and anonymity (e.g., P52). All 28 children completed the task and demonstrated competence in organizing data using bar graphs and pie charts. The following table summarizes the number of correct responses, the number of children who grouped the data collected, and the number of children who formed bar graphs and pie charts.

Table 1. Quantitative results obtained from the task implementation

Bar Graph		Pie Chart		Sample Space	Probability of an Event
Grouped	Formed	Used Center	Formed		
14	5	24	16	22	10

After making 20 draws from the box, each participant was asked to rearrange all the cards on his/her board. Participants were expected to group/categorize the cards according to their colors and arrange them side by side, resulting in a bar graph. As a result, I labeled the responses either "grouped" or "formed". In relation to the second part of the task, after eight draws, each participant was asked to reorganize the cards on her/his board to forma pie chart. Children were expected to first identify a center for placing the cards on the mat. They were then expected to organize the cards by color whether they arranged identically colored cards side by side, creating a pie chart. As a result, I labeled the responses "used center" and "formed".

In terms of sample space, I asked each child if she/he could draw a green ball from the box. A correct response indicated that the child realized the potential outcomes of drawing a green ball from the box was zero. When I asked a child if she/he could draw a green ball from the bag, she/he replied, "No." When I inquired as to why, the child stated, "There is no green ball." According to Table 1, 21 children correctly answered this question, which is a high number. Concerning the probability of an event, a correct response indicated that the child explained why yellow appeared to be the most prevalent color based on the number of yellow balls in the bag. Ten children reasoned about the probability of the event.

3.1. SAMPLE SPACE

Most of the children (*n* = 22 of 28) responded that they would draw red, yellow, or blue balls from the bag, but were unable to draw green or pink balls. When the researcher questioned as to the reason, they stated that the green ball was not in the bag. The rest of the children stated that they were unaware of how to draw a green ball. They stated that they are unaware of the reason. Several of them responded that it might have something to do with magic. The following are some responses to the question, "Do you think we can draw a green ball from the bag?"

P3: I don't know.
P5: Yes, I can.
P15: Yes, I can draw a green ball, if this bag has magic.

Some examples for the question, "Why can't we draw a green ball from the bag?" were:

P6: Because there is not any green ball.
P22: There is no green ball here.
P7: I don't know. (Although, she responded, "We can't draw a green ball.").

Based on the findings here, it could be claimed that young children intuitively know about the sample space. Twenty of the 22 children could reason about the possible outcomes.

3.2. PROBABILITY OF AN EVENT

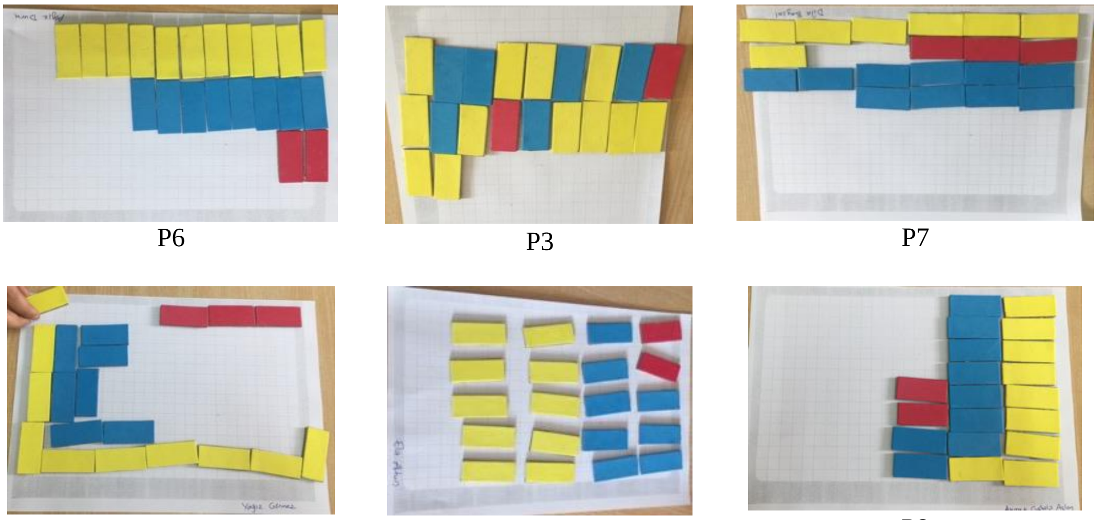
After completing all draws, either 20 or eight balls, the researcher posed the question, "Why did we end up with more yellow balls?" Alternatively, "Why did the red ball occur less?" Nearly half of the children (*n* = 10 of 28) responded that they drew the most yellow balls because the bag contained the most yellow balls. The following are some of the children's responses:

P18: Because there are 3 of the yellow.
P17: I don't know.
P15: Because I closed my eyes and chose [the balls]. Even if the balls weren't yellow, there would never have been so many yellows.
P10: There are a few red (balls) because. There are 3 yellow balls

The remainder of the participants indicated that they were unsure. When the researcher repeated the question, "Is this related to the number of balls in the box?", they responded that they were unsure. P14, for example, explained that she "primarily caught the yellow ones," while focusing on the outcomes she obtained.

3.3. STATISTICAL REASONING – BAR GRAPH

Prior to drawing the balls from the box, the children were told to record the colors drawn by matching them with the colored cards, which were to be collected on the mat. Finally, the researcher requested that they rearrange the cards so that they could more easily see the number of each color without counting. Half of the children (*n* = 14 of 28) grouped the cards by color, and five reorganized them to represent a bar graph. Examples of the children's representations are presented in Figure 1.

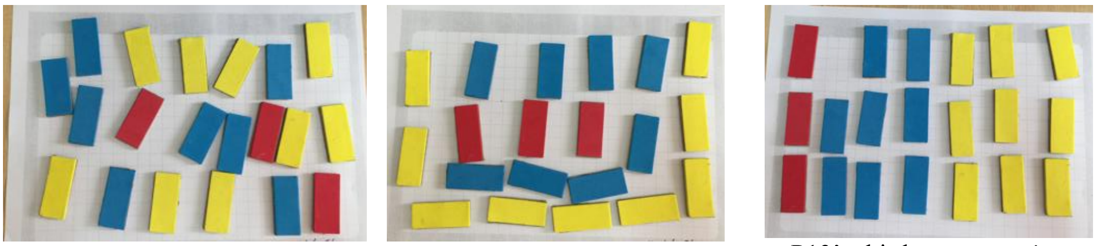


P6 P3 P7

P10 P2 P8

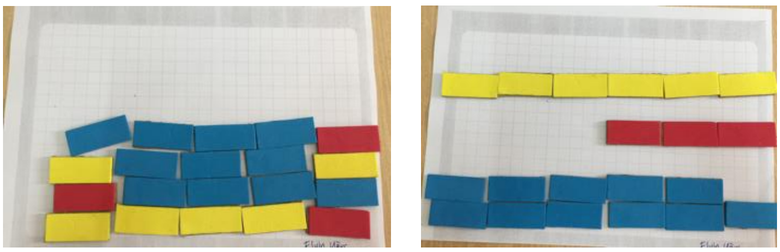
Figure 1. Examples of children's bar graph representations

Another observation concerns the "back and forths" between data and graph (or shape, display, or representation). When the children were unable to answer questions about analyzing the data or interpreting the data, they often to returned to their representation and reorganized the data, sometimes three or four times. For instance, P13 (Figure 2) and P14 (Figure 3) constructed a variety of organizations. Overall, the children were able to respond to the questions in this part of the task and showed their performance in bar graph representations.



P13's first representation P13's second representation P13's third representation

Figure 2. P13's bar graph representations



P14's first representation P14's second representation

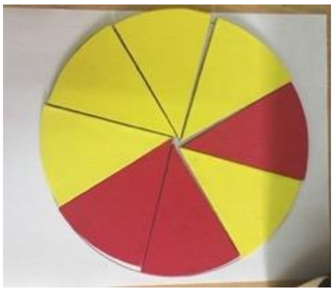
Figure 3. P14's bar graph representations

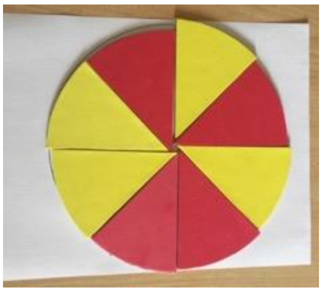
3.4. STATISTICAL REASONING – PIE CHART

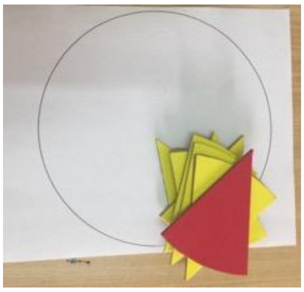
To create a pie graph, children were expected to locate the center point, arrange the cards on the circle, and finally group the cards by color. It appeared the children found it easier to create a pie chart than it was to create a bar graph. Like the bar graph task, the children were asked: "How can we easily determine which color is the most and which color is the least?" Unexpectedly, most children (n = 24 of 28) were aware of the circle's center. Furthermore, they likened the cards to pizza slices. Many of them asked, "Will we play a pizza game?" Sixteen of them further classified the cards by color. Figure 3 contains several examples of children's pie chart representations. Unlike the children who struggled with bar graphs, more than half (*n* = 16 of 28) of the children formed a pie chart and could explain that when the colors were side by side, the groups could be easily compared in terms of magnitude.


P23


P3


P8


P5


P12

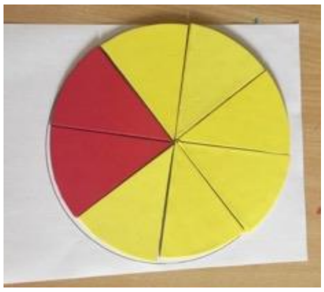
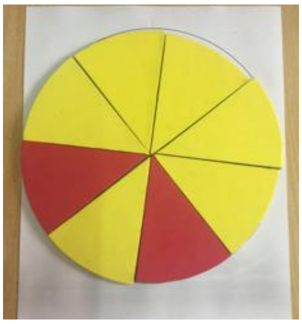
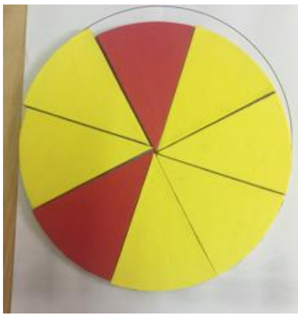

P4

Figure 4. Examples of children's pie graph representations

To make the children think about their pie graph representation, the question, "How can we re-organize the cards to show which color is the most and which is least?" was asked. The following examples illustrate several of their subsequent representations (Figure 5 & 6).


P24's first representation


P27's first representation

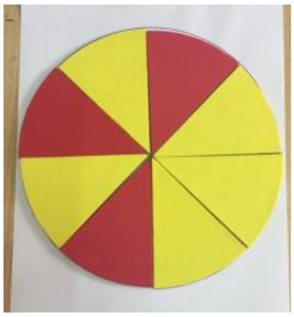
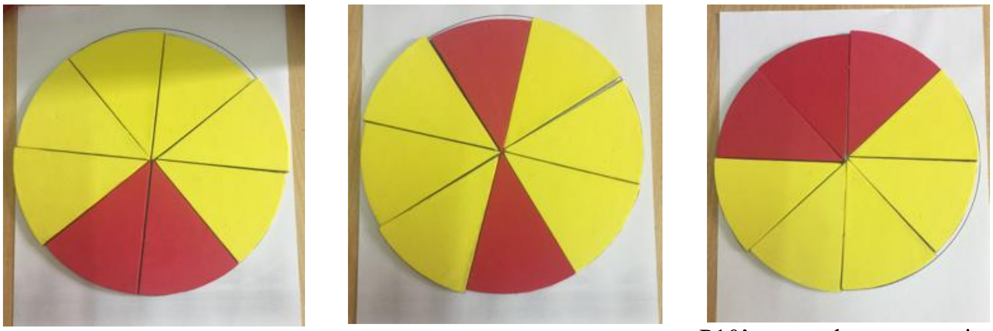

P10's first representation

Figure 5. Examples of children's pie chart representations



P24's second representation P27's second representation P10's second representation

Figure 6. Examples of children's pie chart representations

Additionally, the pie chart representations reflected some pattern thinking. Some children arranged the cards in a repeating pattern, as illustrated in the examples in Figure 7.

P15 P27 P22

Figure 7. Pattern examples of children's pie graph representations

4. DISCUSSION AND CONCLUSION

The overall findings are discussed with an emphasis on three principles of informal inferential reasoning concerning the sample space, probability of an event, and bar and pie graph representations. The findings provided an explanation for how young children's probabilistic and statistical reasoning develops while creating data representations through informal inferential reasoning. Additionally, how young children construct evidence during the task is discussed.

4.1. SAMPLE SPACE

Shaughnessy (1992) asserted that the concepts related with probabilistic reasoning such as sample space, probability of an event, or conditional probability, were not studied together and there is not a consensus regarding their influence on children's probabilistic reasoning. Accordingly, how the instruction of these concepts should be delivered in which environment using which tasks are not clear. This study proposes that children can be observed within a probability context. From this perspective, this study discusses significant findings.

Six children were unable to explain what might happen if they drew a ball from the bag. They mostly replied that they were unaware of the reason or that it was due to magic. Two children were unable to explain the reasoning despite being aware of the possible outcomes, implying that they were incapable of reasoning about this. Twenty children in total could list the possible outcomes. These findings could be explained by the fact that young children intuitively understand the sample space. When children participate in such an experiment, it is possible to assert that they are aware of the possible and impossible outcomes. They were able to use probabilistic language with an emphasis on possible and impossible (Kazak & Leavy, 2018). Additionally, they were able to explain the impossibility with the absence of green or pink ball in the box.

As Nikiforidou (2018) concluded, subjective intuitions can be used to grasp probabilistic concepts. This conclusion implies that probability experiments should be conducted with young children in a classroom setting. These experiments may also reveal information about uncertainty and randomness, which could be discussed further here, given the drawing ball experiment's findings regarding young children's uncertainty comprehension.

Borovcnik (2011) asserted that the relationship between data and probability undervalues probability as a statistical sub-concept. For this study, the participants generated their data set by drawing balls from a box, which was unique for them. The task could easily be replaced by simply presenting a suitable data set. However, the task based on an experiment, served to highlight the basic concepts of data and probability within a more meaningful context. Additionally, the assumption was made that young children could intuitively construct the link between the two concepts through data generation and interpretation.

4.2. PROBABILITY OF AN EVENT

Regarding the concept of the probability of an event, nearly half of the children (n = 10 out of 22) used subjective thinking and the understanding of part-whole relationships (Piaget & Inhelder, 1975). This finding showed that they had difficulty explaining the reason for drawing yellow colors more in both parts of the task. Other participants explained the reasons for drawing yellow the most, such as their unluckiness and loving yellow the most as a color. While they were drawing the balls, they expected mostly a color pattern, for example, Yellow, Red, Blue, Yellow, Red, Blue. This color pattern expectancy was also apparent in their efforts to form a bar graph of the first 20 draws. This finding could be explained by their need "to make sensible decisions in situations of uncertainty" (Borovcnik & Peard, 1996, as cited in Perry & Docket, 2008, p. 95). Therefore, although this task could be an effective example of making children think about uncertainty, more than half of the children could not make a connection between the experiment results and the uncertain nature of such an experiment. Besides, this conclusion strengthened the call to provide children with instances where they can carry out probability experiments with a focus on chance (Lopes & Cox, 2018). With sufficient experience, children could start to think about randomness.

The findings obtained through the question of "Why did yellow appear the most?" could also be discussed regarding likely or unlikely events, since children pointed out that the presence of yellow, blue, and red balls in the box made the drawing of one of those colors a likely event and drawing a green or pink ball an unlikely event. They explained that they came up with the colors because they were drawn to them since the colors were in the box. Moreover, they were also aware that drawing a yellow ball was more likely than drawing a red or blue ball, since they explained this result was because there were more yellow balls in the bag. It could be concluded that children think that "the more something is, the more often we meet," in general. This explained their understanding of the likelihood of an event as likely, less likely, or more likely. However, it was not possible to conclude whether the children were aware that this experiment had random results.

As Nikiforidou (2018) concluded, children from an early age begin to think probabilistically. They may develop a prior understanding of probability if given a probabilistic context in a preschool setting. Thus, a task like the one used in this study could serve as a foundation for children's intuitive probabilistic understanding and provide teachers with pedagogical ideas.

4.3. BAR GRAPH REPRESENTATION

The most frequently seen graphs are bar graphs and pie charts, which can be displayed in a variety of ways and with a variety of shapes (Spence & Krizel, 1994). In statistics education in Turkey, these graph-types are often the first concepts taught. As a result, I assigned tasks such as creating a bar graph and a pie chart. The tasks used in this study focused on organizing and reducing the data. Nevertheless, I analyzed two distinct data visualization techniques: bar graphs and pie charts.

The findings in this section indicated that, while all the children were unable to create a regular appearance for the bar graph at the conclusion, they grouped the cards according to their colors.

Additionally, they stated, “I can easily determine which is the most and which is the least.” P9 (Figure 8), for example, organized the cards similarly to P8 (Figure 1), stating, “I put the yellow ones here, and the blue ones here. It might be better to put in double as there is a lot of yellow.” This means that when the purpose of the display is to show the most and least one, children create the bar graph representation. Thus, the children who organized the cards according to their colors, the task’s objective could easily be said to have been accomplished, creating a bar graph. Graphs or data visualizations serve the purpose of illustrating an important concept. As Friel et al. (2001) stated, each graph or data representation has its own language that enables it to perform its intended function.

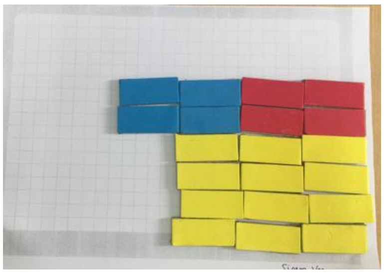


Figure 8. P9’s representation of bar graph

This research demonstrates that young children can form data representations when they have meaningful purposes without any boundaries. They can explain the story the data has to tell with language unique to them. Therefore, it could be concluded that young children can start to develop graph comprehension early in their education, which Friel et al. (2001) explained as “graph readers’ abilities to drive meaning from graphs created by others or themselves” (p. 132).

4.4. PIE CHART REPRESENTATION

The data indicate that more children could form and group the colored cards to form a pie chart compared to the ones to form a bar graph. It may be because of identifying the center of the circle before beginning the task or likening the task to a pizza. Another reason could be the number of cards or the number of colors used in this task. There were eight draws in this part, which is fewer than the ones in the bar graph part. This could be the reason for constructing the representation easily.

Children’s pattern recognition can also be seen in this task, as the examples in Figure 8 demonstrate that children used pattern thinking to organize the cards. According to Friel et al. (2001), graph comprehension is also associated with pattern perception. As a result, children pattern the visualized the information collected on the mat. They were placing the cards on the mat as they drew the balls from the box. The circle sector cards were primarily used by children as if they were pizza slices. As a result, they were not observed to have difficulty creating a pie chart. Placing the cards in drawing order may compel them to think of them in a pattern, impairing their ability to organize the cards according to color. According to Friel et al. (2001), there is no obvious connection between pattern recognition and graph comprehension. Due to the small number of draws, we cannot conclude that children’s pattern recognition ability prevents them from creating a pie chart. This part of the task could be repeated with additional draws to demonstrate the relationship between the two ideas.

4.5. OVERALL DISCUSSION

Makar (2018) asserted that students may become acquainted with statistical and probabilistic concepts prior to formal instruction of these concepts. The findings of this study, which included young children exposed to a task in which statistical and probabilistic concepts were not previously introduced, demonstrated that children could complete the task with an intuitive and informal understanding of statistics and probability concepts. As a result, the study’s findings indicate that young children can complete tasks designed from the perspective of informal inferential reasoning (Makar & Rubin, 2009).

Another conclusion is that children could use this task to develop an intuitive understanding of the big ideas of statistics outlined by Watson et al. (2018). For instance, when children drew balls from the bag, they encountered randomness. Although they anticipated a pattern or magic from the draws, they observed that neither occurred. The participants in this study were also able to consider randomness during an experiment, as they anticipated the colors of the balls in the box prior to drawing out the balls. From this viewpoint, the task at hand incorporates the concept of variation as well (Watson, 2009). The children began to consider variation intuitively within a likely range, and thus encountered an aspect of variation that aided the construction of their data representations. The children were free to organize their data in any way they desired, most notably in their bar graph representations. This resulted in the conclusion that children can engage with the big ideas of statistics in a way that allows them to think about them informally (Watson, 2009) in meaningful and contextual settings (Langrall et al., 2006).

The study presented here contributes to the early statistics and probabilistic education literature by demonstrating that when children are given meaningful teaching environments, such as well-designed course tasks and experiments that require them to collect, organize, reduce, and infer data, they can informally develop an understanding of and reason about fundamental statistical and probabilistic concepts that are also evident in everyday life. Although the early stages of education do not always include explicitly teach probability concepts, young children's experiences embedded in productive tasks that foster institutions have a significant impact on their probabilistic reasoning (Nikiforidou, 2018). One could argue that such tasks, including experiments, expose young children to the concept of randomness. Additionally, such tasks help children develop their sense of expectation.

ACKNOWLEDGEMENTS

I express my gratitude to Professor Katie Makar for her insights that prompted me to consider the study's task and for the perspective she brought to early statistics and probability education.

REFERENCES

Argün, Z., Arıkan, A., Bulut, S., & Sriraman, B. (2010). A brief history of mathematics education in Turkey: K-12 mathematics curricula. *ZDM Mathematics Education*, *42*(5), 429–441. https://doi.org/10.1007/s11858-010-0250-0

Batanero, C., Arteaga, P., & Gea, M. M. (2018). Statistical graphs in Spanish textbooks and diagnostic tests for 6–8-year-old children. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds.), *Statistics in early childhood and primary education: Supporting early statistical and probabilistic thinking* (pp. 163–180). Springer. https://doi.org/10.1007/978-981-13-1044-7_10

Ben-Zvi, D. (2018). Foreword. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds.), *Statistics in early childhood and primary education: Supporting early statistical and probabilistic thinking* (pp. vii–viii). Springer.

Ben-Zvi, D., & Sharett-Amir, Y. (2005). How do primary school students begin to reason about distributions? In M. Pfannkuch (Ed.), *Reasoning about distribution: A collection of current research studies*. Proceedings of the fourth international research forum on statistical reasoning, thinking, and literacy (SRTL-4), Auckland, New Zealand (pp. 2–7).

Borovcnik, M. (2011). Strengthening the role of probability within statistics curricula. In C. Batanero, G. Burrill & C. Reading (Eds.) *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 71-83). Springer. https://doi.org/10.1007/978-94-007-1131-0_11

Borovcnik, M. G., & Bentz, H. J. (1991). Empirical research in understanding probability. In R. Kapadia & M. Borovcnik (Eds.) *Chance encounters: Probability in education* (pp. 73–105). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-011-3532-0_3

diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction, 22*(3), 293-331. https://doi.org/10.1207/s1532690xci2203_2

English, L. D. (1993). Children's strategies for solving two- and three-stage combinatorial problems. *Journal for Research in Mathematics Education,* *24*(3), 255–273. https://doi.org/10.5951/jresematheduc.24.3.0255

Estrella, S. (2018). Data representations in early statistics: Data sense, meta-representational competence and transnumeration. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds), *Statistics in early childhood and primary education. Early mathematics learning and development* (pp. 239–256). Springer. https://doi.org/10.1007/978-981-13-1044-7_14

Fischbein, E., & Schnarch, D. (1997). Brief report: The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education*, *28*(1), 96–105. https://doi.org/10.5951/jresematheduc.28.1.0096

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*(2), 124–158. https://doi.org/10.2307/749671

Greer, B. (2014). Commentary on perspective II: Psychology. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic thinking: Presenting plural perspectives* (pp. 299–309). Springer. https://doi.org/10.1007/978-94-007-7155-0_16

Groth, R. E., Austin, J. W., Naumann, M., & Rickards, M. (2021). Toward a theoretical structure to characterize early probabilistic thinking. *Mathematics Education Research Journal*, *33*(2), 241–261. https://doi.org/10.1007/s13394-019-00287-w

Kazak, S., & Leavy, A. M. (2018). Emergent reasoning about uncertainty in primary school children with a focus on subjective probability. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds.), *Statistics in early childhood and primary education: Supporting early statistical and probabilistic thinking* (pp. 37–54). Springer. https://doi.org/10.1007/978-981-13-1044-7_3

Langrall, C., Nisbet, S., & Mooney, E. (2006). The interplay between students' statistical knowledge and context knowledge in analyzing data. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education.* Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS7), Salvador, Brazil. https://iase-web.org/documents/papers/icots7/2A3_LANG.pdf?1402524964

Leavy, A. (2008). An examination of the role of statistical investigation in supporting the development of young children's statistical reasoning. In O. Saracho & B. Spodek (Eds.), *Contemporary perspectives on mathematics in early childhood education* (pp. 215–232). Information Age Publishing.

Lehrer, R., Kim, M. J., & Schauble, L. (2007). Supporting the development of conceptions of statistics by engaging students in measuring and modeling variability. *International Journal of Computers for Mathematical Learning*, *12*(3), 195–216. https://doi.org/10.1007/s10758-007-9122-2

Lopes, C. E., & Cox, D. (2018). The impact of culturally responsive teaching on statistical and probabilistic learning of elementary children. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds.), *Statistics in early childhood and primary education: Supporting early statistical and probabilistic thinking* (pp. 75–88). Springer. https://doi.org/10.1007/978-981-13-1044-7_5

Makar, K. (2018). Theorising links between context and structure to introduce powerful statistical ideas in the early years. In A. Leavy, M. Meletiou-Mavrotheris & E. Paparistodemou (Eds.), *Statistics in early childhood and primary education* (pp. 3–20). Springer. https://doi.org./10.1007/978-981-13-1044-7_1

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, *8*(1). 82–105. https://doi.org/10.52041/serj.v8i1.457

McPhee, D., & Makar, K. (2018). Early childhood experiences in informal inferential statistics using an inquiry approach. In M. A. Sorto, A. White & L. Guyot (Eds.), *Looking back, looking forward.* Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10), Kyoto, Japan. https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_4B1.pdf?1531364263

Ministry of National Education. (2004). *Ilkogretim matematik dersi ogretim programi ve kilavuzu: 1–5.* sınıflar [Elementary education mathematics curriculum and manual: Grades 1–5]. Devlet Kitapları Mudurlugu.

Ministry of National Education. (2005). *Ilkogretim matematik dersi ogretim programı ve kılavuzu: 6–8.* sınıflar [Primary education mathematics curriculum and manual: Grades 6–8]. Devlet Kitapları Mudurlugu.

Ministry of National Education. (2018). *Matematik dersi ogretim programı: İlkokul ve Ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8.* Sınıflar [Mathematics curriculum: Primary and elementary Grades 1, 2, 3, 4, 5, 6, 7, ve 8.]. Devlet Kitapları Mudurlugu.

Moore, D. S. (1990). Uncertainty. In L. A. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95–137). National Academy Press.

Nikiforidou, Z. (2018). Probabilistic thinking and young children: theory and pedagogy. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds.), *Statistics in early childhood and primary education* (pp. 21–34). Singapore. https:doi.org/10.1007/978-981-13-1044-7_2

Patton, M. Q. (2014). *Qualitative research and evaluation methods: Integrating theory and practice*. SAGE Publications.

Perry, B., & Dockett, S. (2008). Young children's access to powerful mathematical ideas. *Handbook of international research in mathematics education* (pp. 81–112). Routledge.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children* (translated by L. Leake, Jr., P. Burrell & H. D. Fischbein). W. W. Norton.

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). Springer. https://doi.org/10.1007/1-4020-2278-6_9

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 465–494). Macmillan.

Shaughnessy, M. (2010). *Statistics for all: The flip side of quantitative reasoning.* National Council of Teachers of Mathematics.

Sorto, M. A. (2006). Identifying content knowledge for teaching statistics. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education*. Proceedings of the Seventh International Conference on Teaching Statistics, Salvador, Brazil. http://www.stat.auckland.ac.nz/~iase/publications/17/C130.pdf

Spence, I., & Krizel, P. (1994). Children's perception of proportion in graphs. *Child Development, 65*(4), 1193–1213. https://doi.org/10.1111/j.1467-8624.1994.tb00812.x

Watson, J. M. (2009). The influence of variation and expectation on the developing awareness of distribution. *Statistics Education Research Journal, 8*(1), 32–61. https://doi.org/10.52041/serj.v8i1.456

Watson, J. M., & Fitzallen, N. (September 19, 2019). Building understanding of randomness from ideas about variation and expectation. *Statistics Teacher*. https://www.statisticsteacher.org/2019/09/19/building-understanding-of-randomness-from-ideas-about-variation-and-expectation/

Watson, J., Fitzallen, N., Fielding-Wells, J., & Madden, S. (2018). The practice of statistics. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 105-137). Springer. https://doi.org/10.1007/978-3-319-66195-7_4

Watson, J. M., & Moritz, J. B. (2000). Developing concepts of sampling. *Journal for Research in Mathematics Education, 31*(1), 44–70. https://doi.org/10.2307/749819

Yin, R. K. (2014). *Case study research: Design and methods* (Vol. 5). SAGE Publications.

Zieffler, A., Garfield, J. B., Delmas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal, 7*(2), 40–58. https://doi.org/10.52041/serj.v7i2.469

GAMZE KURT
Faculty of Education
Mersin University
Ciftlikkoy Kampusu
Yenisehir, Mersin/ Türkiye