# Investigating the development of informal inferential reasoning in grade 3 through specially designed tasks

## Investigando el desarrollo del razonamiento informal inferencial en el tercer curso de educación primaria, usando tareas especialmente planificadas

Stavroula Saplamidou

University of Athens, Greece

**Abstract**

The present study falls into the field of informal inferential reasoning, by examining the way in which this form of reasoning is being developed through specially designed tasks. By conducting a case study, we investigated and analyzed one grade 3 student's reasoning using grounded theory. Results indicate student's transition from non-statistical to emergent-statistical forms of reasoning.

**Keywords**: informal inferential reasoning, primary education, case study

**Resumen**

El presente estudio se enfoca en el razonamiento informal inferencial, explorando la manera en que ese tipo del razonamiento se desarrolla a través de tareas particularmente planificadas. Llevando al cabo un estudio de caso, identificamos y analizamos el desarrollo del razonamiento de una estudiante de tercera clase, usando la teoría fundamentada. Los resultados indican la transición de no-estadísticos a emergente-estadísticos tipos del razonamiento.

**Palabras clave**: razonamiento informal inferencial, educación primaria, estudio de caso.

## 1. Introduction and theoretical framework

Reflection on the role of statistics in primary education has led many researchers to conclude that statistics in elementary school is being conceptualized in a very confined way, by placing the emphasis on descriptive statistics and simple data display (Ben-Zvi & Amir, 2005). Consequently, a broader interpretation of statistics in school mathematics is needed (Ben-Zvi, Aridor, Makar, & Bakker, 2012). This need brought inferential tasks to the fore of statistics education, as this kind of tasks provides access to statistics through everyday experiences, in unifying important statistical concepts (Paparistodemou & Meletiou-Mavrotheris, 2008).

Formal statistical inference can be viewed as the result and the reasoning process of creating and evaluating probabilistic generalizations from data, as well as drawing conclusions based on formal probabilistic computations about a wider universe by using the data at hand (Makar & Rubin, 2009). It entails parameter estimation and hypothesis testing (Ben-Zvi, Gil & Apel, 2007), procedures that have been proved to be complex and challenging for school students. As a result, a new goal for statistics education is being created, that of making statistical inference accessible to students of all ages. It might be true that students' approaches to inferential tasks do not possess the rigor of formal statistical inference, but their engagement to those tasks can be seen as an intuitive preparation for a formal study of the topic in the future. In order to characterize young students' inferences the term *informal statistical inference* is employed (Makar & Rubin, 2009). Formal and informal statistical inference are governed by similar

Saplamidou, S. (2019). Investigating the development of informal inferential reasoning in grade 3 through specially designed tasks. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística.* Disponible en www.ugr.es/local/fqm126/civeest.html

fundamental principles (generalization, data as evidence, probabilistic language), whilst differ in the way they use statistical methods and procedures.

Zieffler Garfield, Delmas and Reading (2008), after a scholastic review of the existing literature, provide a working definition of informal inferential reasoning as the "way in which students use their informal statistical knowledge to make arguments to support inferences about unknown populations based on observed samples" (p.44). This kind of reasoning consists of the following (Zieffler et al., 2008 ·Ben -Zvi et al, 2007; Ben-Zvi et al., 2012):

a. An evidence based reasoning that leads to predictions and generalizations about a wider universe;

b. Drawing on, utilizing and integrating prior knowledge available (reasoning about variability, distributional reasoning, reasoning about signal and noise, sample reasoning, contextual reasoning, graph comprehension, reasoning about comparing groups, probabilistic reasoning and inferential reasoning), and

c. Articulations of uncertainty, by using probabilistic language or references to the strengths and limitations of the inferred conclusions.

Under this theoretical perspective, the present study aims to explore how informal inferential reasoning of a third grade student evolves, through specifically designed tasks. After explaining the methodological procedures followed in the study, we describe the phases through which the student's reasoning develops. Then, we discuss the identified phases in the light of existing literature on informal inferential reasoning and conclude with limitations and some suggestions for future research.

## 2. Method

With the purpose of investigating the question mentioned above, we conducted a case study, as this seeks to explain "the how" of the phenomena through detailed examination of specific cases (Yin, 2013). Semi-structured observation was the research tool and field notes were used as a method of transcribing the data. The study was carried out in a grade 3 class of a primary school in Attica, Greece. Four students participated in the study, and were selected because of their active engagement during the presentation of the problem in the whole class, as well as their high ability to communicate abstract ideas. The present study focuses on Maria's (pseudonym) reasoning process, an average-achieving student that is interested in problem solving.

The context for the main task given to the students of this study comes from the children's book "Martha Blah Blah" written by Susan Meaddaugh. Martha is a dog able to speak like humans, as her meal consists of a soup containing letters of the English (Greek, for the needs of our study) alphabet. Unfortunately, the company that produces her food is forced to eliminate some letters from the soup, due to financial reasons. Utilizing this plot, we ask students to decide, as if they were the company's directors, which letters should be eliminated from the soup, so that the company reduces the costs and Martha retains a comprehensible level of speaking.

### 2.1 Task design

The task design was guided by the principles of mathematical inquiry. According to this theoretical view, students examine ill-structured problems that are based on mathematical (or statistical) evidence (Makar, 2012). As a coincidence, the main question of the task required students to negotiate the meaning of ambiguous phrases (comprehensible level of speaking) as well as the solution plan and the evaluation criteria concerning their response (for example, how many letters should be eliminated?). Aiming to engage students with powerful statistical ideas, we studied the development of informal inferential reasoning, by using the concept of sampling distribution as main instrument, in considering that this concept is central to a web of interconnected statistical ideas (Garfield & Ben-Zvi, 2008). Given the fact that inferential reasoning leads to generalizations from samples to populations (Makar & Rubin, 2009), we purposefully considered to employ stabilized frequency distribution, which represents the connection between probability and frequency according to the Law of Large Numbers (Konold & Kazak, 2008). Serradò, Meletiou-Mavrotheris and Paparistodemou (2015) noted that this kind of distribution helped secondary school students to perceive distribution as a whole, a view that proves crucial while inferring from data (Rubin, Hammerman & Konold, 2006).

We also supplemented the main problem with sub-tasks that promote informal inferential reasoning according to the existing literature. One of the sub-tasks was based on the growing samples heuristic (Konold & Pollatsek, 2002), by asking the students to make sense of a certain sample and form an informal inference concering the letters that should be eliminated from Martha's soup. After analyzing this sample, students predicted what would remain the same and what would change in a larger sample (Ben-Zvi et al., 2012). This heuristic has proved to be helpful for students, as they get a chance to reflect upon stable features of distributions and compare predictions and speculations with actual data (Braham & Ben-Zvi, 2015). Apart from this heuristic, a large part of the task required the comparison of distributions representing data from different sample sizes (10, 50, 120, 500 letters) and from different samples (different texts). Konold and Pollatsek (2007) claim that the comparison of distributions enables students, both novices and experts in statistics, to examine not only measures of central tendency but also features as variation and shape.

Furthermore, we employed Tinkerplots software (Konold & Miller, 2011) in order to investigate the Greek letters' stabilized frequency distribution. The relative frequency of each letter, determined by quantitative linguistic studies was entered into the software. By doing so, the theoretical distribution of letters in the Greek language was produced, so that letters could be picked randomly a number of times by the software and form a new sampling distribution.

### 2.2. Data analysis

In order to trace the development of students' informal inferential reasoning we employed grounded theory analysis. Field notes and student's responses were combined with her work in order to produce texts that describe Maria's statements and actions while making informal statistical inferences based on sample data of a) 10, b) 50, c)120, d) 500 letters and e) while making inferences based on data produced by Tinkerplots' sampler.

We coded each of these texts line by line, using the procedures that differentiate informal inferential reasoning (way of handling the data, language used while making inferences or predictions and informal statistical knowledge) from other reasoning types as unit of analysis. At this phase, we searched for actions or statements related to the previously mentioned axes, in order to develop initial descriptions for codes. Afterwards, we organized coded data into tables that describe Maria's reasoning in terms of informal inferential reasoning and compared them in order to examine the differentiation of student's reasoning during the different phases of the task. We also examined each table separately, so as to name each stage of reasoning. This procedure was influenced by existing frameworks that characterize statistical conceptions (Shaughnessy, 1992) or describe primary school students' statistical reasoning (Jones et al., 2000) with extensive attention to the definitions provided by those frameworks, in order to ensure that the codes produced by the data complied with those definitions. In case of agreement, the stages of reasoning in this study were named after those frameworks. In case of disagreement, a new name was invented based on the characteristics of each reasoning stage.

## 3. Results

Maria begins to consider the main question of the task with a *blurred focus*, not knowing on which aspect of the data at hand she has to focus on. After posing a question to the researcher, she enters into the stage of *idiosyncratic reasoning*, in which she interprets the data sample (n=10) according to personal believes and adopts a deterministic language while making inferences and predictions. It is when those predictions are being refuted by the data set of 50 letters, that Maria decides to examine carefully the data at hand and not only those compatible with personal beliefs. This decision leads her to focus on each letter of the Greek alphabet and compute the sum of the absolute frequencies available for that letter. Following this procedure during the *additive reasoning* phase, she characterizes the sum as "big", if it is bigger than five and "small" in any other case. The letters that had a "small" sum of absolute frequencies where thought to appear less often in a larger sample size and thus where eliminated from the soup with a high level of certainty.

When this strategy also fails to predict the data set of 120 letters, Maria expresses the need to resign and seems frustrated, stating that she doesn't want to examine the problem anymore. Even when the researcher and her classmates dissuade her from giving up, she decides not to participate in the group discussion but to attend to the argument made by one of her peers who notices the inner variability of the distributions and searches for patterns in data. As a response to this argument, Maria states that in each distribution the letters that appear with high and low frequency differ. During this phase, her reasoning can be characterized as *transitional*.

One of her classmates agrees and proposes the examination of the letters that appear 0, 1, or 2 times and are common to all the three distributions. Maria does not seem persuaded and expresses complete uncertainty about the letters that should be eliminated according to the previous argument refusing, at the same time, to make predictions about the sample of 500 letters. She comments that although the data coming from those three texts indicate that some letters have to be eliminated, there are many other texts that may indicate the elimination of different letters. At this phase, Maria's classmates employ a *proportional* form of reasoning, by connecting the absolute frequencies with sample size (thus referring intuitively to relative frequencies). Maria shows little participation and insists on adopting a high level of uncertainty, due to which *we cannot infer that she also entered into the phase of proportional reasoning*.

When Maria starts to examine the data sample of 500 letters, she states that “the spaces [in the sampling distributions] are filled” and this stems from the bigger sample size, because *“if we count just a few letters, it is possible that some frequently used letters like the letter <alpha> aren’t used”*. During this phase she also sees the need to connect absolute frequencies with the sample size while making inferences, and works, in an informal and intuitive way, with relative frequencies. Using the Tinkerplots software, she decides to investigate larger samples and comments that when the sample size is large the distributions look alike and when sample size is “very very large” the distributions look almost the same and “you need a magnifying glass in order to see differences between them”. She adds that when the sample size is large, it seems that the “text from which we count the letters doesn’t affect the result. That’s why we shouldn’t infer anything when we count just a few letters”. Continuing to work with Tinkerplots’ sampler, she states that when the sample size is large, differences are small “not only if we count 2000 letters from different texts, but also if we count 2000 and 3000 letters from the same text”. During this phase, her reasoning was characterized as *advanced informal inferential reasoning*.

In Figure 1 we gather the phases of reasoning traced in the data, and classify them into Shaughessy’s (1992) types of stochastic conception. We represent the transition from one phase to the other using dotted lines. The phases that belong to the same type of statistical conception are placed into the same rectangle with dotted outline. Rectangles with solid fill represent the different phases of Maria’s reasoning. The rectangle representing proportional reasoning is different from the rest, as data available don’t ensure that Maria entered into that phase of reasoning

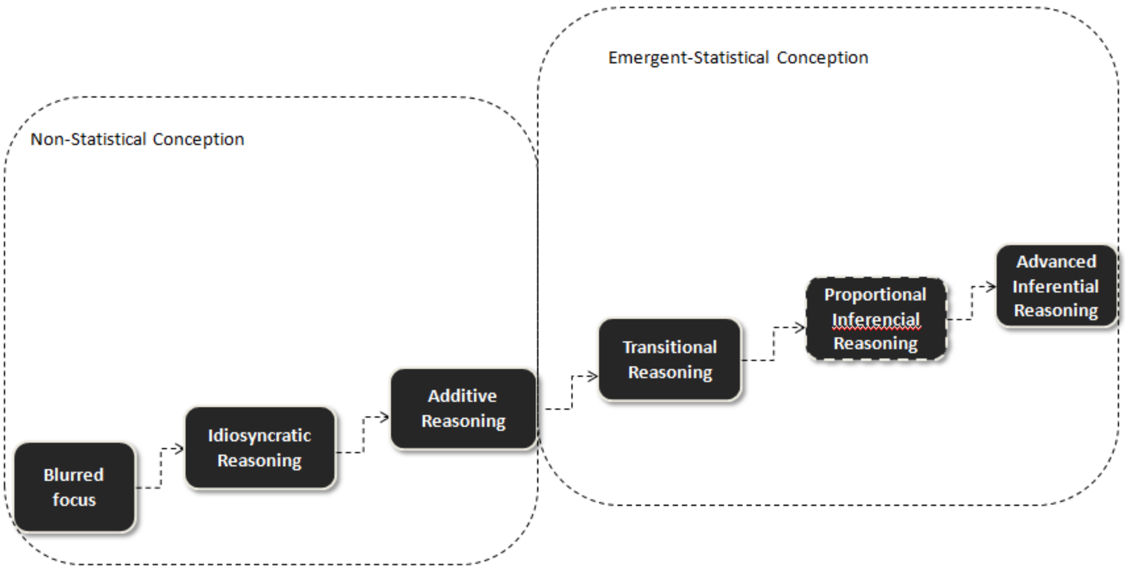


Figure 1. Maria's reasoning process

## 4. Discussion

Examining the results more broadly, we can conclude that Maria’s reasoning during the first phases of its development cannot be characterized as statistical nor as inferential. As the reasoning evolved through carefully designed tasks, its quality improved, by acquiring stochastic characteristics (references to data and uncertainty).

During the phases of non-statistical conception, Maria doesn't know where to focus or focuses on data that corroborate her personal believes. Previous studies have noted young students' lack of focus during initial stages of tasks that required comparing distributions (Ben-Zvi, 2004) and their tendency to confuse data with personal beliefs (Paparistodemou & Meletiou-Mavrotheris, 2008). The lack of focus on statistical information can be attributed, according to Magidson (as quoted in Ben-Zvi, 2004), to the fact that students' opinions about what is relevant and what is not differ from those held by statistic educators.

The difficulty to focus on data may also stem from students' early conceptions about distributions and their difficulty to perceive them as a whole (Garfield & Ben-Zvi, 2008). During this phase, Maria showed complete certainty about her inferences from data, a fact also mentioned by studies using the growing samples heuristic. This deterministic view during the early stages of the task may stem from students' ignorance of variation and uncertainty (Savard, 2014). Rubin and colleagues (2006) give an alternative explanation to this fact, stating that deterministic language comes from the exclusive focus on sample representativeness.

In the phase of Additive Reasoning, Maria adds the absolute frequencies available from different samples and then invents a non-statistical criterion in order to characterize those frequencies as small or high. The sums computation strategy was also acknowledged by studies employing comparison of distributions in secondary school (Shaughnessy, Ciancetta, Best & Noll, 2005; Watson & Moritz, 1999). In these studies, students associated (correctly or not) the sum of frequencies with sample size, contrary to Maria, who evaluated the sum as an absolute number. Maria's strategy might be relevant to her beliefs about problem solving in mathematics. According to Greer, Verschaffel and De Corte (2002) many primary school students hold the belief that problems are solved after performing basic operations between numerical data.

When the strategies employed by Maria in previous stages prove unable to produce valid predictions, she decides to resign from the task in the first place and then places extreme emphasis on sampling variability. Due to this emphasis, she states that she cannot infer anything from data. This viewpoint is common in the literature, as students over relying on sampling variability exhibit a great degree of uncertainty (Braham & Ben-Zvi, 2015). In the phase of transitional reasoning, Maria uses non-deterministic language but perceives the data available as simple information, not as evidence. Therefore, her reasoning cannot be labeled as inferential.

Maria does not seem to change her opinion when her classmates enter into the phase of proportional reasoning, by relating frequencies with the sample size and seeking for patterns in data. The persistence she shows is probably linked to her personal characteristics, as, according to her teacher, she gets easily frustrated after making a mistake. The student reduces the attention she places on sampling variability only when she interacts with Tinkerplots software. The dynamic nature of the software gave her the chance to explore how the distribution of frequencies varies in relation with the sample size and what happens to the sampling variability when the sample size is very big. This improvement in the quality of Maria's reasoning can be related with Tinkerplots software, as according to Braham and Ben-Zvi (2015), its dynamic nature enables students to detect patterns in data that vary.

It is crucial to acknowledge that the present study is not free of limitations. To start with, a case study limits researcher's ability to generalize, as it examines specific cases thoroughly. Apart from the research method, the phases of reasoning traced in this study cannot be generalized due to the fact that they are closely connected to the carefully designed environment in which they were detected. We should also note that the participant in the study possessed certain characteristics· a high level of communication skills and a good academic performance. In the light of those remarks, suggestions for future research could be made. It would be beneficial for the generalizability of the present study, to explore the development of informal inferential reasoning in other grades, in order to detect patterns in students' reasoning. Furthermore, research on the role that teachers play in the development of informal inferential reasoning is needed.

## References


Ben-Zvi, D. (2004). *Reasoning about variability in comparing distributions. Statistics Education Research Journal*, *3*(2), 42-63.

Ben-Zvi, D., & Amir, Y. (2005). How do primary school students begin to reason about distributions? In K. Makar (Ed.), *Proceedings of the Fourth International Research Forum on Statistical Reasoning, Thinking, and Literacy* (pp. 2-7). Brisbane: University of Queensland.

Ben-Zvi, D., Aridor, K., Makar, K., & Bakker, A. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM*, *44*(7), 913-925.

Ben-Zvi, D., Gil, E., & Apel, N. (2007). What is hidden beyond the data? Young students reason and argue about some wider universe. In D. Pratt & J. Ainley (Eds.). *Proceedings of the Fifth International Forum for Research on Statistical Reasoning, Thinking and Literacy*. Warwick, UK: University of Warwick.

Braham, H. M., & Ben-Zvi, D. (2015). Students' articulations of uncertainty in informally exploring sampling distributions. In A Zieffler & E Fry (Eds.), *Reasoning about uncertainty: Learning and teaching informal inferential reasoning* (pp.57-94). Minneapolis, Catalyst Press.

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

Greer, B., Verschaffel, L., & De Corte, E. (2002). "The answer is really 4.5": beliefs about word problems. In G.C. Leder et al.(Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 271-292). Dordrecht, The Netherlands: Kluwer Academic Publishers .

Jones, G. A., Thornton, C. A., Langrall, C. W., Mooney, E. S., Perry, B., & Putt, I. J. (2000). A framework for characterizing students' statistical thinking. *Mathematics Thinking and Learning*, *2*,269–307.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education*, *33*(4), 259-289.

Konold, C., & Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Statistics Education*, 2(1). Available from: https://escholarship.org/uc/item/38p7c94v

Konold, C., & Miller, C. (2011). T*inkerPlots™ 2.0 beta.* Amherst, MA: University of Massachusetts.

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

Paparistodemou, E., & Meletiou-Mavrotheris, M. (2008). Developing young students' informal inference skills in data analysis. *Statistics Education Research Journal, 7*(2), 83-106.

Makar, K. (2012). The pedagogy of mathematical inquiry. In R. Gillies (Ed.), Pedagogy:New developments in the learning sciences (pp. 371–397). Hauppauge, NY: Nova Science.

Paparistodemou, E., & Meletiou-Mavrotheris, M. (2008). Developing young students' informal inference skills in data analysis. *Statistics Education Research Journal, 7*(2), 83-106.

Rubin, A., Hammerman, J., & Konold, C. (2006). Exploring informal inference with interactive visualization software. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics.* Cape Town, South Africa*:* International Association for Statistics Education. Available from: https://www.ime.usp.br/~abe/ICOTS7/Proceedings/PDFs/InvitedPapers/2D3_RUBI.pdf.

Savard, A. (2014). Developing probabilistic thinking: What about people's conceptions? In E. Chernoff, & B. Sriraman (Eds.), *Probabilistic thinking, presenting plural perspectives* (pp. 283-298). Dordrecht, TheNetherlands Springer.

Serradó, A. (2014). Constructing, refining and validating a task for developing reasoning on stabilized relative frequency distributions in the context of informal inferences. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Proceedings of the Ninth International Conference on Teaching Statistics.* Flastaff, AR: IASE. Available from:
https://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_2E2_SERRADOBAYES.pdf

Serradó, A., Meletiou-Mavrotheris, M., & Paparistodemou, E. (2015). Students' reasoning about variability in a horizontal modelling process of the stabilized relative frequencies. In J. M. Contreras, C. Batanero, J. D. Godino, G.R. Cañadas, P. Arteaga, E. Molina, M.M. Gea y M.M. López (Eds.), *Didáctica de la estadística, probabilidad y combinatoria,* 2 (pp. 415-424). Granada: Grupo de Investigación en Educación Estadística.

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* (pp. 465–495). New York: Macmillan.

Shaughnessy, J. M., Ciancetta, M., Best, K., & Noll, J. (2005). Secondary and middle school students' attention to variability when comparing data sets. Paper presented at the*Research Pre-session of the 83rd annual meeting of the National Council of Teachers of Mathematics,* Anaheim, CA: NCTM.

Yin, R. K. (2013). *Case study research: Design and methods*. Sage publications.

Watson, J. M., & Moritz, J. B. (1999). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics, 37*, 145-168.

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.