

# VU Research Portal

## From sample to population

de Vetten, A.J.

2018

### **document version**

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

### **citation for published version (APA)**

de Vetten, A. J. (2018). *From sample to population: Pre-service primary school teachers learning informal statistical inference.*

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

## Summary

The ability to reason inferentially is increasingly important in today's society. To help primary school students acquire the foundations for inferential reasoning, they could be engaged in informal statistical inference (ISI). ISI can be defined as a probabilistic generalization based on, but extending beyond available data. If primary school students are introduced to ISI their teacher should have good content knowledge of ISI (ISI-CK). This dissertation first presents evidence on the ISI-CK of pre-service primary school teachers. Second, it shows how interventions of limited length can foster the ISI-CK of pre-service primary school teachers. The results of the dissertation inform research on ISI-CK of pre-service teachers and suggest how ISI education can be implemented in teacher training.

**Chapter 1** outlines the theoretical framework used in the dissertation and describes several methodological tenets that informed the design of the studies presented in the subsequent chapters. Throughout the dissertation, the ISI framework of Makar and Rubin is used, which covers the key themes of ISI identified in the literature. This framework consists of three components: "data as evidence", "generalization beyond the data", and "probabilistic language".

**Chapter 2** reports on a survey study of the ISI-CK of all three ISI components of beginning pre-service teachers. We surveyed 722 first-year pre-service primary school teachers from seven teacher colleges across The Netherlands. We used a questionnaire that combined open-ended questions with true/false statements. Most respondents understood that descriptive statistics that take the global shape of the distribution into account, such as the mean, can be used as arguments within ISI. Although a majority agreed that random sampling is a valid sampling method, distributed sampling (i.e. purposefully selecting individuals to obtain a distributed sample across critical population characteristics) was the preferred sampling strategy. Furthermore, when asked to make a generalization beyond the data, most pre-service teachers only described the data, rather than generalizing beyond the data. Finally, less than 40% of the respondents understood that generalization is possible because sample-to-sample variability will be low for a sample where an appropriate sampling method and sample size is used. These findings suggest that it may be useful if statistics education for pre-service teachers

places more emphasis on sampling and inference, thereby prompting pre-service teachers to engage in ISI.

**Chapter 3** reports on an exploratory study into the informal inferential reasoning of first-year pre-service teachers who were engaged in a growing samples activity. They were asked to use sample data to generalize to a population and to predict the graph of a larger sample during three rounds with increasing sample sizes. The content analysis revealed that most pre-service teachers used suitable descriptive statistics to compare the two distributions. As in Chapter 2, most pre-service teachers only described the data, without noticing that the activity required a generalization beyond the data. After class discussions, in two classes most participants seemed to agree that it is possible to make probabilistic generalizations based on a sample. However, the majority of the participants could not link the possibility of making generalizations to an understanding of how a sample can represent the population.

Since both in Chapter 2 and 3, the pre-service teachers showed relatively good ability to descriptively analyze sample data, while they tended not to generalize, the intervention of **Chapter 4** put less emphasis on descriptive analyses and more emphasis on sampling and inference. During this intervention, we investigated the ISI-CK development in a class of 21 pre-service teachers. The intervention consisted of five teacher college sessions, half of which focused on the participants' own ISI-CK, and half focused on their pedagogical content knowledge of ISI. Based on qualitative and quantitative analyses of the pretest, posttest and intervention data, the results suggest that most participants acknowledged it is possible to make uncertain inferences. In contrast to most pre-service teachers participating in the studies of Chapter 2 and 3, the pre-service teachers noticed when tasks required an inference. This awareness was possibly sparked by an assignment to search the media for inferential claims. A demonstration of a real-time computer simulation illustrating the law of large numbers may have led to increased understanding of sampling variability and to an increased acceptance of random sampling. The statistical investigation conducted by the participants appeared to have further strengthened their awareness of ISI, but also revealed that many participants continued to favour distributed sampling over random sampling. No development was seen in the participants' knowledge of sufficient sample

sizes. The participants seemed to lack tools to express the certainty of their inferences.

As teachers ultimately have to use their ISI-CK in teaching, **Chapter 5** describes the ISI-CK of three pre-service teachers when teaching an ISI lesson in their primary placement schools. These pre-service teachers also participated in the study of Chapter 4. The results showed that the pre-service teachers made generalizations based on the collected data when teaching their lessons. This result is encouraging, since it shows that pre-service teachers who have limited experience with ISI are able to express their newly acquired ISI-CK in teaching. However, although the sample size to be used was discussed extensively, the sampling method to be employed remained largely undiscussed. Moreover, all pre-service teachers faced difficulties when trying to explain why making inferences is possible. The lack of tools to express the certainty of their inferences (see Chapter 4) became problematic in this study, because without such tools they were not able to explain why making inferences is possible. To allow pre-service teachers to provide such an explanation, the lesson should contain tools that allows them to illustrate the law of large numbers.

**Chapter 6** integrates and critically discusses the findings of the research presented in Chapter 2 – 5. The generalizability of the findings was limited by the first-person perspective of the research and the specific Dutch teacher college context. Also, the short length of the intervention limited the opportunities to familiarize the pre-service teachers with ISI. We point at several directions for future research. First, we propose to validate the findings of our research in other settings and to conduct effect studies to study the learning effects of the intervention. Second, we suggest to study how in initial engagements of primary school students to ISI, the focus can be on one crucial question: Why making inferences based on one sample is possible.

Based on our results, we formulate a number of design heuristics that could be used in the design of ISI education at the primary education teacher college. Examples of these design heuristics are: (1) create awareness of inferential claims and of the distinction between sample and population before engaging pre-service teachers in ISI via statistical investigations; (2) use tangible samples and populations; (3) use data that require little descriptive analyses when conducting ISI investigations with pre-service teachers; and (4)

provide pre-service teachers who introduce primary school students to ISI with a well-structured lesson plan that yields sufficient affordances to engage students in ISI and that contains a tool that helps them to explain why making inferences is possible.

In conclusion, this thesis shows promising results of pre-service teachers developing their ISI-CK within a limited time frame. They were able to make generalizations about a population based on sample data when teaching an ISI lesson. Also, most of them showed the agency to learn and to teach ISI. If in the future ISI would become part of teacher college curricula, the insights from this thesis can be put to fruitful use. It will help in the implementation of ISI education in primary education teacher colleges, and ultimately help to prepare primary school students to participate in society.