



HAL
open science

Informal statistical inference and pre-service primary school teachers: The development of their content knowledge and pedagogical content knowledge during a teacher college intervention

Arjen de Vetten, Judith Schoonenboom, Ronald Keijzer, Bert van Oers

► To cite this version:

Arjen de Vetten, Judith Schoonenboom, Ronald Keijzer, Bert van Oers. Informal statistical inference and pre-service primary school teachers: The development of their content knowledge and pedagogical content knowledge during a teacher college intervention. CERME 10, Feb 2017, Dublin, Ireland. hal-01927861

HAL Id: hal-01927861

<https://hal.science/hal-01927861>

Submitted on 20 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Informal statistical inference and pre-service primary school teachers: The development of their content knowledge and pedagogical content knowledge during a teacher college intervention

Arjen de Vetten¹, Judith Schoonenboom², Ronald Keijzer³ and Bert van Oers⁴

¹Vrije Universiteit Amsterdam, Amsterdam, The Netherlands & Hogeschool iPabo, Amsterdam, The Netherlands; a.j.de.vetten@vu.nl

²University of Vienna, Vienna, Austria; judith.schoonenboom@univie.ac.at

³Utrecht University, Utrecht, The Netherlands & Hogeschool iPabo, Amsterdam, The Netherlands; r.keijzer@uu.nl

⁴Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; bert.van.oers@vu.nl

Teachers who engage primary school students in informal statistical inference (ISI) need to have a good content knowledge (CK) and pedagogical content knowledge (PCK) of ISI themselves. However, little is known of how teacher college education for pre-service teachers can contribute to the development of their ISI CK and PCK – a shortcoming we attempt to address in this paper. A class of 21 pre-service primary school teachers participated in an intervention consisting of five lessons. Design research methodology guided the design of the intervention. The preliminary results indicate that most pre-service teachers seemed to be aware of the inferential nature of research questions and the uncertainty inherent to ISI, but not all of them understood the fundamental idea that if a properly selected sample is representative for the population it can be used for an inference. Lacking this understanding might hinder them in teaching ISI.

Keywords: Informal statistical inference, informal inferential reasoning, primary education, statistics education, teacher education.

Introduction

In daily life, sample data is regularly used to make generalizations that go beyond the data collected; informal statistical inference (ISI) is a form of this phenomenon. The ability to critically evaluate such generalizations is increasingly useful in participation in present and future society. ISI is defined as “a generalized conclusion expressed with uncertainty and evidenced by, yet extending beyond, available data” (Ben-Zvi, Bakker, & Makar, 2015, p. 293). Unlike formal statistical inference, ISI does not make use of formal statistical tests that are based on probability theory (Harradine, Batanero, & Rossman, 2011). In recent years, statistics education researchers have turned their attention to how primary school students can be introduced to ISI. It is hypothesized that if students are familiarized with the concept in primary school, it will help them to understand the processes involved in its reasoning and in statistical reasoning in general (Bakker & Derry, 2011; Makar, Bakker, & Ben-Zvi, 2011). Evidence suggests that meaningful learning environments can render ISI accessible for primary school students (Ben-Zvi et al., 2015; Meletiou-Mavrotheris & Papanastasiou, 2015).

If students are to be introduced to ISI in primary school, future teachers need to be well prepared to conduct this introduction (Batanero & Díaz, 2010). This requires them to have an appropriate content knowledge (CK) of the subject and to have adequate pedagogical knowledge (PCK) (Burgess, 2009). Recent research shows that pre-service teachers have difficulty making generalizations and

understanding sampling representativeness and the logic of sampling (De Vetten, Schoonenboom, Keijzer, & Van Oers, 2016a; De Vetten, Schoonenboom, Keijzer, & Van Oers, 2016b). However, little is known how we can prepare pre-service teachers to teach ISI to primary school students. This paper reports on the design and preliminary results of an intervention at a teacher college for primary education that aimed to foster the development of the ISI content knowledge (CK) and PCK of pre-service primary school teachers.

Teachers' knowledge of ISI

In his seminal article, Shulman (1986) uses two broad categories to categorize the knowledge teachers need to have to teach a particular subject: pedagogical content knowledge (PCK) and content knowledge (CK) of the subject. Shulman's first category, pedagogical content knowledge, includes knowledge of how to present, illustrate and explain new material (Knowledge of Content and Teaching; KCT for short); and knowledge about the students' conceptions and misconceptions (Knowledge of Content and Students; KCS for short). Concerning content knowledge, teachers need to understand the subject at the level of student (Common Content Knowledge; CCK for short), but also need to have specialized knowledge of the subject that enables them to teach the subject and that is specific for the job of teaching (Specialized Content Knowledge; SCK for short) (Ball, Thames, and Phelps, 2008).

For ISI no categorization of CK and PCK is available. We combine the categorization of Shulman with the ISI framework of Makar & Rubin (2009) to conceptualize the CK and PCK needed for teaching ISI. Makar & Rubin (2009) argue that an ISI consists of the following three components:

1. 'Data as evidence': The inference is based on the data, and not on tradition, personal beliefs or experience.
2. 'Generalization beyond the data' (in short: 'Generalization'): The inference goes beyond a description of the sample data to make a claim about a situation beyond the sample data.
3. 'Uncertainty in inferences' (in short: 'Uncertainty'): The inference includes a discussion of the sample characteristics, such as sample size and sampling method, and a discourse on what these characteristics imply for the representativeness of the sample and the certainty of the inference. Moreover, it requires an understanding of basic logic sampling: the understanding that if a properly selected sample is representative for the population it can be used for an inference, because sample-to-sample variability is low. We have subdivided this component into three subcomponents: sampling method, sample size and uncertainty.

There are a limited number of studies that investigate (pre-service) primary school teachers' CK of ISI. The authors of this paper have investigated all three ISI components in two studies (De Vetten et al., 2016a; De Vetten et al., 2016b). In an exploratory design study, De Vetten et al. (2016a) show that regarding 'Data as evidence' most pre-service teachers indeed use data as evidence when comparing two samples to generalize to the population. In a large scale questionnaire study, De Vetten et al. (2016b) found less positive results when pre-service teachers were asked to evaluate whether data can be used as reliable evidence for a generalization. Concerning 'Generalization', De Vetten et al. (2016b) show that pre-service teachers are well able to discern that probabilistic generalizations are permissible, while deterministic generalizations are not. However, De Vetten et al. (2016a) report that pre-service teachers tend to only describe the samples, and do not generalize beyond the data.

The evidence on the ‘Uncertainty’ component suggests that many pre-service teachers show a limited understanding of sampling methods, sample size, representativeness and the logic of sampling and sampling variability (De Vetten et al, 2016a; De Vetten et al., 2016b; Meletiou-Mavrotheris et al., 2014; Mooney, Duni, VanMeenen, & Langrall, 2014; Watson, 2001).

Research on teachers’ PCK of ISI and ways to prepare to teach ISI is even scarcer than research on teachers’ CK of ISI. Leavy (2010) reports that pre-service teachers tended to focus excessively on procedures, spent too much time on descriptive analyses at the expense of discussion of inferences, and failed to stimulate data-based reasoning. Using the same data, Leavy (2015) shows that it is critical that pre-service teachers learn how to pose questions that invite students to reason about inference. Madden (2011) shows that tasks that are statistically, contextually and/or technologically provocative triggered high school mathematics teachers to reason about ISI.

Since little is known how to prepare pre-service primary school teachers to teach ISI, the aim of the present study is to investigate in what way teacher college education for pre-service teachers can contribute to the development of their ISI CK and PCK. The research question is: To what extent, and how, do the ISI CK and PCK of pre-service school teachers develop during, and as a result of, an intervention at primary education teacher college aiming at developing ISI CK and PCK?

Method

Context

In many countries, including the Netherlands, current statistics education curricula in primary and secondary education do not include ISI. Actual teaching practices focus primarily on statistical procedures and graphing skills, where concepts are learned without reference to the need to collect and analyze data (Ben-Zvi & Sharett-Amir, 2005; Friel, Curcio, & Bright, 2001; Meijerink, 2009). When statistical inference does form part of the secondary education curriculum, the ideas of sample and population are often only dealt with on a technical level. Consequently, many students enter tertiary education with a shallow and isolated understanding of the concepts underlying statistical inference (Chance, DelMas, & Garfield, 2004). In contrast to many other countries, where students can only opt for teacher education after completion of a bachelor’s degree, in the Netherlands, initial teacher education starts immediately after secondary school and leads to the attainment of such a degree. For these students, mathematics teaching is usually not their main motive for becoming teachers.

The intervention was part of a course on mathematics education for grade 3 to 6. The course was the fourth course on mathematics education and the second for mathematics in grade 3 to 6. During the semester the pre-service teachers worked in a grade 3 to 6 class in a work placement school. Since in the Dutch mathematics curriculum for teacher college statistics gets only minor attention and since we wanted to have an intervention that would fit in the normal teacher college curriculum, we decided to restrict the length of the intervention to five lessons, out of the 16 lessons of the total course.

Design

We employed design research methodology to study the development in ISI CK and PCK of the pre-service teachers and to explain this development (Van den Akker, Gravemeijer, McKenney & Nieveen, 2006). Previous research (Ben-Zvi, 2006; De Vetten et al., 2016a&b; Paparistodemou &

Meletiou-Mavrotheris, 2008; Saldanha & Thompson, 2007) and our own ISI teaching experiences with primary school students informed us which learning goals are within reach of pre-service teachers and what PCK is necessary to teach ISI to primary school students (see Table 1). These learning goals were categorized into the three ISI components ‘Data as evidence’, ‘Generalization’ and ‘Uncertainty’.

Table 1: ISI learning goals for the intervention at primary education teacher college

ISI component	Knowledge type ^a	Learning goals	Attained? ^b	
<i>General</i>	KCT	To reason with students about ISI in a meaningful way, teachers can (1) have students conduct empirical investigations with an inferential research question about a meaningful topic or (2) have students evaluate research (for example as reported in the media) with an inferential research question about a meaningful topic.	0.75	
<i>Data as evidence</i>	CCK	Use data as evidence, not other sources	1	
	CCK	Sample provides information about likelihood of population parameters	0.5	
	KCS	Many students do not use data as evidence	0.5	
	KCS	Many students think that every sample distribution is evenly likely	0.25	
	KCT	To teach students that data can be used as evidence for answers on inferential questions teachers can (1) have students conduct empirical investigations where sample and population are concretely visible, (2) regularly point at the sample, and (3) ask students on what arguments they or other researchers reached their conclusion.	0.25	
<i>Generalization beyond the data</i>	CCK	It is possible to make claims about population, despite individual differences between subjects	0.75	
	CCK	It is possible to use sample to make claims about population	1	
	CCK	Awareness of inferential nature of research questions	1	
	CCK	Claims about a population are often based on a sample	1	
	CCK	Correctly articulate answers to inferential questions	0.5	
	KCS	Many students answer inferential questions descriptively only	0.5	
	KCT	To teach students that it is not necessary to investigate an entire population, but that a sample suffices for a reliable conclusion, teachers can (1) use examples from media where a sample is used, or (2) use resampling activities where different samples yield similar results.	0.25	
	KCT	To make students aware of the inferential nature of research questions, teachers can (1) use real empirical investigations where sample and population are concretely visible, and (2) ask questions, such as 'Does our result hold for the sample only, or also for the population?'	0.25	
<i>Uncertainty inherent to inferences</i>	<i>Sampling method</i>	CCK	Which of the following sampling methods are (in-)appropriate: convenience sampling, random sampling, quota sampling	0.75
		SCK	Understand why random sampling is appropriate	0.75
		KCS	Many students think random sampling is not an appropriate sampling method	0.5
		KCS	Many students tend to use incorrect matching techniques in sampling	0
		KCT	To teach students that a sample needs to be representativeness of the population, a teacher can (1) ask students to investigate how other researchers have selected their sample and what students think of the representativeness of the sample and (2) let students discuss how they would select a representative sample to answer a specific research question and let the children execute the sampling.	0.75
	<i>Sample size</i>	SCK	Why a larger sample leads to more certainty	0.5
		CCK	Sufficient sample size	0.25
		CCK	Sample size is independent of population size	0.5
		KCS	Many students do not take sample size into account in the certainty of their answers	0.5
		KCS	Many students make very certain inferences, even for small samples	0.5
		KCS	Many students think that sample size is dependent of population size	0.25

<i>Uncertainty</i>	KCT	To teach students the effect of sample size on the certainty of inferences, teachers can use real empirical investigations where resampling is used, select to two samples of different size and ask which sample provides more certainty.	0.5
	CCK	Acknowledge uncertainty of inferences and impossibility of absolute certain inferences	1
	CCK	The larger the sample, the more certain the inference	1
	CCK	The better the sampling method, the more certain the inference	0.75
	CCK	When a sample is properly selected, the probability is small that another but likewise sample gives an entirely different result	0.5
	CCK	Correctly articulate uncertainty in inferences	0.5
	KCS	Many students express complete (un-)certainty in their inferences	0.5
	KCT	To make students aware of the uncertainty of inferences, teachers can use real empirical investigations where resampling is used, and confront the children that different samples lead to different conclusions.	0.25
	KCT	To help students articulate uncertainty in inferences, teachers can reformulate students' responses or ask how much certainty the students have.	0.25

^aCCK: common content knowledge; SCK: specialized content knowledge; KCS: knowledge of content and students; KCT: knowledge of content and teaching. ^bEach learning goal received one of the scores 0, ¼, ½, ¾ or 1 (0: not attained; 1: attained by (almost) all pre-service teachers)

For each lesson, a hypothetical learning trajectory was designed, connecting activities with the learning goals for the lesson, while also explaining in what way the activities were hypothesized to help to attain the learning goals. Using example lessons was one type of activity used to foster the CK and PCK of the pre-service teachers. These lessons introduced many CK concepts and provided opportunities to discuss how children would deal with issues involved in these lessons. Each lesson was evaluated to inform the design of the next lesson. Table 2 provides an overview of the intervention. Part of the intervention was that pre-service teachers would give an ISI lesson in their work placement school. The analysis of these lessons is beyond the scope of this paper.

Table 2: Overview of the ISI intervention at the teacher college

Week	Activity
1	Informed consent, pre-test and homework instruction
Between 1 and 3	CK: Pre-service teachers make homework assignment: 1. Look up an article in the media that makes a claim about a population based on a sample and that somehow appeals to you. 2. Describe how the researchers came to their conclusions. 3. Write a critical evaluation about the quality of the research: to what extent is in your opinion the conclusion justified based on the research conducted?
3	Lesson 1 CK & PCK: Small group and whole class discussion of homework assignment, attention for both CK and PCK aspects of the assignment CK: Explanation of random sampling and appropriate sample size using an ICT demonstration
5	Lesson 2 CK & PCK: Teacher educator models an example lesson which the pre-service teachers could give themselves in their work placement schools PCK: Discussion of teacher educator's experiences with teaching the example lesson in primary school
5	Lesson 3 CK & KCS: Discussion of equiprobability bias using a task CK Short recap of main ISI concepts PCK: Teacher educator presents learning goals of ISI lesson (KCT), discusses typical responses of students (KCS) and provides instructions for the ISI lesson (KCT)
Between 5 and 12	Half of the pre-service teachers teach an ISI lesson in their placement school.
12	Lesson 4 PCK: Discussion of pre-service teachers' experiences with teaching with teaching the ISI lesson PCK: Pre-service teachers provide suggestions for alterations of the ISI lesson and tips for their fellow students.
Between 12 and 16	The other half of the pre-service teachers teach an ISI lesson in their placement school.
16	Lesson 5 PCK: Discussion of pre-service teachers' experiences with teaching the ISI lesson
16	Post-test and evaluation of the intervention

Participants

One class of second year pre-service teachers participated in this study. They studied at a small teacher college for primary education in a large city in the Netherlands. The intervention took place in their second year of study, because statistics is part of the knowledge base that is tested in the third of study. This particular class was chosen, because the pre-service teachers had fewest credits open from their first year of all three second year classes. The class consisted of 23 pre-service teachers. They were asked to provide their informed consent. While all of them were required to participate in the activities and lessons, one pre-service teacher invoked the possibility to have her results excluded from the analysis. The results of another student were also excluded, because of absence during all but one of the lessons. This resulted in a sample of 21 pre-service teachers. The procedure was approved by the ethical board of the Faculty of Behavioural and Movement Sciences of Vrije Universiteit Amsterdam. The average age of the participants was 21 years (SD: 2.19); 6 were male; 10 had a background in secondary vocational education (students attending this type of course are typically aged between 16 and 20), 7 came from senior general secondary education, 2 had been enrolled in university preparatory education, and the educational background of the remaining 2 was either something else entirely or unknown. Their average score on the obligatory first-year mathematics exam for Dutch pre-service teachers was 135 out of 200 possible points (SD: 12.86). A score of 103 equals the 80th percentile of Grade 6 primary school students in the Netherlands. The first author was the teacher educator. He had four years of experience as a mathematics teacher educator, a master degree in economics and experience as a university statistics lecturer. He had taught most of the pre-service teachers during their first year of study.

Data collection

During the lessons, whole class interactions were recorded on video and audio, while small group interactions were recorded on audio. Furthermore, during most lessons one of the co-authors was present as observer, taking notes. Finally, all written work was collected. A pre-test and post-test were used to measure ISI CK and PCK at the start and at the end of the intervention. The results of these tests will be presented during the CERME presentation.

Data analysis

The goal of the analysis is to study the development of the ISI CK and PCK and to explain these developments. This paper reports preliminary analyses. After each lesson, the teacher educator's and observer's notes and reflections were used as for estimating to what extent the learning goals relevant for the particular lesson had been attained by the pre-service teachers as a whole. Each learning goal received a score ranging from 0 to 1 (0: not attained; 1: by large attained by (almost) all pre-service teachers). These estimations per lesson were used to make an overall score for each learning goal. Next, average scores were calculated for the various components and types of knowledge. Based on these notes and reflections, these scores were related to the activities used during the intervention. The final results will show in detail the development of ISI CK and PCK at the level of the pre-service teacher and the role of the activities used in the intervention.

Table 3: Attainment of the learning goals, summarized by ISI component and knowledge type

Component	Average score ^a		Knowledge type	Average score ^a
General	0.75		Common content knowledge	0.73
Data as evidence	0.5		Specialized content knowledge	0.63
Generalization beyond the data	0.61		Knowledge of content and students	0.39
Sampling method	0.55		Knowledge of content and teaching	0.43
Sample size	0.43			
Uncertainty	0.59			

^aEach goal received one of the scores 0, ¼, ½, ¾ or 1 (0: not attained; 1: by large attained by (almost) all pre-service teachers)

Preliminary results

The preliminary results reveal to what extent the learning goals have been attained and shed some light on what activities helped to reach these goals. Table 1 shows to what extent the learning goals are attained. Table 3 shows the average score for the learning goals summarized for components and knowledge type respectively. There are some notable results. First, most pre-service teachers seemed to be aware of the inferential nature of research questions and the uncertainty inherent to the results. The homework assignment at the start of the intervention seemed to have helped to foster this awareness. Second, the modeling of the example lesson seemed to be a good context to reason about sampling methods, because it naturally led to the use and discussion of random and quota sampling methods. Moreover, although average scores on PCK learning goals are lower than on CK learning goals, modeling helped the pre-service teachers to get a better idea how to teach ISI to primary school students. Thirdly, while the presentation of PCK issues by the teacher educator in lesson 8 did not seem to contribute much to the pre-service teachers understanding, the discussion of PCK based on the pre-service teachers' experiences in lesson 12 and 15 did. Finally, an understanding of the fundamental idea that if a properly selected sample is representative for the population it can be used for an inference, might be conditional for understanding other ISI concepts. While part of the pre-service teachers seemed to understanding this idea, part of them did not. Lacking this understanding might hinder them in teaching ISI.

References

- Bakker, A., & Derry, J. (2011). Lessons from inferentialism for statistics education. *Mathematical Thinking and Learning*, 13(1-2), 5–26.
- Batanero, C., & Díaz, C. (2010). Training teachers to teach statistics: What can we learn from research? *Statistique et enseignement*, 1(1), 5–20.
- Ben-Zvi, D. (2006). *Scaffolding students' informal inference and argumentation*. Paper presented at the Seventh International Conference on Teaching Statistics, Salvador, Brazil.
- Ben-Zvi, D., Bakker, A., & Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics*, 88(3), 291–303.
- Ben-Zvi, D., & Sharett-Amir, Y. (2005). *How do primary school students begin to reason about distributions*. Paper presented at the Fourth international research forum on statistical reasoning, thinking, and literacy (SRTL-4), University of Auckland, New Zealand.
- Burgess, T. (2009). Teacher knowledge and statistics: What types of knowledge are used in the primary classroom? *Montana Mathematics Enthusiast*, 6(1&2), 3–24.

- Chance, B., DeMas, R. C., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295–323). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (2016a). *Exploring student teachers' reasoning about informal statistical inference when engaged in a growing samples activity*. Paper presented at the Thirteenth International Conference on Mathematical Education (ICME13), Hamburg.
- De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (2016b). Primary school student teachers' knowledge of informal statistical inference. *Submitted for publication*.
- 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.
- Harradine, A., Batanero, C., & Rossman, A. (2011). Students and teachers' knowledge of sampling and inference. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Joint ICMI/IASE study: Teaching statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. (pp. 235–246). Dordrecht, The Netherlands: Springer.
- Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The reasoning behind informal statistical inference. *Mathematical Thinking and Learning*, 13(1-2), 152–173.
- Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, 8(1), 82–105.
- Meijerink, H. (2009). *Referentiekader taal en rekenen: de referentieniveaus. [Reference levels language and mathematics]*. Enschede: Doorlopende Leerlijnen Taal en Rekenen.
- Meletiou-Mavrotheris, M., Kleanthous, I., & Papanastasiou, E. (2014). *Developing Pre-service Teachers' Technological Pedagogical Content Knowledge (TPACK) of Sampling*. Paper presented at the Ninth International Conference on Teaching Statistics (ICOTS9), Flagstaff, AZ.
- Meletiou-Mavrotheris, M., & Papanastasiou, E. (2015). Developing students' reasoning about samples and sampling in the context of informal inferences. *Educational Studies in Mathematics*, 88(3), 385–404.
- Mooney, E., Duni, D., VanMeenen, E., & Langrall, C. (2014). *Preservice teachers' awareness of variability*. Paper presented at the Ninth International Conference on Teaching Statistics (ICOTS 9), Flagstaff, AZ.
- Papanastasiou, E., & Meletiou-Mavrotheris, M. (2008). Developing young students' informal inference skills in data analysis. *Statistics Education Research Journal*, 7(2), 83–106.
- Saldanha, L., & Thompson, P. (2007). Exploring connections between sampling distributions and statistical inference: An analysis of students' engagement and thinking in the context of instruction involving repeated sampling. *International Electronic Journal of Mathematics Education*, 2(3), 270–297.
- Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research*. New York, NJ: Routledge.
- Watson, J. M. (2001). Profiling teachers' competence and confidence to teach particular mathematics topics: The case of chance and data. *Journal of Mathematics Teacher Education*, 4(4), 305–337.