PRE-SERVICE PRIMARY SCHOOL TEACHERS' (PEDAGOGICAL) CONTENT KNOWLEDGE DURING TEACHING INFORMAL STATISTICAL INFERENCE

Arjen de Vetten, Judith Schoonenboom, Ronald Keijzer, and Bert van Oers
Leiden University, The Netherlands
University of Vienna, Austria
iPabo University of Applied Sciences, The Netherlands
Vrije Universiteit Amsterdam, The Netherlands
a.j.de.vetten@iclon.leidenuniv.nl

We analyzed the lessons of three pre-service primary school teachers (PSTs) who taught a lesson about informal statistical inference (ISI). These PSTs participated in a teacher college intervention. We investigated the appropriateness of their ISI knowledge when teaching to primary school students an ISI lesson provided by their teacher educator. Using the ISI framework by Makar & Rubin and the Knowledge Quartet framework we coded and categorized the teaching actions of the PSTs. The results showed that a large majority of the PSTs' teaching actions was appropriate. Overall, they did well in analyzing children's conceptual suggestions, but PSTs found it more difficult to express and connect ISI conceptual ideas themselves, in particular, to explain how generalizing from a sample is possible. Teacher college education should focus on how PSTs can foster children's understanding of how it is possible that a sample suffices to draw conclusions about a population.

INTRODUCTION

Data and statistics have become pervasive in our society (Gravemeijer, Stephan, Julie, Lin, & Ohtani, 2017). Often, these data are used for inferential reasoning (i.e., the process of drawing a conclusion based on evidence or reasoning). This reveals the need for citizens to be able to make and critically evaluate inferences from empirical data (Liu & Grusky, 2013). To prepare children for their roles in society, it may be beneficial to introduce inferential reasoning to students during their elementary education. One way to introduce inferential reasoning to elementary school students is by means of informal statistical inference (ISI; Groth & Meletiou-Mavrotheris, 2018), which can be defined as "a generalized conclusion expressed with uncertainty and evidenced by, yet extending beyond, available data" (Ben-Zvi, Bakker, & Makar, 2015, p. 293). Research has shown that elementary school students are indeed able to begin reasoning about inferences (Leavy, 2017; Makar, 2016). Teachers are usually the ones to conduct this introduction in ISI. In a previous study, we studied how pre-service primary school teachers' ISI knowledge developed during a short teacher college intervention (De Vetten, Schoonenboom, Keijzer, & Van Oers, 2018). However, the goal of teachers learning content is to teach this content in school. Therefore, we also need to study teachers' content knowledge during actual teaching to see what knowledge they are able to mobilize during teaching (Ball et al., 2001; Fennema & Franke, 1992; Rowland et al., 2009). The present study did this for the case of ISI. We investigated to what extent three pre-service teachers (PSTs) who participated in the teacher college intervention were able to show appropriate ISI knowledge during teaching an ISI lesson in their elementary internship classrooms. This ISI lesson was taught to them during the teacher college intervention.

To describe the ISI-CK of the three PSTs showed during teaching, we used the ISI framework by Makar and Rubin (2009). This framework consists of three components: data as evidence, generalization beyond the data, and probabilistic language. We divided the third component into four subcomponents: sampling variability, sampling method, sample size, and uncertainty (De Vetten et al., 2018). To study how the three PSTs used their ISI knowledge during teaching an ISI lesson, we used the Knowledge Quartet framework (Rowland, Huckstep, & Thwaites, 2005). This KQ framework is specifically designed for empirical evaluations of content knowledge shown during the planning and delivery of teaching of (pre-service) teachers. The KQ comprises of four dimensions of teachers' content knowledge: foundation, transformation, connection, and contingency. Each dimension consists of four to seven aspects (see Appendix 1 and knowledgequartet.org). The research question is: To what extent do pre-service primary school teachers show appropriate ISI knowledge when teaching an ISI lesson to primary school students?

In: R Helenius, E Falck (Eds.), *Statistics Education in the Era of Data Science*
Proceedings of the Satellite conference of the International Association for Statistical Education (IASE),
Aug-Sept 2021, Online conference. iase-web.org/Conference_Proceedings.php

METHOD

Context and participants

This study reports on the ISI-CK of three second-year pre-service elementary school teachers—Celine, Demi, and Alfred (pseudonyms)—as they taught an ISI lesson to primary school students. These teachers were among 21 participants of a teacher college intervention that aimed to improve their ISI-CK and PCK (De Vetten et al., 2018). Celine taught her ISI lesson between the third and fourth meeting of the intervention; Alfred and Demi between the fourth and fifth meeting. Celine (female; 18 years; showing high ISI-CK compared to the other participants during the teacher college intervention) taught in third grade, Demi (female; 19 years; average ISI-CK) in fifth grade, and Alfred (male; 20 years; high ISI-CK) in a combined third/fourth grade class. Participants and placements schools provided informed consent.

Lesson "What is the most frequently used word in a stack of children novels?"

The PSTs all taught the same lesson, called "What is the most frequently used word in a stack of children novels?" The lesson centered on a large collection of Dutch children's novels. The driving question was which word would be the most frequently used in the collection of books. The enormity and visibility of the population was expected to elicit the need to draw a sample and make inferences. The investigation was based on five words that the class expected to find most frequently. Class discussion was used to reach consensus about the preferred sampling method so that separate groups' sample data could later be pooled into one large sample. The groups were to conduct an investigation using the agreed-upon sampling method. The analysis of the sample data was kept simple, as only frequencies needed to be counted. The subsequent discussion dealt with the possibility and certainty of generalization from sample results, both from the individual groups' sample data and from the pooled data. The lesson was also laid out in a lesson plan, including a suggested sequence and questions to ask to children, but the PSTs were expected to adjust the lesson plan to their class. In general, their lesson plans were closely aligned to the provided lesson plan.

Data collection and analysis

Two kinds of data were collected. The elementary data source were audio and video recordings of the classroom interactions. The second data source was comprised of transcripts of audio-taped reflection interviews between the PST and the first author, which were conducted shortly after the lessons. Additionally, PSTs' lesson plans, students' written work, and the observer's notes were used as data. All data from the class discussion and the reflection interviews were transcribed. The unit of analysis was a teaching situation, constituted by a fragment of whole class discussion that concerned one substantive topic. Within each fragment the PSTs' actions were analyzed from the idea that these actions "could be construed to be informed by a trainee's mathematics content knowledge or […] mathematical pedagogical knowledge" (Rowland et al., 2005, p. 258).

Each fragment was coded using both the ISI and the KQ frameworks. The coding process from the ISI framework used a process consisting of deductive and inductive elements. On the deductive side, the ISI framework was used to categorize the data into one or more ISI components. On the inductive side, codes (short summaries of the text) were attached to the text to describe the exact meaning. The coding process from the KQ framework followed the approach of Weston (2013). First, for each fragment the presence of each of the 20 KQ codes was coded (present versus non-present). Second, the appropriateness of the (non-)presence was evaluated (appropriate versus inappropriate). Present teaching actions were coded as appropriate when these teaching actions helped the lesson move towards attainment of the learning objectives. Absent teaching actions were coded as appropriate when this absence did not hinder the attainment of the learning objectives. Present teaching actions were coded as inappropriate when these teaching actions hindered the attainment of the learning objectives. Absent teaching actions were coded as inappropriate when these teaching actions were essential to move the lesson towards attainment of the learning objectives. The coding process was discussed among the first and the second author until consensus was reached about the coding's validity. The analyses involved (cross-)tabulating the codes from the ISI and KQ frameworks and searching for overall patterns and notable results.

RESULTS

Results for Knowledge Quartet

Overall, 91% of all PSTs' teaching actions were appropriate (see Table 1). This percentage, however, obscures the variation within the KQ dimensions and between PSTs. We highlight a number of notable results.

A first notable result is that six KQ codes were present in at most one fragment. Two of these codes, 'concentration on procedures' and 'making connections between procedures', are related to procedures. Given that ISI generally de-emphasizes procedures in order to bring the conceptual reasoning underneath statistical inference to the fore, this is not surprising. Also the codes 'responding to (un)availability of tools and resources' and 'teacher insight' cannot be expected to be present often: in actual classrooms both situations rarely occur. The fifth code that was absent is 'theoretical underpinning of pedagogy'. This can be explained by the limited training and background literature on ISI the PSTs had received. The sixth code, choosing and using examples, occurred only once. This would have required the PSTs to change the entire context of frequently used words in children novels to another context. Such change of the entire context would not have helped the children to attain the learning objectives. Also the use of mathematical language was largely absent from the lessons, apart from Alfred and Demi defining a sample. This is not surprising given that ISI is contextual and informal by nature and stimulates the use of context language rather than formal statistical language.

A second notable result is that in most fragments the PSTs did not deviate from the provided lesson plan, either on own their initiative or in response to children's input ('adherence to textbook', 'deviation from agenda', and 'decisions about sequencing'). Demi was the only one to deviate from the lesson plan regularly, and most of her planned and unplanned deviations did not contribute to the attainment of the learning objectives.

A third notable result is that, while Celine showed to be aware of the purpose of the particular lesson fragment in all but one fragment, this held for Alfred and Demi for around 70% of the fragments. In these cases, Alfred made use of extensive and aimless stocktaking of children's suggestions or failed to provide feedback to children regarding the conceptual appropriateness of their remarks. Demi sometimes discussed irrelevant issues from an ISI point of view, such as the exact order within the hypothesized top five of most frequently used words in children novels.

A fourth notable result is that teaching actions that require active expressions of conceptual insight into ISI ('teacher demonstration', 'making connections between concepts', and 'responding to students' ideas') prove to be more difficult than those that require only passive conceptual insight, ('awareness of purpose', 'identifying pupils' errors', and 'recognition of conceptual appropriateness'). Making connections between concepts is important in a lesson about ISI, in particular during the conclusion phase when children hopefully learn why a sample of books suffices for a conclusion about the entire population of books. However, in the PSTs' lessons, making connections rarely occurs, while the PSTs fail to provide convincing explanations how the various concepts connect. For example, the only instance Celine's subject knowledge falls short is when she fails to counter the children's unanimous opinion that all books need to be read. In general, overt expressions of subject knowledge that clearly extend beyond what can be expected of PSTs with similar training in ISI are scarce: Celine and Demi do so in only one fragment each. On the other extreme, overt shortcomings in subject knowledge, shown in obvious mistakes, are present in 23% of Alfred's and Demi's fragments. In case of conceptual input of children, Celine and Demi are (almost) flawless in analyzing the conceptual appropriateness of this input ('identifying pupils' errors, and 'recognition of conceptual appropriateness'), while Alfred does so correctly in 69% of the fragments.

A fifth notable result is that Alfred and Celine used instructional materials (i.e., pointing at the pile of books as the population, holding one book as a sample) in the majority of the fragments, and Demi only in a minority (23%) of the fragments. This difference may be explained by the fact that Alfred and Celine actually had piled up books in front of their classrooms, while Demi had left the books on the book shelves at the side of the classroom.

A sixth notable result is that in 98% of all fragments, the PSTs responded to students' ideas, underlining the constructivist nature of the lesson. Explanations were only present in a minority of the fragments, often in instances where children failed to come up with useful or correct suggestions or to where the PSTs countered incorrect suggestions. These explanations were correct in a majority of the cases, although Alfred scored lowest of three PSTs in this regard.

Table 1. Proportion present and appropriate KQ codes in fragments[1]

	Alfred		Celine		Demi		Total	
KQ code	**Pres.**	**Appr.**	**Pres.**	**Appr.**	**Pres.**	**Appr.**	**Pres.**	**Appr.**
Foundation	**0.20**	**0.87**	**0.22**	**0.96**	**0.29**	**0.90**	**0.34**	**0.87**
Adherence to textbook	0.15	0.92	0.08	0.92	0.32	0.77	0.21	0.85
Awareness of purpose	0.69	0.69	0.92	0.92	0.73	0.73	0.77	0.77
Concentration on procedures	0	1	0	1	0	1	0	1
Identifying pupils’ errors	0.46	0.69	0.46	0.92	0.73	1	0.58	0.90
Overt display of subject knowledge	0	0.77	0.08	0.92	0.05	0.77	0.04	0.81
Theoretical underpinning of pedagogy	0	1	0	1	0	1	0	1
Use of mathematical terminology	0.08	1	0	1	0.18	1	0.10	1
Transformation	**0.35**	**0.90**	**0.33**	**0.96**	**0.16**	**0.94**	**0.26**	**0.94**
Choice of examples	0.08	1	0	1	0	1	0.02	1
Choice of representations	0.23	0.92	0.23	1	0.14	0.95	0.19	0.96
Use of instructional materials	0.69	1	0.85	1	0.27	0.95	0.54	0.98
Teacher demonstration	0.38	0.69	0.23	0.85	0.23	0.86	0.27	0.81
Connection	**0.15**	**0.78**	**0.23**	**0.95**	**0.19**	**0.86**	**0.24**	**0.84**
Anticipation of complexity	0.15	0.62	0.38	1	0.27	0.77	0.27	0.79
Decisions about sequencing	0.08	1	0	1	0.14	0.91	0.08	0.96
Making connections between concepts	0.15	0.69	0.08	0.85	0.14	0.77	0.12	0.77
Making connections between procedures	0	1	0	1	0	1	0	1
Recognition of conceptual appropriateness	0.38	0.69	0.69	0.92	0.41	0.86	0.48	0.83
Contingency	**0.27**	**0.90**	**0.25**	**0.94**	**0.27**	**0.92**	**0.27**	**0.92**
Deviation from agenda	0.08	1	0.08	0.92	0	0.86	0.04	0.92
Responding to students’ ideas	1	0.62	0.92	0.85	1	0.82	0.98	0.77
Responding to (un)availability of tools and resources	0	1	0	1	0.05	1	0.02	1
Teacher insight	0	1	0	1	0.05	1	0.02	1
Total	**0.23**	**0.86**	**0.25**	**0.95**	**0.23**	**0.90**	**0.28**	**0.91**

[1]Alfred 13 fragments, Celine 13 fragments; Demi 22 fragments.

Relationship between Knowledge Quartet and ISI

Table 2 shows for each PST the relative presence of each of the ISI components across the fragments and the average appropriateness of the teaching actions according to the 14 KQ codes that are at least present twice. Generalization beyond the data was most often discussed (44% of the fragments); Sampling methods and Uncertainty least often (14%). The appropriateness of the teaching actions in fragments where at least one of ISI components was present was 86% on average, with small differences between the components (range: 79% to 91%) and between the PSTs, although Celine scored highest on average. Importantly, the ISI components Sampling variability and Uncertainty scored lowest, whilst these address the central issue why one can rely on the conclusion drawn based on the sample of books and why it is not necessary to read all books. Alfred did not address sampling variability at all.

Table 2. Proportion ISI component present and average appropriateness of KQ codes in fragments[1]

	Alfred		**Celine**		**Demi**		**Total**	
ISI component	**Pres.**	**Appr.**	**Pres.**	**Appr.**	**Pres.**	**Appr.**	**Pres.**	**Appr.**
Data as evidence	0.19	0.85	0.38	0.90	0.08	0.96	0.19	0.90
Generalization beyond the data	0.59	0.80	0.32	0.97	0.42	0.85	0.44	0.85
Sampling variability	0	na	0.17	0.84	0.26	0.78	0.17	0.79
Sampling methods	0.22	0.78	0.07	1	0.13	1	0.14	0.90
Sample size	0.15	0.89	0.24	1	0.15	0.83	0.18	0.91
Uncertainty	0.11	0.95	0.20	0.81	0.13	0.70	0.14	0.79
Total	0.21	0.83	0.23	0.92	0.19	0.84	0.21	0.86

[1]Alfred 13 fragments, Celine 13 fragments; Demi 22 fragments.

DISCUSSION

The results of the analysis of the ISI lessons of three pre-service primary school teachers, who participated in a teacher college intervention and who had little exposure to ISI, shows that a large majority of their teaching actions was appropriate. For some teaching aspects, this was due to the support offered by the provided lesson plan. Teaching actions that required active expressions of conceptual insight into ISI were less often appropriate than those that required only passive conceptual insight. In particular, the PSTs struggled to explain why it is not necessary to sample the entire population of books and that a sample of books suffice for a sufficient certain conclusion. This struggle appeared to be related to insufficient knowledge of the Sampling variability and Uncertainty components of ISI. Therefore, teacher college education should focus on how PSTs can foster children's understanding of how it is possible to a sample suffices to draw conclusion about a population.

REFERENCES

Ben-Zvi, D., Bakker, A., & Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics*, *88*(3), 291–303.

De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (2018). The development of informal statistical inference content knowledge of pre-service primary school teachers during a teacher college intervention. *Educational Studies in Mathematics, 99*(2), 217-234. doi:10.1007/s10649-018-9823-6

Gravemeijer, K., Stephan, M., Julie, C., Lin, F.-L., & Ohtani, M. (2017). What mathematics education may prepare students for the society of the future? *International Journal of Science and Mathematics Education, 15*(1), 105-123.

Groth, R. E., & Meletiou-Mavrotheris, M. (2018). Research on statistics teachers' cognitive and affective characteristics. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 327–355). Cham, Switzerland: Springer.

Leavy, A. M. (2017). Insights into the approaches of young children when making informal inferences about data. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the tenth congress of the European Society for Research in Mathematics Education* (pp. 788-795). Dublin, Ireland: DCU Institute of Education and ERME.

Liu, Y., & Grusky, D. B. (2013). The payoff to skill in the third industrial revolution. *American Journal of Sociology, 118*(5), 1330–1374.

Makar, K. (2016). Developing young children's emergent inferential practices in statistics. *Mathematical Thinking and Learning, 18*(1), 1-24.

Rowland, T., Huckstep, P., & Thwaites, A. (2005). Elementary Teachers' Mathematics Subject Knowledge: the Knowledge Quartet and the Case of Naomi. *Journal of Mathematics Teacher Education, 8*(3), 255-281. doi:10.1007/s10857-005-0853-5

APPENDIX 1. DEFINITION OF KNOWLEDGE QUARTET DIMENSIONS AND CODES

KQ dimensions and codes	Definition
Foundation	**PST's knowledge, beliefs and understanding acquired in the academy, in preparation for their role in the classroom**
Adherence to textbook	PST deviates from provided lesson plan, in planning or delivery of lesson
Awareness of purpose	PST is aware of the purpose of the lesson
Concentration on procedures	PST concentrates on applying or teaching procedures without an attempt to discuss the reasoning behind the procedure
Identifying pupils' errors	PST identifies children's mistake(s)
Overt display of subject knowledge	PST displays subject knowledge that clearly extends beyond or falls below (i.e. makes obvious errors) what can be expected from a teacher with the same background
Theoretical underpinning of pedagogy	PST shows deliberate use of theoretical insights regarding the teaching of the subject in planning or delivery of lesson
Use of mathematical terminology	PST uses mathematical terminology at a for the children appropriate level
Transformation	**PST's capacity to transform content knowledge into pedagogically powerful forms**
Choice of examples	PST selects and uses examples, other than suggested in the lesson plan
Choice of representations	PST selects and uses mathematical representations that fit the content to be explained
Use of instructional materials	PST uses instructional materials, either concrete or symbolic, to explain the content
Teacher demonstration	PST explains or demonstrates content
Connection	**PST's ability to bind together choices and decisions that are made for more or less discrete parts of mathematical content**
Anticipation of complexity	PST anticipates the complexity of the content by assessing whether children will be able to understand the content and by assessing possible misconceptions
Decisions about sequencing	PST introduces content, ideas and strategies in an appropriately progressive order (i.e., in an order that makes the content understandable to children)
Making connections between concepts	PST explicitly discusses (with the children) the connections between various mathematical concepts
Making connections between procedures	PST explicitly discusses (with the children) the connections between various mathematical procedures
Recognition of conceptual appropriateness	PST recognizes the conceptual appropriateness of children's remarks concerning mathematical concepts
Contingency	**PST's ability to act upon unplanned classroom events (i.e., to 'think on one's feet')**
Deviation from agenda	PST deviates from the lesson plan, as a result of what happens in class
Responding to students' ideas	PST responds to children's ideas or suggestions related to the content of the lesson
Responding to (un)availability of tools and resources	PST responds to unplanned and/or unexpected (un)availability of tools and resources
Teacher insight	PST reflects on his or her teaching during teaching ('reflection in action')