Learning opportunities for pre-service teachers to develop pedagogical content knowledge for statistical inference

Per Blomberg[1,2], Per Högström[1] and Yvonne Liljekvist[2]

[1]Halmstad University, Academy of Learning, Humanities and Society, Sweden; per.blomberg@hh.se, per.hogstrom@hh.se

[2]Karlstad University, Department of Mathematics and Computer Science, Sweden; per.blomberg@kau.se, yvonne.liljekvist@kau.se

Recently, researchers have encouraged the teaching of statistical inference to students at all levels. However, what constitutes pre-service teachers' pedagogical content knowledge for statistical inference has not yet been given specific attention in research. This paper presents a qualitative study of pre-service teachers participating in a collaborative learning setup in a mathematics course to be prepared for teaching statistics in primary school aged 6–10 years. The study reported here is the first cycle of a design research project. This first-phase study explores how pre-service teachers' pedagogical content knowledge for statistical inference can be developed during their mathematics course. The findings show that pre-service teachers' learning opportunities regarding pedagogical content knowledge for statistical inference are insufficient. Based on the initial results, an initial conjecture map was constructed that guides the forthcoming design cycle.

Keywords: content representation, design research, pedagogical content knowledge, pre-service teachers, statistical inference.

Introduction

Informal statistical inference (ISI) (Makar & Rubin, 2009) is often considered an essential ability of statistically literate citizens and the root of understanding formal inference (Biehler & Pratt, 2012). Much research focusing on ISI has been carried out with recommendations for powerful statistical ideas to introduce in primary school (Makar & Rubin, 2018). Despite prevalent international research findings, ISI has only made breakthroughs in the primary school curricula of a few countries. Therefore, researchers have called for ISI to be explicitly included in school curricula and teacher education (e.g., Langrall et al., 2017). However, in teacher education, limited time is spent on statistics. Thus, we need to learn to endorse ISI pedagogical knowledge for pre-service teachers through well-designed and effective arrangements (de Vetten et al., 2018). In line with Makar and Rubin (2018), we shift the focus from questions about what constitutes ISI to more knowledge of teaching statistical inference (SI) and design educational settings that provide prerequisites for pedagogical content knowledge (PCK) regarding SI.

Lehrer and English (2018) emphasised that children's informal reasoning about insecurity is formed both in and outside formal schooling. The same can be expected to apply to pre-service teachers, which makes it interesting to investigate the extent to which they pay attention to inference as a learning object in an existing learning environment. This paper presents the first cycle of a design research project (cf. Bakker, 2018), conducted in a mathematics course for pre-service teachers in Sweden. The design aims to provide knowledge to improve teacher education, from an educational setting focusing on statistical literacy to a setting supporting pre-service teachers' development of PCK for SI. We ask: How can a collaborative teaching strategy that mainly promotes statistical literacy skills support pre-service teachers' PCK for SI? Based on the findings, we propose an improved teaching strategy that is beneficial for developing pre-service teachers' PCK for SI.

Theoretical background

Statistical inference in statistics education

SI is usually associated with methods and concepts at the tertiary level, such as point interval estimates and hypothesis testing. By proposing a broader use of the term inference and making inferential reasoning available to all ages, Makar and Rubin (2009, p. 85) identified three key principles of ISI: "(1) *generalization*, including predictions, parameter estimates, and conclusions, that extend *beyond describing the given data*; (2) the use of data *as evidence* for those generalizations; and (3) employment of probabilistic language in describing the generalization, including informal reference to levels of certainty about the conclusions drawn". Acting on these principles can direct teachers' attention to children's abilities to communicate probabilistic predictions and generalisations from data. For example, Shaughnessy (2019) proposes a learning trajectory for the early years where children are initially encouraged to explore the use of statements such as 'likely' or 'unlikely', and thereafter propose and justify conclusions and predictions based on data, followed by designing studies for further investigation. Furthermore, Makar and Rubin (2018) acknowledge the *integration of contextual knowledge* and *consideration of the aggregate* as having key roles in making inferences. The proposed broad approach aligns with frameworks such as *data modelling* and the *statistical investigation cycle*, all of which incorporate the importance of teaching the overall picture of a statistical investigation rather than divided parts (Lehrer & English, 2018).

Several frameworks have been developed to describe a statistical investigative process. At the school level, the GAISE Pre-K–12 Report (Bargagliotti, 2020) proposes a model of four components: 1) Formulate Statistical Investigative Questions; 2. Collect/Consider the Data; 3) Analyse the Data; and 4) Interpret the Results. When composed into a somewhat simplified format that we believe can serve pre-service teachers who are to teach younger children, a statistical process follows three phases: data generation (e.g., problem context, statistical question, collect and organise data), analysis (e.g., use appropriate displays to represent data, distribution, variability, expected value), and communication (e.g., inference, generalisation, based on data, with uncertainty). In line with Shaughnessy (2019), we consider distribution and inference to be the heart and soul of statistics. However, these concepts are nurtured by the generated data, which in turn includes conceptions such as, for example, contextual knowledge, the statistical question, and how to measure, compile and organise the data.

To summarise, findings from prior research (e.g., Blomberg, 2015) suggest that *data*, *distribution*, and *inference*, each associated with the earlier phases, constitute the core of statistics in education at all stages, including younger students. For students of all ages to have access to the aforementioned desirable inferential ideas, pre-service teachers not only need content knowledge of statistics, but they also need PCK for SI. What constitutes PCK and how to measure PCK for SI in teacher education are presented in the following section.

The quality of pre-service teachers' pedagogical content knowledge

Since the introduction of PCK (Shulman, 1986), the concept has been interpreted and operated in various ways in different research domains. The resulting theoretical diversity of PCK models motivated Carlson and Daehler (2019) to develop a refined consensus model (RCM). In addition to *content knowledge, pedagogical knowledge, knowledge of students, curricular knowledge,* and *assessment knowledge*, this model also considers *collective* (*cPCK*), *personal* (*pPCK*) and *enacted* (*ePCK*) as aspects of teachers' professional knowledge. To shed light on how to support pre-service teachers' development, the RCM of PCK is expected to contribute to situating the research and identifying elements to study. Based on the RCM of PCK, Chan et al. (2019) suggested a framework for measuring the quality of teachers' PCK. The framework consists of the following five key components: 1) selection and connection of big ideas, 2) selection of instructional strategies and representations, 3) recognition of variations in student understanding, 4) integration between PCK components and 5) pedagogical reasoning. These PCK components are not intended to constitute a fixed framework but are rather open to adaptation within different contexts. Thus, they are flexible enough to measure the quality of different variants of pre-service teachers' PCK, such as what a group of pre-service teachers knows, what a single pre-service teacher knows and does, and expressed pedagogical reasons for judgement and action.

Methods

Context and participants

The RCM model of PCK centres around the teacher's ePCK, which can be characterised by planning, carrying out that plan and reflecting on instructions and student outcomes. However, in this study, the pre-service teachers' reports did not include an enacted teaching context with students in schools. The empirical data occurred when groups of pre-service teachers planned and reflected upon a hypothetical teaching situation. Thus, this study falls within the realm of cPCK, which is considered a result of transformations of pre-service teachers' pPCK in a learning context in which pre-service teachers initially individually reflect and thereafter collectively reflect and plan. As the study was conducted during the COVID-19 pandemic, a remote teaching setup was used. Lectures and group work were conducted online in digital meetings.

The participants in the study came from one class in teacher education for primary school that focused on becoming a teacher for students aged 6–10 years. The participants studied a course in mathematics, including statistics, during the second of four years of study. The class consisted of 62 pre-service teachers, 33 of whom signed an informed consent form to participate in the study. The mean age of the participants was 29 years, with a standard deviation of 7.7 years. The research project has been ethically reviewed and approved by the Faculty of Health, Science and Technology at Karlstad University.

Design, data, and analysis

The teacher educators have been involved in planning and evaluating the study's design and have provided significant knowledge of the inherent possibilities in hypothetical learning trajectories and design conjectures (Bakker, 2018). Hence, a research model was developed that matched the purpose of the study and the teacher educators' planning. The teaching of statistics was carried out in six steps over two weeks. The learning objectives with these six occasions can be summarised as emphasising statistical literacy, developing statistical thinking and conceptual understanding rather than knowledge of the modelling process and procedural skills. Teaching statistical literacy emphasises statistical knowledge and skills in contextual settings and the need to critically evaluate studies and reports based on statistics and to communicate critical ideas (Sharma, 2017). Such an approach calls for developing pre-service teachers' ability to plan, enact and evaluate teaching-learning settings that focus on developing skills to evaluate statistical studies' quality and findings rather than practical skills that come with performing studies.

The six occasions are briefly depicted next: 1) A homework assignment that consists of reading and performing tasks in teaching material on probability and statistics for teachers. The learning content consists of basic concepts from the descriptive statistics such as mean, median, mode, range, boxplot, and how to compile and represent data with different types of diagrams. In addition to a short section that addresses the challenge teachers face in choosing an appropriate activity for teaching statistics, the teaching material includes no more than a short repetition of what the pre-service teachers have encountered in primary and secondary school. 2) A lecture, which highlights the purpose of statistics in teaching and in everyday life; descriptions of the subject and central concepts (as mentioned above, reliability and validity), drawing on practical examples. 3) A homework assignment involving individual work of self-selected statistical material and reflections based on support questions from the teacher educator. One question concerns inference: "In what way can conclusions be drawn (generalise, predict, etc.) and with what certainty do you think it can be done?" 4) Group work: The participants in the study were randomly divided into 11 groups, with 3–4 members in each group. Each group chooses one of the group members' self-selected statistical materials from the third occasion. Each group completed a description of the content in a hypothetical teaching situation covering three big ideas for teaching statistics for students aged 6–10 years, supported by a so-called *Content Representation* (CoRe) (e.g., Hume & Berry, 2011). A CoRe is a template for teachers to portray big ideas of concepts and skills related to a particular topic with answers to critical pedagogical challenges. 5) Each group presents their planned teaching situation by recording a video. Finally, 6) Group work: Each group compiles written feedback for three films.

This study focuses on data materials from occasion four, the pre-service teachers' CoRes, compiled in groups. Using CoRe as a reflective tool can transform teachers' tacit pPCK into explicit cPCK (Alonzo et al., 2019). Therefore, the knowledge expressed in a CoRe that was put together by a group of pre-service teachers represents their cPCK (Carlson & Daehler, 2019). A *content analysis* approach was used in this study (Robson & McCartan, 2016). The analysis used the five key PCK components (Chan et al., 2019) as operationalised categories. Within each of these components, the pre-service teachers' written outcomes related to statistical conceptual ideas were identified. Our focus was on identifying content related to inference. The first author of this paper is responsible for the analysis. To ensure validity, the findings have been discussed with both teacher educators and co-authors.

Results

The first key PCK component concerns the selection, connection and coherence of big ideas. In total, the 11 groups proposed 29 big ideas. To get an idea of the pre-service teachers' knowledge of big ideas in statistics, their stated outcomes of big ideas were interpreted and placed in one of the following five categories and their accompanying distribution and examples of the stated big idea followed by the expected learning outcome: 1. Data (10): "How to measure and compare data – To be able to formulate a question to be able to find out a specific matter", "With the help of statistics, you can find out various questions", "Statistics is a collection ... – Collection is an important process; it is important for students to get an understanding of who to ask or what to look for", 2. Analysis (9): "Interpret – reading the diagram", "The bars represent data – To be able to create bar charts after collected data", "Analyse – comparison of the data presented in the diagrams", 3. Communication (0), 4. Critical (1): "Statistics can be misused and misunderstood - All available statistics should not be trusted. It is important to be critical of sources", and 5. Other topic (9): "Stress – We think that students should learn that stress can be both positive and negative depending on how you deal with it", "Risks with an increased use of the Internet", "Values – We want our students to gain knowledge about the Convention on the Rights of the Child". Summing up, the analysis of the pre-service teachers' CoRes showed that most of their stated big ideas connected to data generation, followed by analytical skills and other topic. One group stated a critical approach to statistics and no big idea was qualified in the category of communication.

The second key PCK component, teaching strategies, highlights the selection of instructional strategies and representations regarding the stated big ideas. Several groups suggest a modelling approach, meaning that students conduct their surveys and create diagrams. Furthermore, we can see traces of outcomes of PCK for teaching statistics, such as starting from everyday questions and showing examples of statistics that can be interesting, showing films about statistical investigation and practicing interpretations and comparisons of existing diagrams. We also note some outcomes of general pedagogical ideas, such as individual work, classroom discussions and a combination of these (individual-pairs-all), without relating to statistics.

The third key PCK component, students' understanding, draws attention to a student-focused classroom climate based on the individuals' knowledge and recognises the variation in students' learning. In addition to activities that engage students' interests, it is also desirable to uncover their thinking. Overall, pre-service teachers identify a wide range of student-focused factors that affect their teaching ideas at a general level. Raised challenges are, for example, different levels of students' preconceptions, language difficulties, lack of interest and unproductive attitude. Pedagogical ideas to address these issues were, for example, conducting investigation, diagram creation, presenting these to each other, making comparisons and discussing each other's presentations. When it comes to outcomes regarding inference, there are no traces of possible difficulties that students may have.

In the fourth key PCK component, integration between PCK components, attention is drawn to the teachers' ability to plan the next teaching step and adapt this to how the students have received teaching. Since the study does not include a context in which pre-service teachers practice teaching, it is impossible to infer anything about their adjustment skills within this component. However, based on existing data, it is possible to identify pre-service teachers' presumptive ideas about obtaining students' learning big ideas and adjusting teaching practices. For example, some groups expressed ideas about controlling how students design, interpret, and analyse diagrams. However, the present data cannot explain how they intend to use this evaluation in their teaching.

The fifth key PCK component, pedagogical reasoning, focuses on pedagogical issues as a teacher's ability to justify decisions and actions within a teaching situation. Supporting questions in the CoRe evoked pedagogical reasoning and thus became applicable as an empirical basis for this fifth component. The arguments for the selected ideas are anchored in the importance of being a competent consumer of statistics. It was seen as important that the students understand the purpose of using, for example, diagrams and being critical of the statistics presented to them. Some state the importance of students engaging with an investigation, compiling diagrams and interpreting and drawing conclusions based on the results. Both teaching methods and the arguments behind these choices vary.

Some groups highlighted the importance of motivating students to learn and suggested using up-to-date social issues and topics that were close to the students' interests. Others mentioned teaching methods such as classroom discussion to introduce students to making correct interpretations and conclusions and equal collective understanding to reduce the variation in students' knowledge in the classroom. In addition to these outcomes of reasoning, most of the pedagogical reasonings was inadequate or missing.

Conclusion and discussion

Regarding the implications and limitations of the study, we would like to draw the reader's attention to the fact that the study should not be regarded as an isolated case study but as a first sub-study of design research consisting of several iterative studies (Bakker, 2018). The study was conducted in a specific context that limits the generalizability of the results to similar conditions. The study was also limited to one group of students who carried out all work during remote teaching due to the pandemic. The dropout rate was 47 %, most likely due to the pandemic.

Regarding the question of how the collaborative teaching strategy promotes pre-service teachers' PCK, we see, with the support of Carlson and Daehler (2019), that when students in groups plan to teach and then assess each other's work, they are offered the opportunity to exchange knowledge from pPCK to cPCK and back to pPCK on several occasions. In addition, students acquired professional knowledge bases such as content knowledge about statistics and curriculum knowledge via the internet. The results from the study depict a somewhat fragmented picture of the pre-service teachers' knowledge about big ideas in statistics education, as emphasised by, for example, Shaughnessy (2019). The pre-service teachers' outcomes regarding inference in accordance with Makar and Rubin's (2018) and Shaughnessy's (2019) ideas are almost non-existent. However, there is a clear emphasis on compiling and organising data, interpreting data, and being statistically literate. It is noteworthy that the setting triggered nearly half of the participating groups to choose topics from statistical contexts, separate from statistics, as big ideas. Indeed, *knowledge of the problem context* is a central conception in statistics education, and the stated big ideas, which were identified as statements about societal issues, are relevant to students. However, in this learning context, statistics was the subject of focus. Therefore, the results indicate that learning objectives for statistics are partly outperformed by learning objectives for other topics. Based on these results, we infer that a statistical literacy approach to teaching does not provide satisfactory outcomes in pre-service teachers' PCK for SI. Therefore, we suggest a draft of a conjecture map (Table 1) to guide the next cycle.

In accordance with Lehrer and English (2018), our suggestion for an improved strategy for high-level teacher educators with high-level conjecture is to emphasise a more explicit focus on practicing data modelling and highlight what characterises SI in education (Makar & Rubin, 2018). We choose to complement the design conjectures of existing tools (e.g., CoRe) and participant structures (collaborative learning) within the embodiment of putting the conceptual framework of ISI modelling into practice (Blomberg, 2015). The design includes a lesson plan in which the development of content knowledge encompasses PCK. Our ideas of theoretical conjectures, meaning the produced desired outcomes from the mediating processes, include appropriately emphasising inference as a big idea, describing its key principles in accordance with Makar and Rubin (2009), and connecting inference to other core statistical conceptions such as data generating and distribution. Furthermore, according to Table 1, we suggest a designed mediating process that produces outcomes for how SI can be made teachable for the age group in focus, as well as pedagogical reasoning for selected teaching methods at both class and individual levels. The initial conjecture map presented in Table 1 will guide the forthcoming design cycle of our design research project. The plan is to carry out a second-phase study at the end of 2021. We intend to present preliminary results from the forthcoming study at CERME 12.

Table 1: Conjecture Map - Statistical Inference Modelling Teaching

High-level conjecture	Embodiment/design	Mediating processes	Pre-service teachers' outcomes
Conceptual frameworks of ISI modelling in education, and cPCK can contribute to pre-service teachers' ability to plan and reflect on teaching SI appropriate for a particular age.	• Putting ISI modelling into practice. • Completing a CoRe. • Planning for a hypothetical lesson or lesson sequence in statistics presented via video recording. • Using peer review. • The development of content knowledge encompasses PCK. • Collaborative learning as a participant structure.	• Experiencing ISI modelling cycle. • Identifying inference as a core concept and its connections to other big ideas in statistics education. • Expressing in speech and writing about planning for SI modelling teaching. • Reflecting in writing on other students' video presentations.	• Describing inference as a big idea together with other big ideas and concepts in statistics education. • Specifying how SI can be made teachable for the age group in focus. • Arguing for selected teaching methods at both class and individual levels.

Acknowledgments

This work was funded by the Swedish Research Council [2019–04419], a graduate school at Karlstad University, Halmstad University, and University West.

References

Alonzo, A., Berry, A., & Nilsson, P. (2019). Unpacking the Complexity of Science Teachers' PCK in Action: Enacted and Personal PCK. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 271–286). https://doi.org/10.1007/978-981-13-5898-2_12

Bakker, A. (2018). *Design Research in Education: A Practical Guide for Early Career Researchers*. Routledge.

Bargagliotti, A. F., C.; Arnold, P.; Gould, R.; Johnson, S.; Perez, L.; Spangler, D. (2020). *Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II) A Framework for Statistics and Data Science Education*. A. S. Association. https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf

Biehler, R., & Pratt, D. (2012). Research on the reasoning, teaching and learning of probability and uncertainty. *ZDM, 44*. https://doi.org/10.1007/s11858-012-0468-0

Blomberg, P. (2015). *Informell statistisk inferens i modelleringssituationer - en studie om utveckling av ett ramverk för att analysera hur elever uttrycker inferenser* [Informal statistical inference in modelling situations - a study of developing a framework for analysing how students express inferences]. Linnéuniversitet. http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-45572

Carlson, J., & Daehler, K. R. (2019). The Refined Consensus Model of Pedagogical Content Knowledge in Science Education. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 77–94). Springer. https://doi.org/10.1007/978-981-13-5898-2_2

Chan, K., Rollnick, M., & Gess-Newsome, J. (2019). A Grand Rubric for Measuring Science Teachers' Pedagogical Content Knowledge. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 251–269). https://doi.org/10.1007/978-981-13-5898-2_11

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. https://doi.org/10.1007/s10649-018-9823-6

Hume, A., & Berry, A. (2011). Constructing CoRes - a Strategy for Building PCK in Pre-service Science Teacher Education. *Research in Science Education*, *41*(3), 341–355. https://doi.org/10.1007/s11165-010-9168-3

Langrall, C. W., Makar, K., Nilsson, P., & Shaughnessy, J. M. (2017). Teaching and learning probability and statistics: An integrated perspective. In Cai, J. (Ed.), *Compendium for Research in Mathematics Education* (pp. 490–525). The National Council of Teachers of Mathematics.

Lehrer, R., & English, L. (2018). Introducing Children to Modeling Variability. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 229–260). Springer. https://doi.org/10.1007/978-3-319-66195-7_7

Makar, K., & Rubin, A. (2009). A Framework for Thinking about Informal Statistical Inference. *Statistics Education Research Journal*, *8*(1), 82–105.

Makar, K., & Rubin, A. (2018). Learning About Statistical Inference. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 261–294). Springer. https://doi.org/10.1007/978-3-319-66195-7_8

Robson, C., & McCartan, K. (2016). *Real world research*. John Wiley & Sons.

Sharma, S. (2017). Definitions and models of statistical literacy: a literature review. *Open Review of Educational Research*, *4*(1), 118–133. https://doi.org/10.1080/23265507.2017.1354313

Shaughnessy, J. M. (2019). *The Big Ideas in the statistics education of our students: Which ones are the biggest?* XV Inter-American Conference on Mathematics Education, Medellín, Colombia.

Shulman, L. S. (1986). Those Who Understand: Knowledge Growth. *Educational Researcher*, *15*(2), 4–14.