RESEARCH ON EARLY STATISTICAL AND PROBABILISTIC THINKING

SPECIAL ISSUE OF THE *STATISTICS EDUCATION RESEARCH JOURNAL* (*SERJ*)

Guest Editors

AISLING LEAVY
Mary Immaculate College, Limerick
Aisling.Leavy@mic.ul.ie

MARIA MELETIOU-MAVROTHERIS
European University Cyprus
M.Mavrotheris@euc.ac.cy

EFI PAPARISTODEMOU
Cyprus Pedagogical Institute
e.paparistodemou@cytanet.com.cy

Special Issue Editor

DANIEL FRISCHEMEIER
University of Münster
dfrische@uni-muenster.de

"Today's students need to learn to work and think with data and chance from an early age, so they begin to prepare for the data-driven society in which they live" (Ben-Zvi, 2018, viii).

As we consider this quote by Dani Ben-Zvi, we are living in times of the COVID-19 pandemic, the Ukraine crisis, and in the era of "Fake News and Alternative Facts." Within the context of these global events, the need for data and chance literacy is particularly warranted and compelling. A vibrant and strong democracy needs well-educated and active citizens who are competent in reading and understanding data because in the age of Data Science and the omnipresence of Big Data, statistical thinking and competent handling and consideration of data have become critical (Engel, 2017). Indeed, in 2016, the National Science Foundation (NSF) identified, "Harnessing the Data Revolution" as one of the 10 big ideas that will "push forward the frontiers of U.S. science and engineering research, and lead to discoveries and innovations."

The book entitled, *Statistics in Early Childhood and Primary Education,* edited by Leavy et al. (2018), has both drawn attention to and shed a spotlight on the critical importance of formative experiences with data and probability in the early years. The contributions to this book, from all around the world, represent a milestone in this field and attest to the growing appreciation internationally for high-quality teaching and learning experiences that bring statistical and probabilistic thinking into early childhood and primary classrooms. The aim of this *SERJ* Special Issue is to connect with these ideas and synthesize innovative pedagogical approaches and research on early statistical and probabilistic thinking.

The Special Issue includes contributions that consider curriculum issues, interrogate new developments, and discuss the state-of-the-art and future trends that could inspire ideas for teaching, learning, and research on early statistical and probabilistic thinking from early childhood to upper primary levels of schooling. When generating the call for papers, we posed these questions to guide, direct and stimulate focus on what we considered the emerging areas of scholarship:

- How do we prepare young children to deal with the ambiguity and uncertainty that is fundamental to working in data and chance environments?
- What knowledge, skills and dispositions support young children in developing statistical and probabilistic thinking and reasoning?
- What role should data and chance play in mathematics education in the early years of schooling?
- What are new and innovative ways to engage young learners in studying data and chance?
- How might the emerging curriculum area of Data Science influence how we think about early statistical and probabilistic thinking and reasoning?
- What are effective ways to support early childhood and primary teachers in providing high-quality data and chance experiences in the classroom?
- How can we leverage the potential of technology to support the teaching and learning of data and chance in the early years?
- What role can online or hybrid learning play in a post-COVID-19 transformed system?

This *SERJ* Special Issue on Early Statistical and Probabilistic Thinking consists of thirteen contributions which can be categorized as follows:

1. Contributions to research on statistical and probabilistic thinking in the early years (learners aged from 3 to 7 years).
2. Contributions to research on statistical and probabilistic thinking at the elementary level (learners aged from 8 to 12 years).
3. Looking beyond the learner aged 3–12: contributions to teacher education and curriculum on statistical and probabilistic thinking.

In the first category—Contributions to research on statistical and probabilistic thinking concerning learners aged from 3 to 7 years—the focus is on early statistical learners at the pre-elementary level. The research study of Zoi Nikiforidou and Jennie Jones explored the probabilistic thinking of 3–4-year-old children in a school setting and suggested that children's authentic understandings of probability can serve as a foundation for more instructional pedagogical sequences. The study emphasized the importance of providing opportunities for young children to engage in problem-solving, creative play, and consideration of risk issues in developing their probabilistic thinking. The next three papers in this category place a spotlight on the critical role and affordances played by meaningful contexts and representations in supporting young learners in thinking and reasoning about data. The research report of Virginia Kinnear investigated how 5-year-old children consider data represented in tables. She reports that the young children in her sample can successfully interpret data when presented in pictorial and numerical forms and that they draw on their understanding of the context to reason about and make sense of the data. The research of Kinnear highlights the importance of using appropriate and accessible data representations and the context of story books to support children's reasoning with data. Representation of data is also the focus of the study by Gamze Kurt. Her paper reports on the investigation of 5–6-year-old children's ability to reason with bar graphs and pie charts in probabilistic reasoning contexts. She argues that young children are capable of informal inferential reasoning concerning the concepts of sample space, event probability, and data representations. The study highlights the potential of using visual representations to support children's understanding of probability concepts. Finally, in this first category, the study of Lucia Zapata-Cardona examined 7-year-old children's ability to model data and make informal inferences. Her research study found that the context of problem activities can facilitate statistical reasoning and help children develop strategies to identify attributes of data, assess the model created, and make sense of the data. This study emphasizes the importance of authentic contexts in promoting children's statistical reasoning.

Our second category focuses on the research on statistical and probabilistic thinking of older elementary children (learners at the age of 8 to 12 years). The study of Roberta Schnorr Büehring and Regina Grando provides insights into how a Brazilian public school conducts civic statistics teaching practices with multivariate data. More specifically they report on 7–8-year-old children's abilities to reason with multivariate data using *Dollar Street* [https://www.gapminder.org/dollar-street]. Büehring and Grando contend that children can expand their awareness of themselves and the world through statistics and that promoting the learning and use of statistical language can facilitate their reasoning

with data. One important finding is that statistical literacy is fundamental for children to read the world in today's context.

Lyn English investigated 8–9-year-old children and how they selected and justified their preferred representations of data in probability settings. More specifically, the children chose their samples of 12 colored counters, ensuring all colors were represented. They were then required to predict the outcomes of item selection, test their predictions, explain the outcomes, quantify their chances of color selections, and finally create two representations displaying the probabilities. In general, the forms of inscription included revealed a range of probability and statistics understanding of the young learners. The study found that children favored both bar and circle graphs and highlighted the importance of providing children with multiple representations to develop their understanding of data.

Taking on a semiotic perspective, Lara Billion explored whether different materials—analog or digital—can influence learners' mathematical interpretations. The study found that some materials can shorten actions and automatically establish mathematical relationships, while different actions on different materials can also lead to the same diagram interpretations. This study emphasizes the importance of using appropriate materials to support children's reasoning with mathematical concepts.

The study by Celi Lopes, Adriana Augusto, and Sezilia Toledo considered the development of statistical and probabilistic reasoning in childhood from an interdisciplinary perspective at a Brazilian school. Lopes et al.'s study involved the disciplines of life sciences, mathematics, and statistics. One finding of the case study presented was that the approach taken can promote 10-year-old children's development of statistical and probabilistic reasoning and positively influence their eating habits. This study highlights the potential of interdisciplinary approaches to promote children's reasoning abilities and connect mathematical concepts with real-world applications.

Informal statistical inference was the focus of Soledad Estrella's research with Grade 3 students. A learning trajectory was used to orient grade 3 students' introduction to informal inferential reasoning. More specifically, the twelve third-grade students of Estrella's study were introduced to sampling, frequency distribution, randomness, and sampling variation as well as to developing a data sense in online lessons. The results of this type of teaching illustrate that the creation and collection of authentic data in a playful context, together with exploratory analysis of the data as a precursor to utilizing aspects specific to informal inferential reasoning, promoted the statistical reasoning components of students.

The research study presented by Per Nilsson reports on a design experiment for a three-step hypothetical learning trajectory on informal hypothesis testing where students (11–12-year-old) played the *Color Run* game. In the accompanying research, it was observed that the students came to favor sample space reasoning over idiosyncratic reasoning when the sample space was changed between color runs. In the second and third steps, students were using degrees of variation in the distribution of the mode across samples to infer whether an unknown sample space is uniform. One fundamental finding concerning students' reasoning was the learners' argument that the larger the variation, the greater the reason for rejecting a uniform sample space.

The third category extends beyond the research on the early statistical and probabilistic thinking of learners and deals more generally with preservice teachers (de Vetten et al.), curricula analysis (Hijazi), and a systematic literature review (Denton)—all in the context of early statistical and probabilistic thinking. In the research of Arjen de Vetten, Ronald Keijzer, Judith Schoonenboom, and Bert van Oers, preservice primary school teachers are in the spotlight. De Vetten et al. investigated pre-service teachers' ability to teach informal statistical inference. Using the informal statistical inference (ISI) framework developed by Makar and Rubin (2009) and the Knowledge Quartet framework by Rowland et al. (2009) as the theoretical basis, the study found that pre-service teachers engaged in making inferences based on sample data and struggled to correctly interpret students' conceptual input and explain ISI. The study highlights the need for effective teaching strategies to improve pre-service teachers' understanding of statistical inference to ensure they are prepared for teaching statistical concepts to young children in the future.

The two other contributions of this special issue deal with a review of curricula (Hijazi) and a systematic literature review (Denton). Rafiq Hijazi and Ali Shaqlaih's research focused on curriculum analysis of textbooks for Grades 1–4. The study found that the textbooks provided limited opportunities for students to develop their statistical reasoning skills and that the content was often too abstract or too advanced for the targeted age group. This study underscores the importance of effective curriculum design to ensure that students receive appropriate instruction in statistical reasoning.

Finally, David Denton conducted a systematic review of the literature on teacher preparation standards in the context of teacher education in early statistical and probabilistic thinking. One important finding was that the literature emphasizes the development of probability and statistics concepts and procedures on the one hand, but on the other hand, there is an underrepresentation of the development of pedagogical knowledge, of learning from school experiences and of reflection on practice.

All these thirteen contributions from different perspectives, with different backgrounds, with different learners, and with different approaches to early statistical and probabilistic thinking make this *SERJ* issue special. This special issue, however, should only be seen as the beginning of further and deeper design and empirical research in early statistical and probabilistic thinking. More qualitative, design-based, and quantitative research studies are needed to investigate the integration of early statistical and probabilistic thinking for learners at the age group of 3 to 12 years. Also, the implementation of digital tools, the impact of contextual knowledge, and ways of supporting young learners to document and communicate their findings in statistical explorations must be the focus of future research, as well as the potential of teacher education and teachers' professional development in this field. We wish you a pleasant journey when reading this special issue on research on *Early Statistical and Probabilistic Thinking*.

ACKNOWLEDGEMENTS

REFERENCES

Ben-Zvi, D. (2018). Foreword. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds.), *Statistics in early childhood and primary education: Supporting early statistical and probabilistic thinking* (pp. vii-viii). Springer.

Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal, 16*(1), 44–49. https://doi.org/10.52041/serj.v16i1.213

Leavy, A., Meletiou-Mavrotheris, M., & Paparistodemou, E. (Eds.) (2018). *Statistics in early childhood and primary education: Supporting early statistical and probabilistic thinking*. Springer. https://doi.org/10.1007/978-981-13-1044-7

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal, 8*(1), 82–105. https://doi.org/10.52041/serj.v8i1.457

National Science Foundation. (2016). *NSF's 10 big ideas*. https://www.nsf.gov/news/special_reports/big_ideas/index.jsp

Rowland, T., Turner, F., Thwaites, A., & Hurst, C. (2009). *Developing primary mathematics teaching: Reflecting on practice with the knowledge quartet*. SAGE Publications. https://doi.org/10.4135/9781446279571