# Introduction to the work of TWG05: Probability and Statistics Education

Daniel Frischemeier, Orlando Rafael Gonzalez, Aisling Leavy, Caterina Primi, Sibel Kazak

# Introduction to the work of TWG5: Probability and Statistics Education

Daniel Frischemeier[1], Orlando Rafael Gonzalez[2], Aisling Leavy[3], Caterina Primi[4] and Sibel Kazak[5]

[1]Westfälische Wilhelms-Universität Münster, Germany; dfrische@uni-muenster.de

[2]Assumption University, Graduate School of Human Sciences, Thailand; ogonzalez@au.edu

[3] Mary Immaculate College, University of Limerick, Ireland; Aisling.Leavy@mic.ul.ie

[4]Neurofarba - University of Florence, Italy; caterina.primi@unifi.it

[5]Pamukkale University, Turkey; skazak@pau.edu.tr

## Introduction

The working group gathered 39 participants from 14 countries, from Asia to South America, with 26 papers and six posters. The participants were introduced to the three subthemes identified as emerging across all submissions. In particular, submissions focused on "Teacher education", "Reasoning about data" and "Statistical and Probabilistic Thinking and Reasoning". We prepared three guiding questions to support in the discussions on the three subthemes. The questions were: "*Considering these presentations, what contribution do they make to what we now know about teacher education/reading the data/statistical and probabilistic reasoning?*", "*Considering these presentations, what issues (regarding (mis-)conceptions, teaching approaches, resources, assessment etc.) do they highlight/raise about teacher education/reading the data/statistical and probabilistic reasoning*?" and "*What comments or thoughts do you have now with regard to teacher education/reading the data/statistical and probabilistic reasoning, from a research and practice point of view, to advance future research?*".

### Teacher education

There were ten papers focusing on the subtheme of *Teacher Education.* They represented a broad gamut of research examining the disciplinary foci of statistics, probability and STEM. These research studies considered the experiences of students, preservice and practicing teachers across various educational settings, including early childhood, primary, secondary and college-level learning environments. They also explored a variety of learning opportunities to support the development of understanding including the use of innovative technologies, lesson study and a range of professional development opportunities. One of the themes that emerged from the research presentations was the need for researchers to engage in more effective communication with statistics and probability educators. The field is a fast-evolving landscape and new knowledge and understandings about pedagogies are constantly being developed. Our responsibility is to assist teachers in identifying these critical and desirable facets of statistics and probability education. This communication should focus on:

(1) The purpose and rationale for teaching statistics education. We need to remind teachers that our goal as educators is to engage students in statistical thinking and reasoning, develop statistical literacy and learn how to critically read the world and how we represent it statistically.

(2) Identifying the desirable feature of good statistics teaching. This includes teaching statistics through projects and investigations, open questions, and the need to address meaningful problems. We need to better communicate and provide a justification for how these desirable practices relate to the previously expressed purpose and rationale of statistics education

(3) Explaining the limitations of some common approaches to teaching statistics and probability. Examples are a focus on technical procedures, calculations of descriptive statistics in isolation from developing conceptual understanding and the use of closed questions that don't promote inquiry and a questioning stance.

(4) Providing insights into what it feels like to engage in these desirable pedagogical practices. It is important that teachers are provided with rich experiences of engaging in statistical investigations so that inert theoretical knowledge becomes meaningful. It is very valuable for teachers to have the opportunity to experience the practices of statisticians and experience the benefits for the learner of engaging in these practices.

A second theme that arose was recognition that statistics and probability education continue to evolve rapidly. Consequently, we need to continually remain open to and expand our conceptualization of the field. Moving forward the statistics and probability educational research community need to better leverage the use of contexts in our practices. This involves moving beyond meaningless and trite textbooks problems to a focus on addressing meaningful problem that may have personal relevance to learners or engage them in considering real life global problems and challenges. Many of these types of problems occur in rich interdisciplinary contexts such as STEM and embracing these new contexts provides numerous opportunities for the development and use of innovative pedagogies. Some of these opportunities include a move to new technologies and tools that allow us to explore large and open-source data sets that represent real problems and provide new lenses and ways of visualising data. Considering these responsibilities as researchers to embrace and manage change, we identified the emerging challenge for us as statistics and probability educators to balance context and at the same time ensure a focus on desirable learning outcomes. We communicated this concern and culminated our discussions in the form of a driving question for future consideration: How can we foreground statistics and probability when engaged in interdisciplinary collaborations and at the same time not lose sight of the important contexts that drive and underpin investigations and inquiry?

**Statistical and Probabilistic Thinking and Reasoning**

Eleven papers were presented under the subtheme of Statistical and Probabilistic Thinking and Reasoning. Five of those focused on probabilistic reasoning in relation to risk perception, decision making under uncertainty, random variation and sample space, covariation tasks in a Bayesian situation, and combinatorics. In these studies, the participants were primary, secondary, and university students. The other six papers investigated students' statistical conceptions with regard to frequency tables, statistical content in terms of statistical literacy and reading levels of statistical

graphs in the mathematics textbook, the exploration of real and rich data with the use of technology tools and communication of results. Most of these studies involved secondary school students, but one study focused on primary school students. Textbook analysis research included primary and secondary grades.

From the discussions of these presented papers, three main themes emerged: 1) The emphasis on statistical reasoning in all phases of statistical investigation cycle and consideration of individual's dispositions like attitude towards statistics led to a broader view of statistical reasoning; 2) The research suggested the importance of rich learning experiences in support of critical thinking through the use of meaningful context and real data sets, the implementation of project-based learning in statistics education, and promoting probabilistic reasoning in data-based decision-making processes. 3) Fostering informal ideas with regard to statistical and probabilistic reasoning with younger learners is still relevant. In addition, some concerns have been expressed about the readiness of teachers and students for open-ended tasks suggested by research as well as the limited sources and time for implementing statistical projects. Critical discussions on the future of statistics and probability education have raised the questions of (1) relying on "better" textbooks versus "more" digital tools in school education, (2) communicating appropriate use of probability language, especially related to the everyday language (e.g., chance, luck, randomness), in the classroom, and (3) increasing the role of researchers in task design, textbook and curriculum development related to probability and statistics.

**Reasoning about data**

Within the five papers in the section "reasoning about data" four papers have had a specific focus on Data Science Education and to related fields like big data and machine learning. We discussed several core ideas and fundamental aspects to develop a competent reasoning about data in a sustainable way and across all age levels.

(1) At first we identified that it is important that there is a continuous development of data competence (from primary school to adult education) in the sense of a spiral curriculum.

(2) Second, the appropriate use of digital tools (like Gapminder, CODAP and TinkerPlots) can reduce the extraneous load in working processes and make learners able to explore large and multivariate data and to explore their data with regard to their specific inquiry questions. One crucial point in this respect is the choice of the digital tool. Educational software like TinkerPlots do not need a specific programming language (but are limited in some sense with regard to the data exploration capabilities). In contrast professional software like R or Python offer a broad range and landscape of statistical activities, but are more difficult to learn and learners may concentrate on programming and technical issues rather than on the content and the statistical exploration. So there is the danger that technical issues distract from the content issues.

(3) A third core idea to develop a competent reasoning about data, which was raised in the discussion, was the cooperation with other disciplines, e.g., the STEM disciplines.

(4) A fourth point is that the kind of data which is used for teaching and learning issues plays an important role: learners should be given real, meaningful data, which is authentic and offers

multivariate explorations. In addition to that the role of task design is central. In the discussion it was mentioned that teachers are often not comfortable with complex open problems which do not show clear steps to solve the task. To prepare teacher to consider using open tasks and problems and to make them familiar with these kinds of tasks was identified as a huge task for teacher education.

(5) A last, but fundamental issue with regard to reasoning about data which was mentioned was Data Science. Data Science was figured out to be an emerging field in statistics education and includes aspects like Big data, Open data, other data collection methods (e.g., Sensors, Webscraping). These new concepts, issues and ideas of Data Science led to the re-interpretation of fundamental ideas and concepts in statistics education (e.g., PPDAC cycle).

Looking ahead in the context of Data Science new approaches like Machine Learning in education are very new topics and one should use the opportunity to share all the different approaches arising and to include all the things we "know" by now (open projects, problem oriented learning, using new technology…). Given that four of five paper in the rubric "Reasoning about data" have tackled issues with regard to Data Science, we see that that this topic becomes more and more important in the statistics education landscape. Looking forward three big issues were identified which seem to be very relevant for a future perspective on reasoning about the data. Specifically more qualitative, design-based and quantitative research is needed in the following three fields: integration of Data Science into the classroom, connection of informal and formal concepts for reasoning about data and the connection of data, chance and context.

**Organization of the TWG sessions**

In the first session, we explained the organization of the Sessions. We were divided into two groups for some sessions: TWG5a with Caterina Primi, Sibel Kazak and Orlando Rafael Gonzalez and TWG5b with Aisling Leavy and Daniel Frischemeier. Each group used breakout rooms to create smaller groups. In this way, also with virtual modality, we tried to promote the famous three Cs of CERME Communication, Cooperation and Collaboration. To create a collaborative atmosphere that would support the discussions and feedback over the following days, we started with an ice-breaker activity "Speed dating" involving all the participants. Participants got to know their TG5 colleagues in a friendly context through this activity. We created groups of three randomly and for 3 minutes each participant introduced themself shortly before swapping to a new random group. At the end of each session for each group, a Padlet has been created to document and to write down the results of the discussion. There was a corresponding card in the Padlet for each of the guiding questions. In addition to that, there were two further cards for raising other issues, thoughts, etc. regarding the papers. All the contributions we collected were significant for a successful and substantial concluding discussion on the last day of CERME-12 when we summed up all our insights during the CERME-12 week. The last day we had a culminating session with all participants with the aim being to engage in an in-depth discussion on subthemes and to share the contributions discussed during each session.