THE INTERNATIONAL RESEARCH FORUMS ON STATISTICAL REASONING, THINKING AND LITERACY: SUMMARIES OF PRESENTATIONS AT STRL-2[1]

SUMMARY

This report includes a brief introduction to the Statistical Reasoning, Thinking and Literacy (SRTL) Research Forum Program, an overview of SRTL-2 and STRL-3, and summaries of each of the SRTL-2 presentations.

Keywords: *statistics education, SRTL forum, statistics, reasoning, thinking, literacy.*

1. INTRODUCTION

CHRIS READING
University of New England, Australia
creading@metz.une.edu.au

1.1. THE SRTL RESEARCH FORUM PROGRAM IS BORN

The International Conferences on the Teaching of Statistics (ICOTS), held every four years, beginning in 1982, helped to progressively link an informal research network of people interested in carrying out research on the teaching and learning of statistics at all age levels. Several papers at ICOTS-5, held in June 1998 in Singapore, focused on the related topics of Statistical Reasoning, Statistical Thinking, and Statistical Literacy. Discussions about these papers and the need to distinguish between the terms “literacy”, “reasoning” and “thinking” led to the First International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-1) which was held in Israel in July of 1999.

In July of 1999, sixteen statistics educators from six different countries met for five days at Kibbutz Be’eri in Israel to discuss the topics of statistical literacy, reasoning, and thinking. The Forum was co-chaired by Joan Garfield (University of Minnesota, USA) and Dani Ben-Zvi (Weizmann Institute of Science, Israel). One unique feature of this Forum was the use of videos of classroom work or interviews with students, as a way to present, discuss and argue about research related to these topics. Sponsors of SRTL-1 were the International Study Group for Research on Learning Probability and Statistics, the International Association for Statistics Education, the University of Minnesota, the Maurice and Gabriela Goldschleger Conference Foundation at the Weizmann Institute of Science, and Kibbutz Be’eri.

Discussions and research at this first forum revolved around a number of issues. An attempt was made to clarify what constitutes statistical reasoning, thinking and literacy (SRTL), how are these cognitive processes and/or outcomes different and how are they related? In particular, the first forum aimed to address the following questions:

- What does research on SRTL tell us about the learning and teaching of statistics? What are the cognitive, socio-cognitive, or developmental aspects of learning SRTL in different age/grade levels?
- What theoretical frameworks and methodologies are appropriate for researching SRTL?
- What types of qualitative and quantitative research studies are needed to help us better understand these ways of processing information and to help promote them in educational settings? Particularly, how do we collect, use and analyse video material for research on SRTL?
- What are the implications of research into SRTL for learning goals, curriculum design, and assessment?

Participants were asked to bring video clips and transcripts to illustrate different types of statistical literacy, reasoning and thinking. The group met all together to view videos and discuss the various research projects. SRTL-1 was a first attempt to focus a research forum on the interrelated and often poorly defined topics of statistical literacy, reasoning, and thinking. Participants of SRTL-1 indicated that the meeting’s format enabled good discussion, and that it was stimulating and enriching to become acquainted with key researchers in this area

[1] *Statistics Education Research Journal 1(1), 30-45, http:/fehps.une.edu.au/serj*
International Association for Statistical Education

and to view their work in progress. Many especially valued the small size that allowed plenty of time for interaction and discussion, the use of video in the talks and the extended discussions of videotaped sequences.

It was clear from participants' evaluations that this was only the beginning of a very exciting and promising line of research, and that there was much work to be done, so a second gathering (SRTL-2) was planned in Australia, with a similar format but improved structure.

1.2. SRTL-2 - ARMIDALE - 2001

The Second International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-2) was held at the University of New England (UNE), in Armidale, Australia from August 15 to August 20, 2001. The Forum was co-chaired by Joan Garfield (University of Minnesota, USA), Dani Ben-Zvi (Weizmann Institute of Science, Israel) and Chris Reading (University of New England, Australia). Just prior to SRTL-2 Bob delMas (University of Minnesota, USA) kindly agreed to also co-chair when Joan Garfield was prevented from coming to Australia due to poor health. Sponsors of the Forum were The IASE Statistical Education Research Group (IASE-SERG), The Centre for Cognition Research in Learning and Teaching (UNE), The School of Curriculum Studies (UNE), the Faculty of Education, Health and Professional Studies (UNE), and The Department of Educational Psychology (The University of Minnesota).

This was a unique opportunity for twenty researchers from six countries to meet for six days to share their work, discuss important issues, and initiate collaborative projects. For the first three days participants were divided into four small working groups, each with three or four presentations and one discussant. The research presented focused on the challenges in describing, teaching, and assessing statistical reasoning, thinking, and literacy, with the emphasis on reasoning. Most presentations included data presented on short videotape or audiotape. The next two days included sharing a synopsis of group findings with all participants and a discussion of the implications of the research presented. A panel session, with four invited speakers, entitled 'what we can learn from mathematics education research' on the final day was followed by an evaluation session and planning for SRTL-3. Following is a summary of the various sessions in the scientific program.

1.3. SRTL-3

SRTL-2 concluded with an evaluation session and discussion on future directions. Plans are now underway for the staging of SRTL-3. It is anticipated that this International Research Forum will be held in August 2003 in Lincoln, Nebraska. Dani Ben-Zvi and Joan Garfield are currently working on an edited book of research papers, some of which were presented at SRTL-2. For more information about SRTL please visit the website at http://www.beeri.org.il/srtl.

CHRIS READING
School of Curriculum Studies
University of New England
Armidale NSW 2351
Australia

2. REASONING ABOUT DATA AND DISTRIBUTION THROUGH THE STATISTICAL INVESTIGATIONS OF A THIRD GRADE CLASSROOM

RUTH HEATON and WILLIAM MICKELSON
University of Nebraska-Lincoln, Lincoln, Nebraska, U.S.A.
rheaton@unl.edu, wmickelson2@unl.edu

We contend that statistical reasoning occurs in the context of applying the process of statistical investigation to solve a specific problem or better understand a topic. In particular, this entails utilizing statistical ideas and information as well as presenting or summarizing data in meaningful ways so that inferences or interpretations from the data can be made that lead to greater understanding about the topic of the investigation. In this session, we gave a description of a third grade teacher's reasoning with data through graphs in the context of applied statistical investigations. The teacher attempted to thematically use the process of statistical investigation as a tool to teach other k-6 curriculum topics, such as language arts, social studies, health and science. We discussed what this reasoning looks like in the elementary classroom, how it is used, and why it is important to study.

The participating third grade teacher took part in a professional development workshop on the process of statistical investigation and the development of curriculum merging this perspective with elementary curriculum topics in Summer 2000. The participating teacher in this study was a best case scenario in that the teacher had considerable prior experience with statistics through a local implementation of Quantitative Literacy and the American Statistical Association's project competitions. The object of our investigation was the teacher's knowledge about data and distribution in the context of teaching during the school year following the summer workshop. An open-ended, hands off approach was used to observe teaching practice. All subsequent interventions regarding statistical content were made after a teaching episode in response to teacher questions. The statistical investigation units were planned in advance and video-taped by researchers in the elementary classroom.

Five artifacts on classroom teaching with supporting video were shared to illustrate the contextual features influencing statistical reasoning in the context of applied statistical investigations into topics across the third grade curriculum. We found our group's discussion and interaction with our data extremely valuable. Three main ideas were raised in the discussion of our data that point to the need to focus on pedagogical content knowledge in teaching statistical investigation.

The first idea focused on the relationship between questions posed, data collected and the purpose of the investigation. The pressing question was: "when should the teacher and children consider the investigation question in the process of reasoning with data?" One perspective presented is that one ought to consider the question in the planning phase of an investigation, then with the question in mind proceed to think through the data collection process. Another perspective argued that one should simply gather data without a pre-planned purpose in mind and then ask what is the data telling me now?

The second idea centred on the role of central tendency and variability, in making interpretations, inferences and predictions based on data. We repeatedly saw the teacher and children interpret data based on frequency of occurrence, in terms of which group had more and which had less. There was little/no evidence of the teacher or children using variability in their interpretations, inferences or prediction. Considering the interplay of central tendency, variability, interpretation and prediction, and the underlying story from the data raises interesting questions about the nature of learning opportunities offered and possibly missed. In regards to these first two ideas, it was pointed out during discussion that when teachers do not know or understand the content sufficiently well, they fall back to a comfortable position in terms of what they do know. In the case of teaching and implementing a statistical investigation in the classroom, we repeatedly saw this phenomenon where the nature of the posed investigation questions were of the type which group has more versus which has less, and the sole use counts as a means to analyse data.

In the process of reasoning with data, our teacher naturally attempts discussions on interpretations of data. When one looks closely at the dialogue, it is apparent that the teacher does the talking, asks the questions (usually with particular answers in mind), tries to lead students to specific answers, but often answers her own questions. This observation raises interesting questions about the interplay between this teacher's knowledge of statistics and her ability to be fluent in the different ways she might want to be, or need to be, responsive to students' learning needs. A related research question that was generated is: In what ways does the teacher's statistical knowledge influence his/her manner of leading class discussions with children? We left our session with new insights for further data analysis.

From the analysis we have done so far, we can say that it is extremely difficult to make a judgement about a teacher's statistical reasoning by looking in isolation at their understanding of statistical concepts like distribution, centre, and spread. In other words, knowing that a teacher knows how to pose a question, collect data, and summarize the data graphically does not necessarily predict what the teacher will be like in the classroom, teaching the process of statistical investigation in an applied context. The teacher in this study was a 'best case' scenario since she had previous statistics experience in a number of different contexts, and had demonstrated more than a sufficient command of the statistics in a de-contextualised situation removed from the immediate act of teaching (i.e. the summer workshop). In one sense she knows the content, however, we continually see the limits of what she knows or how she knows it when she tries to use the knowledge in teaching to help children learn the same ideas about statistical reasoning.

Our case highlights the need to understand the differences between statistical content knowledge and pedagogical content knowledge in statistics when conducting applied statistical investigations. What the data analysis session at SRTL-II did for us is begin to define in greater detail what pedagogical content knowledge in statistics is needed when working with K-6 teachers.

RUTH HEATON
University of Nebraska-Lincoln
118 Henzlik Hall, Lincoln, NE 68588-0355
USA

2. STATISTICAL INVESTIGATIONS: YEAR 7 AND 8 STUDENTS' REASONING WITH MULTIVARIATE DATA

MAXINE PFANNKUCH
University of Auckland, New Zealand
m.pfannkuch@auckland.ac.nz

The 4-dimensional framework for statistical thinking in empirical enquiry created by Wild and Pfannkuch (1999) provided the theoretical model for examining and describing students' reasoning with data. This model was built up through examining the statistics discipline itself and can be thought of as a way of describing the type of thinking that should be fostered in students. Some interviews with 11 and 12 year-olds (Rubick, 2000), while conducting their own statistical investigation with a small multivariate data-set created by Watson, Collis, Callingham and Moritz (1995), were presented as well as the students' written tables, graphs and conclusions.

The data were analysed through the four lens of transnumeration, consideration of variation, reasoning with statistical models and integrating the statistical with the contextual (Yoon, 2001). These four aspects were identified in the Wild and Pfannkuch framework as being fundamental types of statistical thinking. The analysis of the data attempted to describe the students' reasoning through these four lens. For example, data were provided to demonstrate students' noticing local and global variation, explaining local and global variation, controlling variation and quantifying variation. All these data showed students' emergent understandings of variation.

The following four aspects were highlighted as a result of group discussion about the data presented. First that students' views of data throughout a statistical investigation involve not only an intertwinement of local and global statistical thinking but also an intertwinement of local and global contextual thinking. Second that students created their own representations for displaying multivariate data which seemed to be fostering statistical thinking and seemed to be part of learning how to represent data. Third that developing an awareness of the need to converse with the data as well as to have different conversations with data in their various representations is part of the reasoning process throughout an investigation. These conversations build up an understanding of relationships in a data-set and enable students to learn more in the context sphere. Four that students need to play the dual role of corroborator and discoverer. The corroborator uses data to justify a claim whereas the discoverer is the explorer or data-detective or hypothesis generator who looks at data for possible interesting patterns, features, anomalies and so forth.

REFERENCES

Rubick, A. (2000). *The statistical thinking of twelve year 7 and year 8 students.* Unpublished Master's thesis. The University of Auckland, New Zealand.

Watson, J., Collis, K., Callingham, R. & Moritz, J. (1995). A model for assessing higher order thinking in statistics. *Educational Research and Evaluation*, 1, 247-275.

Wild, C. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry (with discussion*). International Statistical Review*, 67(3), 223-265.

Yoon, C. (2001). *An analysis of students' statistical thinking.* Unpublished Master's dissertation, The University of Auckland, New Zealand.

MAXINE PFANNKUCH
Department of Mathematics, The University of Auckland
Private Bag 92019, Auckland
New Zealand

4. JUNIOR HIGH SCHOOL STUDENTS' CONSTRUCTION OF GLOBAL VIEWS OF DATA AND DATA REPRESENTATIONS

DANI BEN-ZVI
University of Haifa, Israel
dbenzvi@univ.haifa.ac.il

Becoming competent in a complex subject matter domain, such as statistics, "*may be as much a matter of acquiring the habits and dispositions of interpretation and sense making as of acquiring any particular set of skills, strategies, or knowledge*" (Resnick, 1988, p. 58). This involves both cognitive development and 'socialization processes' into the culture and values of 'doing statistics' ('*enculturation*'). This study of Ben-Zvi & Arcavi (2001) is intended as a contribution to the understanding of these processes in the area of Exploratory Data Analysis (EDA). It focuses on the ways in which junior high school students begin to develop a global view of data and their representations, while investigating time series.

Data were collected in a progressive experimental school during a 10-week EDA course: 1) videotapes of one pair of seventh grade students capturing their learning during the first four class periods; and 2) responses of 80 students – working in pairs and using a spreadsheet – to a data analysis assessment task, which was administered after the end of the EDA course. The analysis of the videotapes was based on *interpretive microanalysis*: a qualitative detailed analysis of the protocols, taking into account verbal, gestural and symbolic actions within the situations in which they occurred. The goal of such an analysis is to infer and trace the development of cognitive structures and the socio-cultural processes of understanding and learning. The data of the assessment task were categorized and analysed taking into account the type of explanations associated with them.

This study shows how students' novice views slowly changed and evolved towards an expert perspective by making use of local information in different ways as stepping-stones towards the development of global points of view of data sets in different representations. At the beginning they persistently emphasized local views of data in tables and graphs. They were attentive to the prominence of 'local deviations', which kept them from dealing more freely with global views of data. Only later, the focus on certain point wise observations, the gradual adoption of the notion of *trend*, and the exercise of scaling, helped them to direct their attention to the shape of the graph as a whole, taking into account the variability in the data.

Students' learning involves: a) prior knowledge is engaged in multifaceted and sometimes unexpected ways – possibly hindering progress in some instances, but making the basis for construction of new knowledge in others; b) during the learning process, many questions either make little sense, or, alternatively, are interpreted and answered differently from the original intention; c) students' work is inevitably based on partial understandings, which grow and evolve towards more complete meanings; d) most of the learning takes place through dialogues between the students themselves and in conversations with the teacher; and e) sophistication in students' understanding of data develops within each point of view (local and global) and within the dynamic and flexible integration of those views.

This study confirms that even if students initially do not make more than partial sense of their data analysis tasks, through the support of appropriate teacher guidance, class discussions, peer work and interactions, and ongoing cycles of experiences with realistic problem situations, students slowly build meanings and develop experts' points of view on local-global approaches to data and data representations.

REFERENCES

Ben-Zvi, D. & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics, 45*, 35–65.

Resnick, L. (1988). Treating mathematics as an ill-structured discipline. In R. Charles & E. Silver (Eds.), *The teaching and assessing of mathematical problem* solving (pp. 32-60). Reston, VA: NCTM.

DANI BEN-ZVI
University of Haifa, Faculty of Education
Mt. Carmel, Haifa 31905
Israel.

5. FROM DATA VIA 'BUMP' TO DISTRIBUTION

ARTHUR BAKKER
Utrecht University, The Netherlands
A.Bakker@fi.uu.nl

Students tend to see data as individual values and find it hard to reason with data sets as a whole that has certain characteristics such as an average representing the group, a majority and outliers, or a constant shape. An end goal of our teaching experiments was therefore that students could reason with distributions on an intuitive level in relation to shape (hill, bump). For this classroom-based research in six seventh-grade classes, the so-called statistical minitools initially designed by Cobb, Gravemeijer, and colleagues of the Vanderbilt University (Cobb et al 1997) were used.

At the presentation, three episodes were discussed. The first video fragment showed how students found the mean visually in a value bar graph by a compensation strategy (figure 1). Student: *"I cut off the long bits and give them to the shorter ones"*. The presenter argued in the discussion that this representation is more suitable for developing understanding of the mean than the balance model.

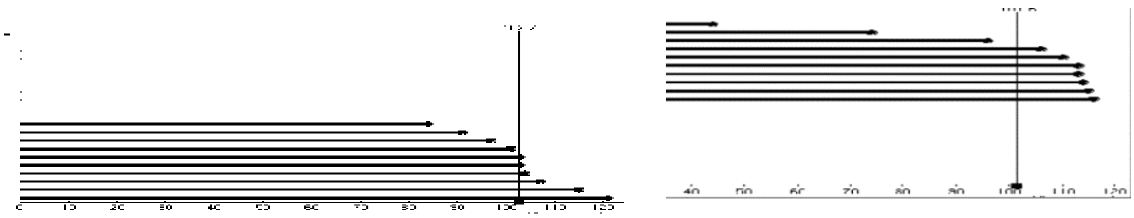

Figure 1. Compensation strategy on battery life span data in hours.

The second video fragment was an example of what the presenter calls 'extended sampling' or 'growing samples'. Students investigated sample size starting with four data points and adding new data to it up to 67 data points with dot plots. After that they predicted the shape of still larger samples. Their predictions were smoother and more hill-like than the investigated dot plots from smaller samples. In the last episode, students started to reason with 'bumps' after student-made graphs had been discussed. They predicted shapes and used terms like majority, outliers, sample size to explain their predictions.

Figure 2. Student graph of weight in kg leading to discussion of' bumps'.

REFERENCES

Cobb, P., Gravemeijer, K. P. E., Bowers, J. & McClain, K. (1997). *Statistical minitools*. Vanderbilt University, TN. Programmed and revised (2001) at the Freudenthal Institute, Utrecht University, Netherlands.

ARTHUR BAKKER
Freudenthal Institute
Postbus 9432, 3506 GK Utrech
Netherlands

6. 'VARIATION' FROM A STUDENT'S PERSPECTIVE

CHRIS READING
University of New England, Australia
creading@metz.une.edu.au

As statistics is the methodology needed to make decisions under uncertainty and variation is the source of this uncertainty, the understanding and analysis of variation is critical to students' reasoning in statistics. Although a number of global questions were posed relating to how students dealt with variation in different situations and some factors affecting their responses, the two questions deal with in most detail were: How is this variation described by students? and Is there an hierarchy of understanding of variation?

During the session a variety of student responses, which had been collected from a number of different sources, were used to help build up a Matrix of Description and Understanding of Variation (summarized in Table 1). This matrix, which gradually unfolded as the session progressed, is only in an embryonic form with no column headings as yet. The student responses, in italics, may help to elaborate on the titles of each cell in the matrix.

First, examples of the language used by students when asked to describe variation in general terms were examined and allocated to three cells forming the beginning of the matrix. Second, examples of any variation-related aspects of responses to open-ended general questions in various areas of statistics were presented and added to the matrix, creating some new cells. Third, variation-related aspects of student responses to a variation in sampling problem were slotted into the matrix. Fourth, and final, video excerpts of students working on the same sampling problem, but in a small group situation, were viewed and appropriate response segments added to the matrix.

Table 1. Summary of suggested Matrix of Description and Understanding of Variation

SUBSTITUTE ANOTHER WORD different changing variety	MOVE FROM ONE VALUE TO THE NEXT drop or go higher from high to low and back up again	IDENTIFY EXTRANEOUS SOURCE OF VARIATION I might get less than my friends might pick all over the place
CONCERN WITH MIDDLE VALUES size of the numbers doesn't matter just have to be different	CONCERN WITH EXTREMES minimum and maximum varies from 2- 7 stay out of the extremes	INDIRECTLY IDENTIFY SOURCE OF VARIATION not pick 7 – have got other colours in there
LOOK FOR A PATTERN no regular eating habits overall are gradually increasing	CONCERN WITH EXTREMES AND WHAT HAPPENS BETWEEN most …and least …and on average …	DIRECTLY IDENTIFY SOURCE OF VARIATION depends on what day it is for different ages
	DISCUSS CHANGE COMPARED TO SOMETHING come close to most of the non-extreme values	
	DISCUSS CHANGE COMPARED TO 'CENTRE' VALUE average around half…go over occasionally	

Discussion then centred on the various categories in the matrix and possibilities for rearrangements and headings for the various columns but no definite conclusions were reached. However, the student responses presented did demonstrate some aspects of students' describing variation and the matrix should help to provide a basis for developing an hierarchy of understanding of variation.

CHRIS READING
School of Curriculum Studies
University of New England
Armidale NSW 2351
Australia

7. ASPECTS OF STUDENTS' UNDERSTANDINGS OF VARIATION

MIKE SHAUGHNESSY
Portland State University, USA
mike@mth.pdx.edu

A repeated samples task and a probability task were presented to mathematics students in both written and interview format in order to study students' thinking and reasoning in contexts where repetitions of experiments lead to variation in the results of the trials. Over 300 middle level and secondary level students (predominantly ages 12 – 16) were given a survey to explore their thinking. Students predicted outcomes from a series of repetitions of the sampling activity and of the probability task. They were then asked to explain their reasoning, why they thought the results would occur as they predicted. Student responses to the written survey led to the development of an interview script for each of the two tasks in order to more explicitly address some of the issues that were surfacing in the written task. Audio taped interviews were conducted with 24 secondary school mathematics students.

Results indicated that a variety of types of thinking arise among students who are given such tasks to tap their conceptions of variation, including either too narrow or two wide an expectation for the range of results in the distribution of data from repeated samples. This session also shared evidence of students' beliefs about how they feel repeated experiments "should" come out, about their confidence or lack of confidence in predicting the results of sampling experiments. A necessary, but possibly not sufficient condition for students to make good predictions about repeated trials of a probability experiment is accurate knowledge about the sample space. Students may actually be able to list all possible outcomes prior to conducting experimental trials, or they might "learn" what the sample space is *while* conducting trials, but in either case knowledge of the sample space is critical for making reasonable predictions. Similarly, an understanding of what would be a *reasonable* spread around the population mean--some intuitive understanding of a confidence interval--is necessary for making good predictions for the results of repeated samples from a known population.

MICHAEL SHAUGHNESSY
Department of Mathematics
Portland State University
Portland, OR 97207
USA

8. COMPARING TWO DATA SETS: REASONING AND THE INFLUENCE OF COGNITIVE CONFLICT

JANE WATSON
University of Tasmania, Australia
Jane.Watson@utas.edu.au

This research looked specifically at school students' reasoning associated with comparing two data sets presented in graphical form. Of particular interest were comparisons where the data sets are of the same or different size. The reasoning was shown in the arguments presented by students to show that one or the other data set represents a class that had performed "better" on a test. As well many students were shown conflicting arguments of other students (on digital video) and asked to choose between the other student's argument and their own. Again arguments were documented and final decisions noted.

Individual interviews were conducted with 60 students from four Tasmanian government schools: 10 third- and sixth-grade students from each of two primary schools, and 10 ninth-grade students from each of two secondary schools. Students were selected on the basis that they would be willing to talk in interview and not be threatened by the complexity of listening to other students' ideas and evaluating them. Hence it might be expected that the students interviewed were more willing and able than would normally be expected for their grade levels. All interviews were video taped.

Following the data collection of Watson and Moritz (1999), a digitised video clip research resource was created using selected student responses from the 88 student interviews. The objective was to present new students being interviewed with conflicting ideas selected from interviews with the earlier 88, mimicking to some extent an ideal classroom setting where students engage in dialogue and debate. Eight prompts were included in the final protocol.

Among the research questions considered was the following: What is the evidence for visual strategies used in comparing the variation between data sets; is it based on individual attributes of the sets or on features across the entire sets; and do the strategies change after prompting? A clustering technique used on the responses produced five categories of increasing complexity of argument in using the visual evidence in the graphs, both when the size of the two sets was the same and when the size was different. Examples from the categories were given based on the video extracts of the students explaining their reasoning.

REFERENCES

Watson, J. M. & Moritz, J. B. (1999). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics*, 37, 145-168.

JANE WATSON
School of Education, University of Tasmania
GPO Box 252C-66, Hobart
Tasmania 7001
Australia

9. COMPARING TWO DISTRIBUTIONS: INVESTIGATING SECONDARY TEACHERS' STATISTICAL THINKING

JERE CONFREY and KATIE MAKAR
University of Texas at Austin, U. S. A.
jere@mail.utexas.edu, kmakar@mail.utexas.edu

While many schools are increasing their emphasis on statistics, few are taking the necessary steps to help teachers master the statistics they are expected to teach. Furthermore, U.S. teachers have little experience with data analysis and inferential statistics, yet in an era of accountability, are required to make instructional decisions based on large quantities of data about their students' performance.

Texas, where the study took place, has a high-stakes accountability system. Students are tested annually with the Texas Assessment of Academic Skills (TAAS), and high school graduation depends on passing this test. In addition, schools and teachers are held accountable for their students' performances. In this climate of high stakes accountability, urban schools that serve less academically advantaged children are under constant scrutiny to ensure they do not receive unacceptable ratings. As a result, much of the schools' professional development time is spent focusing on their TAAS results. Teachers feel that the accountability system creates a situation over which they feel very little power. This context seemed ripe to invite teachers to examine the statistical data as investigators.

A research team at The Systemic Research Collaborative for Education in Mathematics, Science, and Technology (SYRCE) in the College of Education at the University of Texas – Austin designed the NSF-funded research. Our focus group in the project was the mathematics department at our partner school, an urban middle school that feeds into a low-performing high school in the district. The professional development workshops followed an immersion model, allowing teachers to *do* statistics by investigating their own questions, and was conceived as a mathematical parallel of the Writers Workshop from the National Writing Project, where teachers learn to write rather than how to teach writing. The research project had a set of four related objectives to: (1) Strengthen *teacher content knowledge in statistics* by giving them the opportunity to learn statistics well beyond their curriculum; (2) Immerse teachers in *focused investigation and chains of reasoning* about student data in a high-stakes accountability environment; (3) Build *teacher confidence and facility in using dynamic software* (*Fathom*); (4) Orient teachers with a *healthy mindset about data and inquiry*: the acceptance of uncertainty when searching for solutions, the limitations and misuses of statistics and inferential reasoning.

The project took place in Spring and Summer 2001. During the spring, teachers learned the basics of the software *Fathom*™, unique in its application as a teaching and inquiry tool. Also during the early stages of our interactions with teachers, we examined introductory descriptive statistics and became acquainted with their student data. Throughout the latter phase of the project, a two-week intensive summer institute, teachers built a richer conceptual understanding of sampling distributions and inference through discussion, problem-based investigations with their student data, and simulations using *Fathom*. Sampling distributions were used frequently in problems to provide evidence for differences in groups and to imbue a tolerance for variation. As the study progressed, increasing amounts of time were dedicated to the teachers' own explorations. During the second week of the summer institute, teachers investigated a problem of their own choosing and presented their findings on the final day of the project to their peers and a group of researchers. At this time, clinical interviews were conducted to further probe teachers' reasoning about group comparisons. Our particular interest was in inquiry surrounding the question: How do you decide if two groups are different?

In the SRTL2 presentation, the videoed responses of four teachers in a clinical interview were examined. In the interview, which followed the project, teachers were asked to compare the relative performances of males and females on the state competency exam, given raw test scores for each group. Beyond the computational distinction made through descriptive statistics, teachers' analysis of comparing two groups was examined using several other important concepts: *tolerance for variability*, *understanding of the context*, and an ability to *draw conclusions*, perhaps inferentially. A categorization for statistical thinking about comparing two groups was described with five levels of reasoning that teachers use when comparing two groups.

Feedback during the presentation revealed commonalities with other areas of research in the focus group. Interestingly, the four members of the focus group (Jones, Moritz, Biehler, and Makar) represented research covering four different age levels: lower primary, upper primary and middle school, upper secondary, and professional; yet every age level struggled with similar concepts: understanding graphical representations and their connection to context, and conceptualizing variation.

JERE CONFREY
Department of Curriculum and Instruction
Sanchez Building The University of Texas at Austin
Austin, TX 78712
USA

10. CONFLICTING REPRESENTATIONS OF STATISTICAL ASSOCIATION

JONATHAN MORITZ
University of Tasmania, Australia
Jonathan.Moritz@utas.edu.au

10.1. BACKGROUND

Social- and physical-sciences often aim to reach verbal conclusions of causation by collecting bivariate data that involve statistical association and by controlling for other variables. It is important to be aware of the translation processes among raw numerical data, graphical representations, and verbal summaries, and an understanding of what constitutes a statistical association when presented in these forms (Fig. 1).

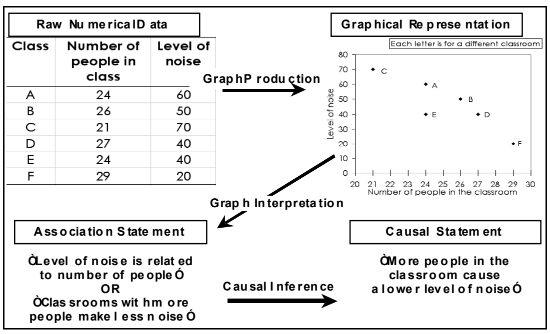

Figure 1. Forms of representing statistical association and skills of correlational reasoning to translate them.

10.2. TASKS

Students were provided with brief contexts of data collection involving at least two variables, and were given verbal statements of association to represent graphically:

Q1(a). "People grow taller as they get older".

Q1(b). "People do grow taller. But when you are 20 years old, you stop growing".

Q1(c). "For 10 year olds, girls and boys are about the same height. But men usually grow to be taller than women."

Q2. "People who studied for more time got lower scores."

Q3. "An almost perfect relationship between the increase in heart deaths and the increase in use of motor vehicles".

These tasks were intended be more informative of student understanding of association than asking students to judge whether a given graph represents an association. Q1 was based on Mevarech and Kramarsky (1997) who observed students' difficulties with linear relation of zero slopes.

10.3. SURVEY RESPONSES

Responses from previous research have been coded into 3 levels for Q1 (Moritz, 2000) and 4 levels for Q3 (Moritz & Watson, 2002). For this study, a total of 184 student surveys (grades 3, 5, 7, and 9) were gathered. Coding responses to Q2 was discussed with respect to causal reasoning about the topic context and beliefs about the direction of the association (see Figure 2).

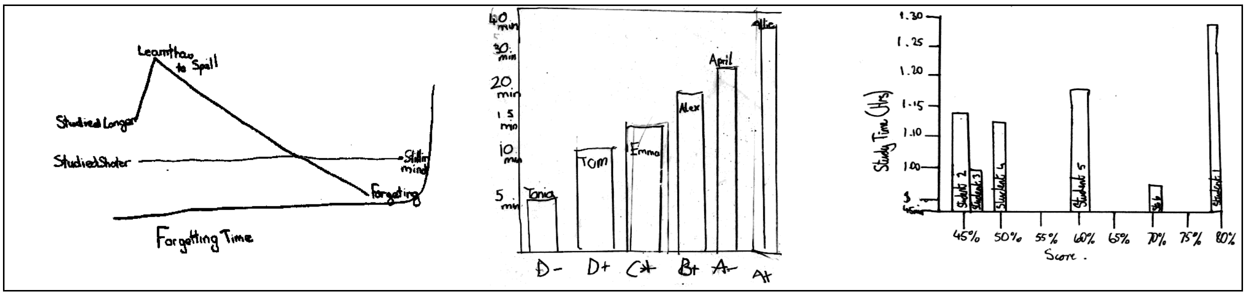

Figure 2. Student responses to Q2: (left) causal, grade 7; (middle) direction, grade 5; (right) direction, grade 7.

10.4. INTERVIEW DIALOGUE

In videotaped individual interviews, 34 students in grades 3, 5, 7, and 9 were first asked to explain their graphs, in particular how the graphs show the information and why they chose to represent the verbal statement the way they did. In an attempt to create cognitive conflict to explore how students might learn from new ideas, interviewees then were shown graphs drawn by other students, and asked to compare the different responses to decide which better represented the verbal statement. Selected extracts of dialogue illustrated how some students ignored the specifications of the survey task in order to represent what they believed about the topic context.

REFERENCES

Mevarech, Z. R. & Kramarsky, B. (1997). From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions. *Educational Studies in Mathematics*, 32, 229-263.

Moritz, J. B. (2000). Graphical representations of statistical associations by upper primary students. In J. Bana & A. Chapman (Eds.), *Mathematics education beyond 2000. Proceedings of the 23rd Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 440-447). Perth: MERGA.

Moritz, J. B. & Watson, J. M. (2002). *Representing and questioning statistical associations*. Manuscript in preparation.

JONATHAN MORITZ
University of Tasmania
GPO Box 252-66, Hobart 7001, Australia

11. DEVELOPING AND ASSESSING STUDENTS' REASONING IN COMPARING STATISTICAL DISTRIBUTIONS IN COMPUTER SUPPORTED STATISTICS COURSES

ROLF BIEHLER
Universität Gesamthochschule Kassel, Germany
biehler@mathematik.uni-kassel.de

The paper summarizes results from some of our studies of students' reasoning with data. We interviewed students after a computer supported course in statistics, which had an emphasis on exploratory data analysis (EDA). Our major goal was to support students' thinking in terms of "distributions". One of the issues we looked at was strategies and tools students used for comparing two data sets. Students had learned various displays and summaries including dot plots, box plots, histograms, mean, median, quartiles, interquartile range, variance and standard deviation.

We consider a cultural practice of using just means for group comparison as critical and often misleading. The origin of EDA is closely related to this criticism. Before one uses certain summary statistics for comparisons, distributional assumptions have to be checked, data displays have to be used for becoming aware of distributional behaviour. Box plots were introduced as an exploratory tool which provide a multifaceted initial distributional summary including a robust measure of the center and information about the amount of spread above and below the center. The difference :

$$q_3 - \tilde{x} = median\left(x_i - \tilde{x} \mid i \geq \frac{n+1}{2} \right)$$

where $\tilde{x}$ is the median of the whole data set can be interpreted as an average deviation from the median in the upper half (similarly the difference $\tilde{x} - q_1$). In this sense the box plot is intended a center $\pm$ spread display. In our research we identified many "non-standard" uses of box plots.

Students often frame group comparison tasks as hypothesis testing tasks such as: *Is X larger in group 1 than in group 2?* Example: Do boys (tend to) watch longer TV per week than girls? The expectation that this question has a definite answer is one of the obstacles that have to be overcome. Students are looking for a single comparison number, are irritated when quantiles in the box plot do not all point into the same direction. An interpretation of quartiles as medians of the lower (upper) half that could help is often not available. Students have difficulties in relating spread information to aspects of the context of the data. We think that the conscious introduction of the "uniform shift model"(group 2 distribution is just group 1 distribution uniformly shifted by a fixed amount) might help students. Looking for deviations from a shift model can draw attention to more complex distributional relation can occur.

Some students interpret the box as representing the "majority" of the data although it contains only 50%. The quantiles of the box plot are used for quantile by quantile comparisons but students do not understand why only this quantile selection. The quantile comparison of distributions was introduced by Galton and we consider this use as different from the use intended in EDA. Other students can see the varying data density in a box plot and can relate this to the different density representation in a histogram. It seems however to be difficult to see the box plot as a center$\pm$ spread display at the same time.

Generally, group comparisons need much more conceptual underpinnings than usual courses seem to offer (including our own). Different uses and interpretations of box plots have to be developed with adequate contexts. The presentation showed some tasks and data sets we used in order to achieve this goal. One of the formats was to ask the students to sketch the distribution in group 2 when a graph of the distribution in group 1 was given. This task opened an instructive window on students' thinking in terms of distributions and the interplay between representation and contextual knowledge.

ROLF BIEHLER
Universität Kassel FB17-Mathematik/Informatik
Hienrich-Plett-Str. 40, 34132 Kassel
Germany

12. STATISTICAL REASONING USED BY ELEMENTARY AND MIDDLE SCHOOL STUDENTS WHEN THEY ANALYZE AND INTERPRET DATA

GRAHAM A. JONES, CAROL A. THORNTON, CYNTHIA W. LANGRALL, EDWARD MOONEY(1), BOB PERRY(2) and IAN PUTT(3)
(1)Illinois State University, USA
jones@ilstu.edu, thornton@ilstu.edu, mooney@ilstu.edu, langrall@ilstu.edu
(2)University of Western Sydney, Australia
B.Perry@uws.edu.au
(3)James Cook University, Australia
IanPutt@bigpond.com

The session focused on elementary and middle school students' statistical reasoning when they faced tasks that engage them in analysis and interpretation. Analysis and interpretation incorporates recognizing patterns, trends, and exceptions in the data and making inferences and predictions from the data. It includes what Curcio (1987) refers to as *reading the data*, *reading between the data*, and *reading beyond the data*. Hence, when observing students analyzing and interpreting data, we were interested in the following processes: (a) how they extracted and described information explicitly stated in the data (reading the data), (b) how they compared and combined data (reading beyond the data), and (c) how they made predictions from the data (reading beyond the data). Consistent with these processes, we generated clusters of tasks like the following to assess children's statistical reasoning when they analyzed and interpreted data: (a) What does the picture tell you? (describe the data) (b) Which day had the lowest number of visitors? (compare) (c) How many friends came to visit during the week? (combine) (d) About how many friends would you expect to visit during the next 4-week month? (predict).

Having established some meaning for analysis and interpretation, we looked at videotapes of individual students (Grades 1 through 6) as they engaged in tasks involving analysis and interpretation. These structured interviews revealed four hierarchical levels of statistical reasoning: idiosyncratic, transitional, quantitative and analytical. Students who exhibit *idiosyncratic* reasoning consistently focus on ideas that are unrelated to the given data and frequently focus on their own personal data banks. Students characterized as *transitional* have begun to recognize the importance of quantitative thinking and generally provide relevant but limited responses to tasks. Students who exhibit *quantitative* reasoning can analyze and interpret data from more than one perspective; however they do not make connections between different aspects of the data. Consequently, they do not detect inconsistencies in their reasoning. Students characterized as *analytical* interpret data from different perspectives and are able to make connections between different aspects of the data.

We also examined Grade 2 students' analysis and interpretation of data during a teaching experiment. Our analysis revealed that these children were able to read between the data and beyond the data under certain conditions. Context plays a key role, and by providing opportunities for children to describe and investigate themes like a butterfly garden for an extended period, the teacher was able to build up a stronger contextual background for tasks involving analysis and interpretation. Children had difficulty focusing on subsets of data and this, in turn, affected their ability to make comparisons between two subsets of data. When looking at two subsets, we found that children focused on individual data values like the mode rather than examining the data subsets as a whole.

REFERENCE

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, 18, 382-393.

GRAHAM A. JONES
Mathematics Department
Illinois State University
Normal, IL 61790-4520
USA

13. METHODS FOR ASSESSING AND RESEARCHING STUDENT REASONING ABOUT SAMPLING DISTRIBUTIONS

MARK EARLEY
Bowling Green State University, USA
mearley@mco.edu

The main objective of this session was to present a discussion of how we as statistics education researchers could capture statistical reasoning. What does it look like? What should be assessed? How can we assess it? Why are we assessing it? All of these questions were addressed during the session. What follows is a basic outline of our discussion. The general outcome of the session, as I anticipated, was not any new specific knowledge, but rather a set of ideas that we as researchers should consider when investigating statistical reasoning in any context.

The presentation began with a brief review of literature summarizing what other researchers have done with sampling distributions. There is a lack of consistent measurement tools used to measure "understanding" or "reasoning" about sampling distributions – each researcher has traditionally used a course quiz or exam, idiosyncratic to the class and professor. There is an obvious need for statistics education research to pursue consistent, reliable, and valid ways to measure reasoning about sampling distributions. This led the group to two questions: (a) what are we calling statistical reasoning, and (b) what are we building towards? These lead us back to the question of "how do we assess statistical reasoning?"

One discussion thread identified one of our goals as professors is to build procedural and process knowing with our students … what can they do with the knowledge they have about sampling distributions? Looking at the concept of "sampling distribution," what is it we want them to know exactly? And if we look at the behavior or actions of our students for research data, how do we identify what knowledge is behind those behaviors and actions? Is this knowledge emerging with the task in which they are engaged? Or is this knowledge already in place in their minds, and they are simply accessing and using that knowledge?

Methods of assessing statistical reasoning that we discussed included memory and recall tasks (can you tell me what this is) and image making (e.g., concept mapping). There are a variety of methods for assessing this type of relational knowledge (Jonassen, Beissner & Yacci, 1993; Olson & Biolsi, 1991; Schau & Mattern, 1997). Does each of these methods give us reliable and valid data? Are any of them transferable to classroom assessment practices? These are questions statistics education researchers need to address before we can come up with consistent and comparable results across age levels and throughout the world.

By the end of the session, the group had not looked at specific examples of data as other groups and sessions did. What we had done instead was look one-step before collecting data to evaluate (a) what data should we collect? (b) How should we collect it?, and (c) what will this data tell us about what students know, how they know it, and how they came to know it? The discussion generated many more questions that we could have possibly answered, but I believe we have set up some interesting points to ponder as we move forward in our statistics education research efforts.

REFERENCES

Jonassen, D. H., Beissner, K. & Yacci, M. (1993). *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge.* Hillsdale, NJ: Erlbaum.

Olson, J. R. & Biolsi, K. J. (1991). Techniques for representing expert knowledge. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: prospects and limits* (pp. 240 - 285). Cambridge: Cambridge University Press.

Schau, C. & Mattern, N. (1997). Assessing students' connected understandings of statistical relationships. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 91-104). Amsterdam: IOS Press.

MARK A. EARLEY
2142 Fairfax Road
Toledo, OH 43613-5119
USA

14. A FRAMEWORK FOR THE DEVELOPMENT OF STUDENTS' STATISTICAL REASONING

ROBERT DELMAS, JOAN G. GARFIELD[1], and BETH L. CHANCE[2]
[1]*University of Minnesota, USA*
delma001@maroon.tc.umn.edu, jbg@maroon.tc.umn.edu
[2]*California Polytechnic State University, USA*
bchance@uop.edu

At the first SRTL, held in 1999 at Kibbutz Be'eri in Israel we summarized our research data and presented our initial framework for modeling students' levels of statistical reasoning. At SRTL-2 in Australia we focused on our current attempts to use videotaped interviews with students to modify the framework and better describe students' reasoning. The presentation started with a general description of our theoretical framework, which is motivated by the work of several researchers in cognitive development (Biggs & Collis, 1982; Case, 1985; Case, Okamoto & Griffin, 1996; Jones, Langrall & Thornton, 1997). This was followed by a description of our most recent research methods, which used students enrolled in several undergraduate introductory statistics courses within both the Statistics Department of Cal Poly and two departments of the University of Minnesota, as well as a graduate course at Minnesota. The procedure involved (1) administration of a pre-test on reasoning about the behavior of sampling distributions, (2) identification of a small sub-sample of students that appear to have different levels of reasoning, and (3) videotaping interviews with these students as they worked on problems designed to probe their reasoning about sampling distributions.

Our current analysis of the interviews attempts to identify aspects of students reasoning in order to more fully describe students thinking. Currently we have identified six categories of students' reasoning about sampling distributions:

- *Fluency*: Understanding and appropriate use of terms, concepts, and procedures
- *Rules:* Identification and use of a rule for prediction or explanation
- *Consistency:* Presence or absence of contradictory statements
- *Integration:* Extent to which ideas, concepts, and procedures are connected
- *Equilibrium:* Awareness of inconsistencies or contradictions
- *Confidence:* Degree of certainty in choices or statements

During our presentation we showed video clips that served as examples of each of these six aspects of students' reasoning about sampling distributions. Copies of the problems and interview protocol, as well as complete transcripts of the video clips were distributed to SRTL-2 participants. We gained valuable feedback on how to refine our methodology from the discussion and made new contacts that we hope will lead to collaborative research projects across institutions in the future.

REFERENCES

Biggs, J. B. & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.

Case, R. (1985*). Intellectual development: birth to adulthood*. Orlando: Academic Press.

Case, R., Okamoto, Y. & Griffin, S. (1996). The role of central conceptual structures in the development of children's thought (with discussion*). Monographs of the Society for Research in Child Development*, 61(2), 1-295.

Jones, G. A., Langrall, C. W &Thornton, C. A. (1997). A framework for assessing and nurturing young children's thinking in probability. *Educational Studies in Mathematics*, 32 (3), 101-125.

ROBERT DELMAS
109 Appleby Hall
University of Minnesota
128 Pleasant Street S.E.
Minneapolis, MN 55455
USA

15. CONCEPTUAL ISSUES IN UNDERSTANDING SAMPLING DISTRIBUTIONS AND MARGIN OF ERROR

PAT THOMPSON
Vanderbilt University, USA
pat.thompson@vanderbilt.edu

This study presented at SRTL-2 employed two teaching experiments to probe essential difficulties in students' constructions of schemes and imagery that might underlie their ability to reason powerfully about distributions of sample statistics. The study's methodology involved analyzing the idea in terms of conceptual operations it might entail, designing objects and situations with the intent of bringing those operations into play within conversations around them, and employing those objects and situations within the teaching experiments. Analysis of the teaching experiment data entailed using the conceptual analysis to guide initial explanations of students' successes and difficulties and feedback into the conceptual analysis in those instances where it failed to provide satisfactory explanations of critical events.

The first teaching experiment comprised 9 instructional days with 27 junior and senior high school students followed by 60-minute interviews of 10 students. Instruction in the teaching experiment focused on having students build multi-level images of repetitively sampling from a population and tracking the sample statistics to form distributions generated there from, and determining invariant properties of those distributions. Videos from Teaching Experiment 1 revealed several, possible essential, difficulties students encountered. The first was their disposition to participate in lessons unproductively. The conversations that actually took place sometimes confused students who anticipated that the instructor would demonstrate procedures that they would then commit to memory. The conversations most often focused on how to understand important issues and on reasonable ways to conceive of them so that people might develop reasonable ways to approach problems entailing them (such as, what does it mean that a particular event is "unusual" and how to determine whether it is). The second difficulty, grounded more in conceptual operations, was some students' predilection to conceive of samples as "some of" a population, instead of as a proportional mini-version of the population. The third difficulty, also grounded in conceptual operations, was some students' inability to keep in mind processes occurring at multiple levels, and their concomitant difficulty conceiving products of those completed processes (e.g., thinking of the distribution of sample statistics where each statistic comes from a sample collected randomly from a population).

Teaching Experiment 2, conducted with 8 juniors and seniors over 20 lessons, with all 8 students interviewed twice during the experiment and once afterward, focused specifically on having students develop the orientations and operations that were found to be problematic in Teaching Experiment 1. The teaching experiment was successful in addressing the problems of students not being able to reason and track the results of multi-level processes. But one additional essential difficulty became evident that probably was at play in Teaching Experiment I but which was confounded with the other difficulties. It was students' predilection to think of outcomes non-stochastically. Thus, repeating a process many times might produce varying outcomes, but students tended to think of outcomes per se as being unassociated with some repeatable process that might produce them. Video segments from the teaching experiments served as points of departure for discussions during the presentation.

PAT THOMPSON
Department of Teaching and Learning
Peabody College Box 330
Vanderbilt University
Nashville, TN 37203
USA

CONCEPTIONS OF VARIATION: A LITERATURE REVIEW[(1)]

MARIA MELETIOU
<*meletiu@spidernet.com.cy*>

SUMMARY

There are two parts to this literature review. The first part includes bibliography directly focusing on variation: meaning of variation, role of variation in statistical reasoning, research on conceptions of variation, as well as literature discussing the neglect of variation. The second part lists references belonging to four bodies of literature which, although not having the study of intuitions about variation as their main object of study, do offer rich insights into people's thinking about variation: literature on sampling and centers, on intuitions about the stochastic, on the role of technology, and on the effect of the formalist mathematics tradition on statistics education.

Keywords: *statistics education, variation, bibliography*

1. BIBLIOGRAPHY ON VARIATION

Pupils in the future will bring away from their schooling a structure of thought that whispers 'variation' matters (Moore, 1992, p.426).

1.1. MEANING OF VARIATION

The first consideration for someone interested in the study of intuitions about variation is the concept of variation itself. The following article by Wild and Pfannkuch identifies the issues necessary for the understanding of variation in data:

1. Wild, C. J. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223-265.

1.2. IMPORTANCE OF VARIATION

The papers listed below emphasize the central role that variation plays in statistical reasoning and that consequently it should also play in statistics instruction and research of students' understanding of statistics:

2. Azcárate, P. & Cardeñoso, J. M. (1994). Why ask why? *Research papers from the Fourth International Conference on Teaching Statistics.* Minneapolis: The International Study Group for Research on Learning Probability and Statistics.
3. Ballman, K. (1997). Greater emphasis on variation in an introductory statistics course. *Journal of Statistics Education, 5*(2).
4. Biehler, R. (1999). Discussion: Learning to think statistically and to cope with variation. *International Statistical Review, 67*(3), 259-262.
5. Hawkins, A. (1997). Discussion. *International Statistical Review, 65*(2), 141-146.
6. Hoerl, R., Hahn, G. & Doganaksoy, N. (1997). Discussion: Let's stop squandering our most strategic weapon. *International Statistical Review, 65*(2), 147-153.
7. Kettenring, J. (1997). Discussion. *International Statistical Review, 65*(2), 153.
8. Moore, D. (1990). Uncertainty. In L. Steen (Ed.), *On the shoulders of giants: new approaches to numeracy* (pp. 95-137). USA: National Academy Press.

(1) *Statistics Education Research Journal 1(1), 46-52, http:/fehps.une.edu.au/serj*
International Association for Statistical Education

9. Moore, D. (1991). Statistics for all: Why? what and how? In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics:*(pp. 423-428). Voorburg, Netherlands: International Statistical Institute.
10. Pfannkuch, M. (1997). Statistical thinking: one statistician's perspective. In J. Garfield & J. Truran (Eds.), *Research papers on stochastics education* (pp. 171-178). Minneapolis, MN: The International Study Group for Research in Learning Probability and Statistics.
11. Rubin, A., Bruce, B. & Tenney, Y. (1991). Learning about sampling: trouble at the core of statistics. Learning about sampling: Trouble at the core of statistics. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 314–319). Voorburg, Netherlands: International Statistical Institute.
12. Smith, T. M. F. (1999). Discussion. *International Statistical Review, 67*(3), 248-250.
13. Snee, R. D. (1999). Discussion: Development and use of statistical thinking: a new era. *International Statistical Review*, *67*(3), 255-258.

1.3. RESEARCH ON ROLE OF VARIATION IN STATISTICAL REASONING

Very few studies exist in the research literature, which focus directly upon students' conceptions of variation. I have located the following:

14. Meletiou, M. (2000). *Developing students' conceptions of variation: an untapped well in statistical reasoning*. Ph.D. Thesis, University of Texas at Austin.
15. Nicholson, J. (1999). Understanding the role of variation in correlation and regression. Presented at the *First International Research Forum on Statistical Reasoning, Thinking and Literacy*, Be'eri, Israel.
16. Reading, C. (1999). Variation in sampling. Presented at the *First International Research Forum on Statistical Reasoning, Thinking and Literacy,* Be'heri, Israel.
17. Reading C. (2001). *Variation from a student´s perspective*. Presented at the *Second International Research Forum on Statistical Reasoning, Thinking, and Literacy*, Armidale, Australia.
18. Reading, C. & Shaughnessy, J. M. (2000). Student perceptions of variation in a sampling situation. In T. Nakahar & M. Koyama. (Eds.), *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 89-96) Hiroshima, Japan.
19. Robinson, G. (2000). *Selling a course on experimentation*. Presented at the *OZCOTS-3 Statistical Education Workshop*, Swinburne University of Technology, Australia.
20. Shaughnessy, J. M. (2001). *Conflict between students´ personal theories and actual data: The spectre of variation*. Presented at the *Second International Research Forum on Statistical Reasoning, Thinking, and Liter*acy, Armidale, Australia.
21. Shaughnessy, J. M. (2002). Aspects of students´ understandings of variation. To be published in: *Proceedings of the Sixth International Conference on Teaching Statistics.* Cape Town, South Africa.
22. Shaughnessy, J. M., Watson, J., Moritz, J. & Reading, C. (1999, April). School mathematics students' acknowledgment of statistical variation. NCTM Research Presession Symposium: *There's More to Life than Centers*. Paper presented at the 77th Annual NCTM Conference, San Francisco, California.
23. Watson, J. (2002). Can grade 3 students learn about variation? To be published in: *Proceedings of the Sixth International Conference on Teaching Statistics.* Cape Town, South Africa.

The following articles, which I have submitted for publication and are available upon request, discuss the findings of a study conducted in a college level introductory statistics course that adopted a non-conventional approach to statistics instruction with variation as its central tenet:

24. Meletiou, M. (2002). Technological tools in the introductory statistics classroom: effects on student understanding of inferential statistics. *Submitted for publication in: International Journal of Computers for Mathematical Learning.*
25. Meletiou, M. & Lee, C. (2002). Redefining statistics instruction: moving away from the formalist tradition. Submitted for publication in: *Mathematical Monthly.*

26. Meletiou, M. & Lee C. (2002). Student understanding of histograms: a stumbling stone to the development of intuitions about variation. To be published in: *Proceedings of the Sixth International Conference on Teaching Statistics.* Cape Town, South Africa.

1.5. NEGLECT OF VARIATION

The references that follow acknowledge the almost complete absence of research on variation and provide possible explanations for the tendency of textbooks, instruction, and research to look at centers in data while ignoring variability: In addition to [1], [22]:

27. Batanero, C., Estepa, A. & Godino, J. D. (1997). Evolution of students' understanding of statistical association in a computer-based teaching environment. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 198-212). Voorburg, The Netherlands: International Statistical Institute.
28. Biehler, R. (1994). Probabilistic thinking, statistical reasoning, and the search for causes: Do we need a probabilistic revolution after we have taught data analysis? In J. B. Garfield (Ed.), *Research papers from the Fourth International Conference on Teaching Statistics*. Minneapolis: The International Study Group for Research on Learning Probability and Statistics.
29. Biehler, R. (1997). Students' difficulties in practicing computer-supported data analysis: Some hypothetical generalizations from results of two exploratory studies. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 176-197). Voorburg, The Netherlands: International Statistical Institute.
30. Loosen, F., Lioen, M. & Lacante, M. (1985). The standard deviation: some drawbacks of an intuitive approach. *Teaching Statistics, 7*, 29-39.
31. Shaughnessy, J. M. (1997). Missed opportunities on the teaching and learning of data and chance. In J. Garfield & J. M. Truran (Eds.), *Research papers on stochastics education* (pp. 129-145). Minneapolis: University of Minnesota.
32. Truran, J. M. (1994). Children's intuitive understanding of variance. *Research papers from the Fourth International Conference on Teaching Statistics.* Minneapolis: The International Study Group for Research on Learning Probability and Statistics.

2. OTHER RELATED BIBLIOGRAPHY

2.1. RESEARCH ON SAMPLING AND CENTERS

Although not directly studying people's understanding of variation, the existing research literature on students' understandings of samples offers useful information regarding people's thinking about variation. In addition to [11], [29]:

33. Bar-Hillel, M. (1982). Studies of representativeness. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 69-83). Cambridge: Cambridge University Press.
34. Garfield J. & delMas, R. (1991). Exploring the stability of students' conceptions of probability. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics*. Voorburg, Netherlands: International Statistical Institute.
35. Hawkins, A. (1997). Children's understanding of sampling in surveys. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 1-14). Voorburg, The Netherlands: International Statistical Institute.
36. Jacobs, V. (1997). Missed opportunities on the teaching and learning of data and chance. In J. Garfield & J. Truran (Eds.), *Research papers on stochastics education* (pp. 3-37). Minneapolis: University of Minnesota.
37. Konold, C., Pollatsek, A., Well, A. & Gagnon, A. (1997). Students analyzing data: research of critical barriers. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 159-175). Voorburg, The Netherlands: International Statistical Institute.
38. Lord, C. G., Ross, L. & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of

prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37* (11), 2098-2109.
39. Metz, K. E. (1999). Why sampling works or why it can't: Ideas of young children engaged in research of their own design. *Proceedings of the XXI Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Cuernavaca, México.
40. Mokros, J. & Russell, S. J. (1995). Children's concepts of average and representativeness. J*ournal for Research in Mathematics Education, 26* (1), 20-39.
41. Nisbett, R., Krantz, D., Jepson, C. & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90*(4), 339-363.
42. Pfannkuch, M. & Brown, C. M. (1996). Building on and challenging students' intuitions about probability: Can we improve undergraduate learning? *Journal of Statistics Education, 4*(1).
43. Shaughnessy, J. M. (1992). Research in probability and statistics: reflections and directions. In D. Grows (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: Macmillan.
44. Well, A. D., Pollatsek, A. & Boyce, S. (1990). Understanding the effects of sample size on the variability of the mean. *Organizational Behavior and Human Decision Processes*, *47,* 289-311.

2.2. LITERATURE ON INTUITIONS ABOUT THE STOCHASTIC

Heuristics Literature

The very large and influential body of research that examines the informal strategies or heuristics people use when making judgments under uncertainty and which has come to be known as the heuristics literature, is also extremely useful in giving insights regarding students' intuitions about variation. References such as the following point to people's deterministic mindset and their limited ability to cope with uncertainty and variation. In addition to [41], [43]:

45. Batanero, C., Godino, J. D., Vallecillos, A., Green, D. R. & Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology, 25*, 527- 547.
46. Cohen, L. J. (1979). On the psychology of prediction: whose is the fallacy? *Cognition, 7*, 385-407.
47. delMas, R. & Garfield, J. (1990). The use of multiple items to identify misconceptions in probabilistic reasoning. In J. B. Garfield (Ed.), *Research papers from the Third International Conference on Teaching Statistics.* University of Otago, Dunedin, New Zealand.
48. Garfield, J. B. & delMas, R. C. (1994). Students' informal and formal understanding of variation. In J. B. Garfield (Ed.), *Research papers from the Fourth International Conference on Teaching Statistics.* Minneapolis, MN: The International Study Group for Research on Learning Probability and Statistics.
49. DelMas, R. C., Garfield, J. & Chance, B. L. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. Journal of Statistics Education, 7*(3).* [On-line]. Available: http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm
50. Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgement under uncertainty: heuristics and biases.* Cambridge: Cambridge University Press.
51. Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237-251.
52. Kahneman, D. & Tversky, A. (1982). On the study of statistical intuitions. *Cognition, 11,* 123-141.
53. Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6,* 59-98.
54. Nisbett, R. E. & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice Hall.
55. Shaughnessy, J. M. (1977). Misconceptions of probability: An experiment with a small-group, activity-based, model building approach to introductory probability at the college level. *Educational Studies in Mathematics, 8,* 285-316.
56. Tversky, A. & Gilovich, T. (1989). The cold facts about the "hot hand" in basketball. *Chance, 2(1),* 16-21.
57. Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232.

58. Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185,* 1124-1131.
59. Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement. *Psychological Review, 90*(4), 293-313.

Criticisms of the Heuristics Literature – Intuitions as Dynamic

A growing number of researchers have lately become critical of the heuristics literature because of its emphasis on discovering fallibilities in peoples' reasoning. The work of these researchers who, although acknowledging that our intuitions often run counter to stochastic reasoning, view those intuitions as dynamic and subject to development, can also be very valuable to the researcher and educator interested in building student conceptions of variation. In addition to [31], [41]:

60. Borovcnik, M. (1991). A complementarity between intuitions and mathematics. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (*Vol.1,* pp. 363-369). Voorburg, The Netherlands: International Statistical Institute.
61. Borovnik, M. & Peard, R. (1996). Probability. In A. J. Bishop (Ed.), *International Handbook of Mathematics Education* (pp. 239-287).Dordrecht, the Netherlands: Kluwer.
62. Confrey, J. (1991). Learning to listen: A student's understanding of powers of ten. In E. von Glaserfeld (Ed.), *Radical constructivism in mathematics education* (pp. 111-136). Dordrecht, the Netherlands: Kluwer.
63. Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. Dordrecht, the Netherlands: Reidel.
64. Fischbein, E. (1987). *Intuition in science and mathematics.* Dordrecht, the Netherlands: Reidel.
65. Gigerenzer, G. (1996). On narrow norms and vague heuristics: A rebuttal to Kahneman and Tversky. *Psychological Review, 103* (3).
66. Jacobs, J. E. & Potenza, M. (1991). The use of judgement heuristics to make social and object decisions: a developmental perspective. *Child Development, 62,* 166-178.
67. Jones, G., Thornton, C., Langrall, C. & Mogill, A. T. (1997). Using students' probabilistic thinking to inform instruction. In J. Garfield & J. Truran (Eds.), *Research papers on Stochastics Education* (pp. 171-178).
68. Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education*, *3*(1).
69. Pratt, D. C. (1998). *The construction of meanings in and for a stochastic domain of abstraction.* Ph.D. Thesis, University of London.
70. Shaughnessy, J. M. (1997). Discussion: Empirical research on technology and teaching statistics. In J. Garfield & J. Truran (Eds.), *Research papers on stochastics education* (pp. 217-219).
71. Smith, J. P., diSessa, A.A. & Rochelle, J. (1993). Misconceptions reconceived - a constructivist analysis of knowledge in transition. *Journal of Learning Sciences, 3*(2), 115-163.
72. Wilensky, U. (1993). *Connected mathematics - building concrete relationships with mathematical knowledge*. Ph.D. Thesis, Massachusetts Institute of Technology.
73. Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. In R. Noss (Ed.), *Educational Studies in Mathematics. Special Issue on Computational Environments in Mathematics, 33*(2), 171-202.

2.3. THE ROLE OF TECHNOLOGY

The computer's ability to repeat experiments through simulations, provides a potential advantage that could be exploited to help develop students' skills in recognizing uncertainty and variation and distinguishing among the different types of variation. The extensive literature which describes how technology is employed or could be employed in the statistics classroom to support the development of didactic materials related to variation ideas such as randomness, sampling distribution, and central limit theorem, provides information about the role of computers and other technological tools in shaping student conceptions of variation. Examples include, in addition to [6], [[15], [43], [69]:

74. Behrens, J. T. (1997). Toward a theory and practice of using interactive graphics in statistical education. In J.

B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 111-121). Voorburg, the Netherlands: International Statistical Institute.

75. Ben-Zvi, D. & Friedlander, A. (1997). Statistical thinking in a technological environment. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 54-64). Voorburg, the Netherlands: International Statistical Institute.
76. Blumberg, C. J. (1997). Discussion: How technology is changing the teaching of statistics at the college level. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 279-283). Voorburg, the Netherlands: International Statistical Institute.
77. Burrill, G. (1997). Graphing calculators and statistical reasoning at the secondary level through the use of technology. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 15-28). Voorburg, the Netherlands: International Statistical Institute.
78. Burrill, G. (1997). Discussion: Technology, reaching teachers, and content. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 71-74). Voorburg, the Netherlands: International Statistical Institute.
79. delMas, R. C. (1997). A framework for the evaluation of software for teaching statistical concepts. In J. B. Garfield & G. Burrill (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics* (pp. 75-90). Voorburg, the Netherlands: International Statistical Institute.
80. Garfield, J. (1997). Preface. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. ix-xi). Voorburg, the Netherlands: International Statistical Institute.
81. Garfield, J., delMas, B. & Chance, B. L. (1999). *Tools for teaching and assessing statistical inference: simulation software* [On-line]: htpp://www.gen.umn.edu/faculty_staff/delmas/stat_tools/stat_tools_softare.htm
82. Glencross, M. J. & Binyavanga, K. W. (1997). The role of technology in statistics education: A view from a developing region. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 301-308). Voorburg, the Netherlands: International Statistical Institute.
83. Ghosh, J. K. (1997). Discussion. *International Statistical Review, 65*(2), 154-155.
84. Hawkins, A. (1997). Myth-conceptions. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. vii-viii). Voorburg, the Netherlands: International Statistical Institute.
85. Konold, C. (1995). Confessions of a coin flipper and would-be instructor. *The American Statistician*, *49*(2), 203-209.
86. Lipson, K. (1997). What do students gain from simulation exercises? An evaluation of activities designed to develop an understanding of the sampling distribution of a proportion. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 137-150). Voorburg, the Netherlands: International Statistical Institute.
87. Meletiou, M. & Lee, C. (2000). Technology Impact on undergraduate statistics instruction. *Proceedings of the 2nd Mediterranean Conference on Mathematics Education*. Nicosia, Cyprus.
88. Meletiou, M., Lee, C. & Fouladi, R. (2002). The role of technology on student understanding of inferential statistics. Submitted for publication in: *Journal of Computers in Mathematics and Science Teaching.*
89. Meletiou, M., Lee, C. M. & Myers, M. (1999). The role of technology in the introductory statistics classroom: reality and potential. *Proceedings of the International Conference on Mathematics/Science Education and Technology* (pp. 292-297). San Antonio, Texas.
90. Meletiou, M. & Martin, W. (2001). Collaborative learning and the use of technology: experiences gained from a statistics course. Abstract published in: *Abstracts of papers presented to the Joint AMS-MAA meeting*. Louisiana, New Orleans. [On-line]. Available: http://www.ams.org/amsmtgs/2025_abstracts/962-s1-1359.pdf.
91. McCloskey, M. (1997). QERCUS and STEPS: The experience of two CAL projects from Scottish universities. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 99-109). Voorburg, the Netherlands: International Statistical Institute.
92. Moore, D. (1997). New pedagogy and new content: the case of statistics. *International Statistical Review, 65*(2), 123-165.
93. Rossman, A. J. (1997). Using technology to promote learning by self-discovery. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 226-237). Voorburg, the

Netherlands: International Statistical Institute.

94. Scheaffer, R. L. (1997). Discussion. *International Statistical Review, 65*(2), 156-158.
95. Schuyten, G. & Dekeyser, H. (1997). Computer-based and computer-aided learning of applied statistics at the department of psychology and educational sciences. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 213-222). Voorburg, the Netherlands: International Statistical Institute.
96. Starkings, S. (1997). How technological introduction changes the teaching of statistics and probability at the college level. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 233-254). Voorburg, the Netherlands: International Statistical Institute.
97. Wilder, P. (1994). Students' understanding of computer-based simulations of random behavior. *Research papers from the Fourth International Conference on Teaching Statistics.* Minneapolis: The International Study Group for Research on Learning Probability and Statistics.
98. Wilder, C. (1994). Embracing the "wider view" of statistics. *The American Statistician, 48*, 163-17.
99. Wood, M. (1997). Computer packages as a substitute for statistical training? In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 267-278). Voorburg, the Netherlands: International Statistical Institute.

2.4. BELIEFS ABOUT THE NATURE OF MATHEMATICS: IMPACT ON STATISTICS

The references below discuss how the formalist mathematics culture with its over-emphasis on determinism and its "orientation towards exact numbers" (Biehler, 1997, p. 187) affects statistics education. They help explain why, in contrast to the varied and extremely rich models of central tendency found in the literature, sterile approaches to the notion of variability (Shaughnessy, 1997) dominate both the curriculum and the research literature. In addition to [28], [29], [31], [72], [83], [93]:

100. Vallecillos, A. & Holmes, P. (1994). Students' understanding of the logic of hypothesis testing. In J. B. Garfield (Ed.), *Research papers from the Fourth International Conference on Teaching Statistics.* Minneapolis: The International Study Group for Research on Learning Probability and Statistics.
101. Metz, K. E. (1997). Dimensions in the assessment of students' understanding and application of chance. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 223- 238). Amsterdam: IOS Press.
102. Steinbring, H. (1990). The use of chance-concept in everyday teaching - aspects of a socially constituted epistemology of mathematical knowledge. In J. B. Garfield (Ed.), *Research papers from the Third International Conference on Teaching Statistics.* University of Otago, Dunedin, New Zealand.

3. CONCLUDING REMARKS

The bibliography I have presented here is not by any means comprehensive. Since I have a special interest in research on variation, I would appreciate the sharing of references by other researchers and statistics educators.

MARIA MELETIOU
4 Halkidikis St., No. 302
1057 Lykavitos, Nicosia
Cyprus