Statistics Education Research Journal

Volume 1 Number 2 December 2002

International Association For Statistical Education
http://www.cbs.nl/isi/iase.htm

International Statistical Institute
http://www.cbs.nl/isi

Statistics Education Research Journal

The Statistics Education Research Journal is published jointly by the International Association for Statistical Education and the International Statistical Institute to:

- encourage research activity in statistics education;
- advance knowledge about students' attitudes, conceptions, and difficulties as regards stochastic knowledge;
- improving the teaching of statistics at all educational levels.

It encourages the submission of quality papers, including research reports, theoretical or methodological analyses, literature surveys, thematic bibliographies, summaries of research papers and dissertations. Contributions in English are recommended. Contributions in French and Spanish will also be accepted. All the papers will be refereed.

ISSN: 1570-1824

Submissions

Manuscripts should be sent to Flavia Jolliffe, by e-mail (frjolliffe@yahoo.co.uk), as an attached document, in rtf format. Two files are required. In one of them the author's name, affiliation and references to the author(s)' previous work should be removed to ensure anonymity in the reviewing process. Manuscripts should conform to the style specified at the Journal website: http://fehps.une.edu.au/serj

TABLE OF CONTENTS

Editorial 2

New Associate Editors 3

Maxine Pfannkuch and Amanda Rubick. An Exploration of Students' Statistical Thinking with Given Data 4

Maria Meletiou-Mavrotheris and Carl Lee. Teaching Students the Stochastic Nature of Statistical Concepts in an Introductory Statistics Course 22

Beth L. Chance and Joan B. Garfield. New Approaches to Gathering Data on Student Learning for Research in Statistics Education 38

Nigel Smeeton. Undergraduate Courses in Dental Statistics in Britain and Ireland 45

Peter Holmes. Some Basic References for the Teaching of Undergraduate Statistics 49

ICOTS-6 Report
1. Report from the International Programme Committee Executive 54
2. Report from the Local Organisers 56
3. The Honourable Trevor Manuel, Minister of Finance South Africa. Opening Address 57
4. Summary of Plenary Sessions 61
5. Summaries of Topic Sessions 62

Recent Publications 75

Recent Dissertations 81

Information on Past Conferences 84

Forthcoming Conferences 88

Statistics Education Research Journal Referees January-November, 2002 91

EDITORIAL

Welcome to the second issue of the *Statistics Education Research Journal (SERJ)*! We hope that you agree that it was worth waiting for.

If you have looked at the first issue you know that it was not so very different from the *Statistics Education Research Newsletter (SERN)*, partly because it included material which had already been prepared for *SERN.* In this issue, in line with the statement on the *SERJ* web page (http://fehps.une.edu.au/serj) that papers submitted to *SERJ* are peer-reviewed by referees selected by the editors, we are pleased to publish four papers which have been refereed both by an associate editor and two external referees. The paper by Beth Chance and Joan Garfield makes an important contribution to the methodology of obtaining data on how students develop an understanding of statistics. This is complemented by a paper by Maxine Pfannkuch and Amanda Rubick that is concerned with a study of students' statistical thinking, and a paper by Maria Meletiou and Carl Lee about a study of students' reasoning. The fourth paper, by Nigel Smeeton, is a little different and describes research into statistics courses given to dental students.

We are, however, retaining some of the features of *SERN* in *SERJ*, so this issue includes a useful bibliography of references for teaching undergraduate statistics that was prepared by Peter Holmes, and, with an emphasis on statistical education research, summaries of publications, and information about recent dissertations, and about past and forthcoming conferences. We feel that these sections are valuable to the statistics education research community.

We are pleased to report that we are receiving a steady flow of papers and that currently there are several papers either being refereed or revised. This has encouraged us to increase the size of the editorial board and we were delighted that, in spite of their many other commitments, Joan Garfield, Dick Scheaffer, David Green, and Iddo Gal all accepted our invitation to become associate editors. They are all well known in the statistical education world. In particular, Joan started the newsletter which has evolved over time to become *SERJ* and she must be one of the most cited statistical education researchers, Dick has served as a vice-president of IASE and has been very active in the American Statistical Association, David was one of the first to undertake a large-scale study into the probabilistic understanding of school children and the questions he devised have since been used by many others, and Iddo is heavily involved with the international Adult Literacy and Lifeskills survey. You can find more details about these new associate editors later in this issue. Details of all members of the editorial board are on the web page, including an update for Flavia Jolliffe who has retired from the University of Greenwich but continues to be active in many professional activities.

The sixth International Conference on the Teaching of Statistics, ICOTS6, was held in Cape Town, South Africa, in July, twenty years after the first ICOTS in Sheffield in the UK. The second to fifth conferences were held in Victoria (Canada), Dunedin (New Zealand), Marrakesh (Morocco), and Singapore respectively. The seventh ICOTS will be held in Brazil in 2006. It was good to meet so many people active in statistical education at ICOTS6, and we were especially pleased that the attendance at the open meeting of *SERJ* was well supported in spite of the early hour. Many useful suggestions were made at this meeting and we would like to thank those who offered their help. You will be hearing from us if we have not already approached you! There is a full report on ICOTS6 in this issue, including the opening address given by South Africa's Minister of Finance, the Honorable Trevor Manuel. We are extremely pleased that he has agreed that we may reproduce this.

SERJ is still under development and the editorial board has been very busy behind the scenes, for example discussing the refereeing process and the guidelines for authors. Here we should like to take the opportunity to mention that Carol Joyce Blumberg, the IASE vice-president with responsibility for IASE publications, keeps a watchful eye on the email exchanges between members of the editorial board and makes helpful interventions. We value her input. Just as we were about to go to press we received the news that the Executive Committee and Council of the ISI have approved our request that *SERJ* is to be a joint publication of the ISI and the IASE.

The IASE executive has decided that access to *SERJ* should continue to be free to all until the journal is well known. The IASE is paying the cost of hiring a student to help Associate editor Chris Reading, who looks after the Website, prepare the final version of each issue. We are very grateful to Chris for these contributions to the success of *SERJ.* We urge you to join IASE if you are not already a member. This will be an indication of your support for its activities and enable you to play a full part in these. Please also tell others about IASE. The web address is http://www.cbs.nl/isi/iase.htm

FLAVIA JOLLIFFE and CARMEN BATANERO

NEW ASSOCIATE EDITORS

Iddo Gal lectures at the Department of Human Services, University of Haifa, Israel, where he teaches courses on quantitative research methods and statistics, among other topics. His publications cover issues in the acquisition, teaching, and assessment of statistics skills and of numeracy skills, general aspects of adult learning and empowerment, and problems with readability of informational materials. He co-edited with Joan Garfield "The assessment challenge in statistics education" (IOS Press / International Statistical Institute, 1997), and later edited "Adult numeracy development: Theory, research, practice" (Hampton Press, 2000). Presently he leads the team developing the numeracy assessment component of the international Adult Literacy and Lifeskills survey, is a member of the International Statistics Literacy project of the IASE, and serves as a Senior Research Advisor, International Literacy Institute, University of Pennsylvania.

Joan Garfield is a Professor of Educational Psychology at the University of Minnesota in the USA. She heads a new graduate program in statistics education and teaches courses in statistics and statistics education. She was former editor of the Newsletter of the International Study Group for Learning Probability and Statistics, and former member of the editorial board of the International Statistical Review. She is a former Vice President of IASE. She currently serves as President of the AERA Special Interest Group for Educational Statisticians, and is Chair Elect of the ASA Section on Statistics Education. She is Associate Director for Research for the new formed Consortium for the Advancement of Undergraduate Statistics Education, and heads a NSF Project with Bob delMas and Beth Chance (Web Assessment Resources Tools for Improving Statistical Thinking). She is co-chair of the International Research Forum for Statistical Reasoning, Thinking and Literacy (SRTL).

David Green is Reader in Mathematical and Statistical Education at Loughborough University but devotes 80% of his time to the post of Associate Dean (Teaching) for the Faculty of Science, being responsible for teaching quality across seven departments. He has a BSc in Mathematics and MSc in Automatic Computation, both from Manchester University, a PGCE and MEd from London University, and a PhD in Probability Concepts from Loughborough University. His previous teaching posts have been in a College of Education, two secondary schools (11-18), and he has also worked as a research mathematician for a large electronics company. He has a total of about 100 publications - books, conference papers and journal papers. His main area of interest has been Probability Concepts, and more recently also the use of ICT particularly Dynamic Geometry. He counts among his collaborators Carmen Batanero and colleagues at Granada. His main current research activity is as project director for the £250,000 HELM project - **H**elping **E**ngineers **L**earn **M**athematics, for which Loughborough is the lead institution in a consortium of five English universities. David Green has been Editor (and more recently Assistant Editor) of the international journal *Teaching Statistics*. Until recently, he was a member of the Executive Committee of the Joint Mathematical Council of the United Kingdom, and a member of the Education Committee of the Royal Statistical Society.

Richard Scheaffer received his Ph.D. in statistics from Florida State University, whereupon he joined the faculty of the University of Florida and has remained on that faculty ever since. Now Professor Emeritus of Statistics, he was chairman of the Department for a period of 12 years. Research interests are in the areas of sampling and applied probability, especially with regard to applications of both to industrial processes. He has published numerous papers in the statistical literature and is co-author of four college-level textbooks covering aspects of sampling, probability, and mathematical statistics. In recent years, much of his effort has been directed toward statistics education throughout the school and college curriculum. He was one of the developers of the Quantitative Literacy Project in the United States that formed the basis of the data analysis emphasis in the mathematics curriculum standards recommended by the National Council of Teachers of Mathematics and directed the task force that developed the Advanced Placement Statistics Program, which he also served as its first Chief Faculty Consultant. He continues to work on educational projects at the elementary, secondary and college levels. Dr. Scheaffer is a Fellow and past president of the American Statistical Association, from whom he has received a Founder's Award.

AN EXPLORATION OF STUDENTS' STATISTICAL THINKING WITH GIVEN DATA[1]

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

AMANDA RUBICK
Westlake Girls High School
Auckland, New Zealand
amandarubick@hotmail.com

SUMMARY

This paper examines how two twelve-year-old students built up their recognition and understanding of relationships in a set of data. Using a small multivariate dataset created by Watson, Collis, Callingham and Moritz (1995), the students conducted an investigation of their choice in a pencil-and-paper environment. The students' thinking across the three representations of cards, tables and graphs is analysed from the perspectives of transnumeration, consideration of variation, reasoning with statistical models, and integrating the statistical with the contextual, which were identified as fundamental statistical thinking elements in empirical enquiry in the framework of Wild and Pfannkuch (1999). The ways of thinking within each element across the representations are identified. In the analysis, references are also made to the types of statistical thinking present in the other ten students in the study. From the analysis we identified five issues that should be considered for determining how students construct meanings from data. They are: prior contextual and statistical knowledge; thinking at a higher level than constructed representations; actively representing and construing; the intertwinement of local and global thinking; and the changing statistical thinking dialogue across the representations.

Keywords: Statistical thinking; Statistical investigation; Middle-school students; Constructing meanings from data.

1. INTRODUCTION

Applied statistics is part of an information gathering and learning process which is undertaken to inform decisions and actions. Industry, medicine, government agencies and many other knowledge-based endeavours are increasingly relying on data for decision making and thus statistics and statistical thinking are becoming an integral part of a societal way of thinking. A statistical investigation is conducted in order to learn more in the context sphere. This learning is much more than collecting information, it involves synthesising new ideas and information with existing ideas and information into an improved understanding. The question is then raised as to how students develop this type of learning.

A statistical investigation requires learners to take on a data-detective role. The learners' thinking activates a constant dialogue between the data and themselves. Of course this dialogue is an internal conversation that resides within the learners as an "understanding" is built up of the real situation. To build up this "understanding" of the context reality, the statistical models that have been constructed, and the learners' statistical knowledge are "interrogated" to provide information that will feed into their mental models (Wild & Pfannkuch, 1999). Furthermore, as the learner engages in a dialogue with the different representations of the real system the dialogue necessarily changes as new meanings and understanding are conveyed about the problem under consideration. The resultant new learning about the context does not reside within one representation but rather

[1] *Statistics Education Research Journal, 1(2), 4-21, http:/fehps.une.edu.au/serj*
International Association for Statistical Education

is built up through engagement with a number of representations.

The overall aim of this exploratory study was to examine how twelve Year 7 and 8 (11-13 year-old) students reasoned with data within each representation and how they built up their recognition and understanding of relationships in a given set of data over the course of an investigation. The main purpose, however, was to establish a baseline position in terms of the students' statistical thinking using a theoretical framework. Therefore the first interest lay with what the students could construct for themselves without having experienced any teaching in how to deal with and reason with multivariate data. The second interest lay with the framework itself and the degree to which its use could help in uncovering and illuminating the statistical thinking going on in the students.

A four-way analysis was performed which entailed examining the data from the perspective of four of the five fundamental elements of statistical thinking – recognition of the need for data, transnumeration, consideration of variation, reasoning with statistical models, and integrating the statistical and the contextual. These five fundamental thinking types are now elaborated upon.

Recognition of the need for data. The foundations of statistical enquiry rest on the assumption that many real situations cannot be judged without the gathering and analysis of properly collected data. Anecdotal evidence or one's own experience may be unreliable and misleading for judgements and decision-making. Therefore properly collected data are considered a prime requirement for reliable judgements about real situations.

Transnumeration. For this type of thinking the word transnumeration was coined which means "changing representations to engender understanding". Transnumeration occurs in three specific instances. If the real system and statistical system are thought of from a modelling perspective then transnumeration thinking occurs when (1) measures that "capture" qualities or characteristics of the real situation are found, (2) the data that have been collected are transformed from raw data into multiple graphical representations, statistical summaries, and so forth, in a search to obtain meaning from the data, and (3) the meaning from the data, the judgement, has to be communicated in a form that can be understood in terms of the real situation by others.

Consideration of variation. Adequate data collection and the making of sound judgements from data are required for an understanding of how variation arises and is transmitted through data, and the uncertainty caused by unexplained variation. It is a type of thinking that is started from noticing variation in a real situation, and then influences the strategies that are adopted through every stage of the investigative cycle. Applied statistics is about making predictions, seeking explanations, finding causes, and learning in the context sphere. Therefore patterns in the data are sought and characterised in an attempt to understand them in terms of the context of the real situation.

Reasoning with statistical models. The predominant statistical models are those developed for the analysis of data. Most people interpret the term "statistical models" as meaning, for example, regression models or time series models. Even much simpler tools, however, such as statistical graphs can also be thought of as statistical models since they are statistical ways of representing and thinking about reality. When statistical models are employed to reason with, the focus is more on aggregate-based rather than individual-based reasoning, although both types of reasoning are used. A dialogue is set up between the data and statistical models. The models may allow patterns in the data to be found, group propensities to be found, and variation to be seen about these patterns via the idea of distribution. The variety of models available enable data to be explored in multiple ways dependent upon the nature of the data. For example, graphs, centres, spreads, clusters, outliers, residuals, confidence intervals, p-values are read, interpreted and reasoned with in an attempt to find evidence on which to base a judgement.

Integrating the statistical and contextual. Although the above types of thinking are linked to contextual knowledge, the integration of statistical knowledge and contextual knowledge is an identifiable fundamental element of statistical thinking. The statistical model must capture elements of the real situation and thus the resultant data will carry their own literature base (Cobb & Moore, 1997), or more generally, their own body of context knowledge. Information about the real situation is contained in the statistical summaries and therefore a synthesis of statistical and contextual knowledge must operate to draw out what can be learned from the data about the context sphere.

Although the student interviews could be analysed from the perspective of each of these elements there was nothing in the given task that really encouraged "the need for data" so that element is not part of this particular analysis.

2. OTHER RESEARCH

2.1. REASONING FROM DATA

Lesh (2000) believed that teaching should aim to help students develop conceptual technologies by focussing on mathematics-as-communication for constructing and making sense of complex systems. For statistics this might mean students should be actively engaged in empirical enquiry from the formulation of questions from a real situation to the drawing of conclusions, and be focussing on multivariate data rather than univariate data (Shaughnessy, Watson, Moritz, & Reading, 1999).

A number of research studies from different theoretical perspectives (e.g., Watson, Collis, Callingham, & Moritz, 1995; Konold, Pollatsek, Well & Gagnon, 1997; Ben-Zvi & Friedlander, 1997) have analysed how students reason with and handle data, and recognise relationships within a multivariate data-set. Konold et al. (1997, p.165) concluded that the students they interviewed had not made *"the transition from thinking about and comparing properties of individual cases, or properties of collections of homogeneous cases, to thinking about and comparing group propensities*". Ben-Zvi and Friedlander (1997) and Watson et al. (1995) focussed on developing a hierarchy of statistical thinking that ranged from using selected pieces of information to a full synthesis of information.

During statistical empirical enquiry, when students are given a multivariate data-set, they must undergo a process of understanding the data, posing questions to investigate, determining appropriate analysis methods, interpreting the data and drawing conclusions. How students interconnect each phase of the enquiry cycle and how reasoning at one phase impinges on the next phase is not known. If representational fluency, which according to Lesh (2000, p. 80), "*is at the heart of what it means to "understand" most mathematical constructs*" then more attention should be paid to elaborating on how students reason with graphs and other models at every stage of empirical enquiry. Kaput (2000, p. 8) stated "*knowledge is co-constituted by the means through which it is represented and used – it does not exist independently of its representation*". Therefore following the chain of representations that are used by students in statistical empirical enquiry might build up a picture of how statistical knowledge is constructed.

2.2. UNDERSTANDING THE GIVEN DATA

Statistical questions arise in real situations and are usually prompted by noticing an incident or phenomenon and wondering "Why?". Before the actual statistical questions are formulated, however, there is a long phase of understanding the real situation system and all the phenomena that impinge on it. Even if the data are given, there should still be a phase of "noticing and wondering" (Shaughnessy, 1997) during the process of understanding the data before there is a transition to defining statistical questions. During the process of understanding given multivariate data sets it has been found by Watson et al. (1995) and Biehler (1997) that students make sense of such information by conjecturing linear causal chains acting on individual cases. It is unclear how or if students take a group propensity view at this stage. Biehler (1997) observed that questions were chosen with regard to the variables and he thinks this might act as a constraint on the types of questions posed. This raises issues about how the posing of questions is linked to the prior perceptions and understandings of raw data, and linked to constructions and interpretations of statistical summaries.

2.3. STUDENTS' BUILDING AND INTERPRETATION OF TABLES

The construction of tables from raw data involves sorting and classifying data into groups. Hancock, Kaput, and Goldsmith (1992) found that students lacked awareness of the need to represent implicit information in the data, such as gender from the given names, which they attributed to a lack of knowledge about data structures. Recognising, developing and implementing criteria for an effective classification procedure is not easy for students (Lehrer & Schauble, 2000).

Little research appears to have been conducted on students' construction and interpretation of statistical data tables. This raises questions about how students perceive tables. For instance Biehler (1997) observed that students frequently dealt with messy tabular data rather than thinking that they would see more structure in the data through graphical representations. It seems, however, that students find it a struggle to re-represent data into another grouping thereby perhaps demonstrating this tension between focussing on individual cases and focussing on variables that may describe group behaviour. Bright and Friel (1998, p. 67) conjectured that *"tables* *may play an important role as an intervening representation that can smooth the transition between representing raw and reduced data*". We think, however, that the representations that re-aggregate or reclassify the raw data, which may pertain to the future analysis or to the question posed, play a role in table representations and are another transition point.

2.4. STUDENTS' BUILDING AND INTERPRETATION OF GRAPHS

Students' construction of graphical representations from verbal descriptions has been probed by Moritz (2000) and Mevarech and Kramarsky (1997). Moritz (2000) studied how 9-12 year-olds represented bivariate and multivariate data given verbal descriptions. Although the forms of the graphs were not conventional there were distinguishable responses of single comparison, double comparison, and series comparison. Construction of graphs is considered to require a different set of cognitive skills to interpretation. When tertiary students were given a multivariate data table Chick (2000) was surprised at their limited use of basic statistical techniques and use of inappropriate techniques.

Much research attention has focussed on the interpretation of graphs. Students' ability to recognise a trend or to compare variables when analysing data is an important part of determining relationships among variables. The ability of teaching programmes to shift attention from individual cases to group propensities is described by Ainley, Nardi & Pratt (2000) and Cobb (1999) in their teaching experiments. According to Ainley et al. (2000) their active graphing method using computing technology enabled 8-9 year-old students to construct meanings for trend. Their method encouraged students to plot their data during collection. Students learning to perceive a trend in a scatterplot started from a pointwise interpretation, and then moved through a series of transitions to being able to look through individual points to identify a trend. Cobb's (1999) teaching approach is based around structuring and organising data. Activities were designed where the purpose was to compare two or more data sets to make a judgement. Cobb concluded that the students were analysing data from a mathematical point of view since the focus was on the rate of occurrence of some set of data values within a range of values which he believes is at the heart of a statistical perspective. Also the notion of distributions rather than collections of data-points emerged for the students, pointing to this shift from an individual case to a group propensity perspective. These teaching approaches clarified how students can learn to reason with statistical models within a structured situation.

Curcio (1987) identified three levels of graphicacy which relate to the kinds of questions graphs can be used to address. These levels are: reading the data, reading between the data and reading beyond the data. We believe, however, that if the purpose of drawing a graph is to glean information about the real situation problem, to interrogate the data, then more levels need to be added such as 'looking behind the data' (Shaughnessy, Garfield & Greer, 1996). Most research on the interpretation of graphs seems to have functioned at a statistical analysis level and has yet to include a full statistical enquiry perspective. As Biehler (1997) noted students' interpretation of their graphs seemed to reflect what they had learned in the elementary classroom and this did not include upgrading their statistical-causal modelling capability. When studies such as Ben-Zvi (2000) did not include set questions but rather an exploration of data then it is found that students interpret data through using comparison methods and their contextual knowledge about the situation to explain interesting phenomena in the data. Many studies have set the statistical tasks and questions. It would be interesting to find out more on what students can do and why within an exploratory data-based environment.

2.5. DRAWING CONCLUSIONS FROM DATA

Research has found a range of responses for drawing conclusions from data. At one end of the continuum students will: give their personal opinion without reference to the data (Hancock et al. 1992; Chick, 1999); derive their conclusion from a single case (Konold et al., 1997); or use their graph as an illustration rather than a means of analysing the data (Ben-Zvi & Friedlander, 1997). At the other end of the continuum are students who are involved in an "ongoing search for meaning and interpretation to achieve sensible results" (Ben-Zvi & Friedlander, 1997, p. 50). Ben-Zvi and Arcavi (1998), Lesh, Amit and Schorr (1997), and Cobb (1999) in their teaching experiment research describe how students can be pushed to high levels of discussion on statistical information when they have to defend their interpretation. Cobb (1999) used Toulmin's theoretical scheme to describe how students argued for their interpretation and conclusions. This raises the issue as to whether there are other theoretical schemes for drawing conclusions from statistical data.

The research literature available indicated, that while much has been learnt about the interpretation of graphs there is still little research concerning how students reason about data when engaged in empirical enquiry. Also little research has been conducted in statistics on how students think about and reason with variation (Shaughnessy, 1997; Torok & Watson, 2000), think with a contextual knowledge base or think transnumeratively. In our opinion there is much more work to be done on understanding students' reasoning with data during statistical investigation.

Therefore our analysis will be not only interpreting students' thinking from the perspective of the statistical thinking elements but also suggesting how the elements are manifested in the thinking of middle school students who have had no experience with tools for interpreting multivariate data. Furthermore, research has not been conducted on the thinking of students using a framework that is specific to statistical thinking in empirical enquiry. This study might provide some insights into the thinking pathways and thinking behind the artifacts that students produce.

3. METHOD

In schools statistical investigations occur that range from open investigations in which students pose their own question and collect data to closed investigations in which data are given with specific questions. The investigation given to the students was one in which the data were given and the students chose what to explore. Although the data-set could be considered unrealistic and too small a sample it is nevertheless a useful school activity to introduce students to EDA (exploratory data analysis) within the constraints of a pencil-and-paper environment. Students in New Zealand begin statistics in Year 1 and by Year 8 have experienced conducting investigations, displaying discrete and continuous data, calculating averages, and comparing data. The focus is entirely on univariate data.

The second author chose a group of students where she was able to obtain the permission from the school and parents to interview them. The twelve Year 7 and 8 students (11-13 year olds) were randomly selected from the mathematics-extension students in a New Zealand intermediate school. Based on her knowledge of the students they were put into pairs so that each pair was the same year level, same gender and was known to be able to work together. Using an investigation and protocol created by Watson et al. (1995), the students were given sixteen cards. Each card contained information about one person: the name, age, weight, eye colour, favourite activity and number of fast food meals eaten per week (Fig.1). The information on these cards was all that they were given. Such a data set can be easily understood by students as the contextual knowledge required is within their own life experiences and hence questions and relationships between variables can be generated.

Name: Simon Kahn
Age: 18

Favourite activity: TV
Eye colour: Brown
Weight (kg): 74
Fast food meals per week: 12

Figure 1: Example of a Data-card

The students had not worked with multivariate data sets before and were not familiar with scatterplots. The students were first required to read and understand the information on the cards before thinking about what they could investigate. A pencil-and-paper environment was provided with materials such as calculators and graph paper being available for the students. They were interviewed by the second author and audio-taped for approximately one hour while they were conducting an investigation of their choice. The students were asked and prompted to think aloud as well as explain their actions as they progressed through the investigation. At times the interviewer pushed the students to think more deeply about what they were saying. However, the investigator did not suggest techniques to use rather used phrases such as “Can you tell me what your graph has shown you?” or “Anything else?”.

Three people separately analysed the interviews qualitatively. Rubick (2000) analysed all the interviews, Yoon (2001) three interviews, and the first author four interviews. All the interviews had at least two independent analyses. Each interview incurred a four-way qualitative analysis. For example, interview data would be analysed for possible instances of variation-type thinking and this was recorded first as an annotation on the script, and after consideration, in a table format with the data and description of the thinking. Since it was not known what variation-type thinking to expect the data was used to suggest possible instances. Therefore the analysis and findings are explorative and tentative.

4. RESULTS

There was no uniformity of reasoning across all the students. Hence for this paper the dialogue of two Year 7 twelve year-old boys, Andy and Ron, was selected because it covered a good range of the thinking displayed across the set of students. The results are presented in an unconventional format since an excerpt from a transcript may convey all four ways of thinking and since the focus of the study is on the chain of reasoning that these students used to build up an understanding of the information contained in the data. In this particular statistical investigation three types of representations were used: the cards, tables of data, and graphs.

The dialogue summary given below was chosen on the basis of demonstrating the progression of the students' thinking to the next representation and of highlighting potentially interesting stages in their thinking. After the dialogue summary an analysis is conducted from the four perspectives of transnumeration, consideration of variation, reasoning with statistical models, and integrating the statistical with the contextual, which were identified by Wild and Pfannkuch (1999) as being fundamental types of statistical thinking. References will be made to the other five interviews in the analysis to highlight aspects that could be considered about the statistical thinking of students.

4.1. THE DIALOGUE SUMMARY

The numbers, for example, (1) or (5), in the dialogue summary act as a cross-reference for the dialogue analysis that follows this summary.

(1). Andy and Ron were introduced to the information on the cards. They immediately focussed on the number of fast foods eaten per week by each person. In particular Ron suggested they find out who ate the most. On finding out that Simon Kahn ate the most and being aghast that he ate 12, Andy remarked "*he is a bit of a fatso isn't he?*".

(2). Ron was then asked to choose a card (Fig. 1) and to say what information was on it. "*Okay his name is Simon Kahn and he is age of 18, he weighs 74 which is pretty big for an 18-year-old I guess. And his eye colour is brown … favourite activity is TV. He probably sits down and does nothing … that's why his weight is 74kgs. And he takes 12 fast food meals per week which is pretty incredible. And by eye colour being brown it seems as if he could be … but he doesn't look like he's a native of Australia or New Zealand*".

(3). On being asked what sort of questions they thought they could answer using these cards, Andy said: "*Quite a lot of statistics. Like you could compare favourite activity with each other … like how much people weigh in accordance with their ages, like that might have something to do with … (searches for card to support his argument) … David Jones he weighs 30kgs and he's only 8 whereas (searches for another card) … whereas Andrew Williams is 14, he weighs 60kgs. So that might be something to do with age and weight*".

(4). Later on Ron remarked about the Simon Kahn card "*74kgs. My Mum doesn't weigh that much*".

(5). Andy: '*We could do something with age and weight like compare them like two 17 [year-old] people at 56 and 65 [kgs]. There's sort of a difference … . Fast food meals per week and weight like compare that. If there is more fast food and if they weigh more there could be a connection there. Eye colour - probably won't be much to do because it doesn't compare with anything else*".

(6). The interviewer then asked what the connection was, to which Ron replied: "*the more fast foods you eat the more weight you get [put on]*".

(7). Then Andy thought: "*there might be something to do with the sex of … like David Jones is like a male and since he develops more muscle as he grows that might contribute … and like Rosemary is a girl so she might be less heavier … so that might be something*" .

(8). Ron immediately thought of a girl in their school who disproved this conjecture: *"Girls nowadays are not like girls in the olden days. Like I mean look at [name of girl]".*

(9).The interviewer intervened at this stage and stated that they have now mentioned four factors. She asked what the fourth factor would be called to which Andy responded with “gender” and “we could do two [factors] each may be … we could split it up”.

(10). Andy decided: “*first we will have to analyse the cards and write down on a piece of paper the things what everyone has. Like I could do the age groups … like from youngest to oldest … probably that would be the best idea and then we could transfer them on to graph paper … we could do two [graphs] on each [piece of] graph paper”.*

(11). Andy took half the cards and gave the other half to Ron. He decided that they would each record their card information on their separate tables. He recorded age, weight, and fast food in approximate columns (Fig. 2) whereas Ron abbreviated the card layout (Fig.3). Throughout the transfer of the card data to the tables they continued to compare two people at a time and suggest reasons for the weight difference. Ron decided to do fast food and age, because the difference in the measurements was not so great, and Andy decided to do gender and weight.

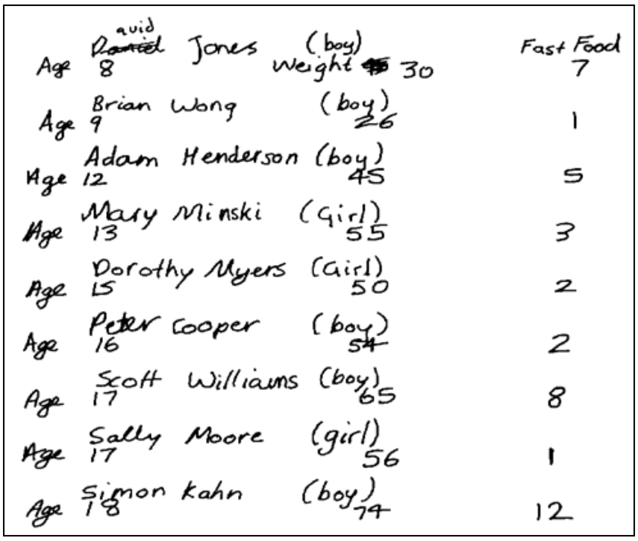

Figure 2: Andy’s table

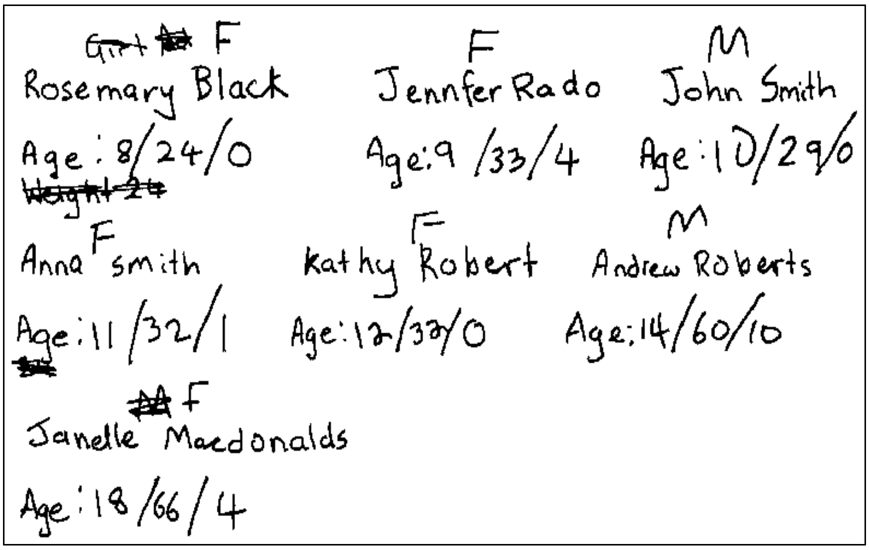

Figure 3: Ron’s table

(12). Andy then asked Ron whether they should do a bar graph. Ron quickly agreed.

(13). After Andy had written the initials along the x-axis he decided that he could actually do gender on the same graph because he could use a separate colour for each. Andy started to interpret his graph (Fig. 4) as he was constructing it: "*Okay, I've noticed that from 6 people I've already done they're actually about the same … between 33 and 26…24 [kgs]. So there's actually quite small [differences] between them and that's between the ages of 11 and 8, so it's quite even at the moment. … We would expect that".*

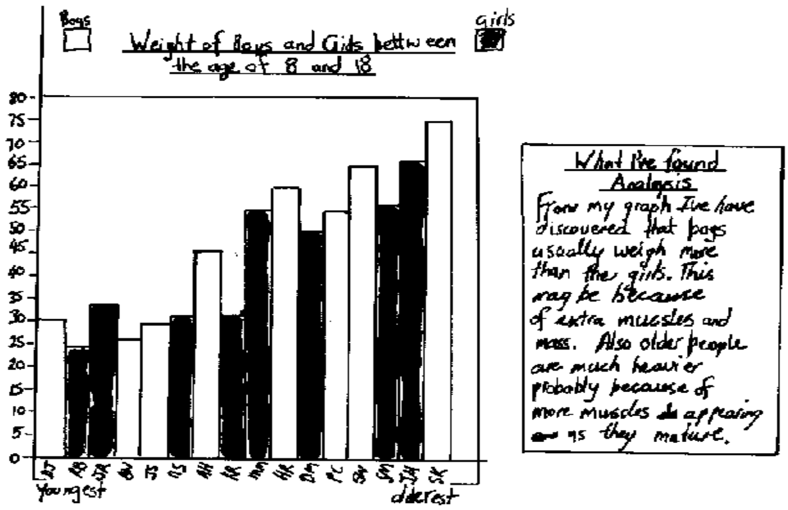

Figure 4: Andy's graph display for weight, age, and gender

(14).They continued graphing and discussing the people on the cards. Meanwhile Ron was constructing all the green bars for age first on his graph (Fig. 5). He noticed that the bars were climbing and he remarked: *"weeee there goes the beanstalk"* and if he "*attached a line to each end of this it would look like a mountain*".

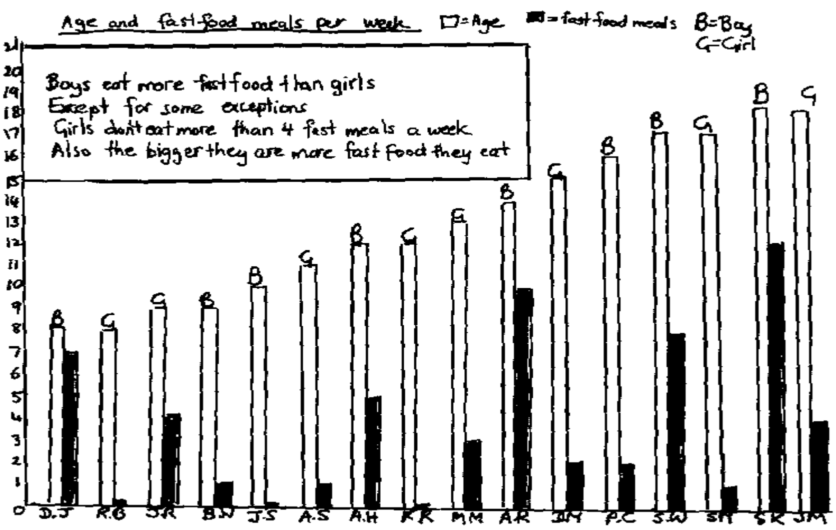

Figure 5: Ron's graph display for age, fast food eaten per week, and gender

(15). Andy announced he would write down what "*the graph tells me" like "how as the age gets more it's increased gradually…it [weight] keeps going up*".

(16). Ron objected to him mentioning age since he was doing that factor. Andy agreed "*but I did it in order of age so it sort of looks like that*". When Andy was asked if the weights followed a rule he replied: "*probably not, there's just a few exceptions where they go down … but that's because of like gender differences, but mainly they go up.*"

(17). Ron chimed in with: "*Obviously the boys weigh more than the girls*". The interviewer objected to Ron's statement since he had not looked at the graph. But Ron was confident of his ground since: "*Janelle MacDonald and Simon Kahn…they are the same age but obviously he, Simon, weighs more. Maybe it's because of diet (looks at list). Yeah diet - the fast food - he takes 12, Janelle takes 4*".

(18). Finally Andy read out his written conclusion: "*From my graph I have discovered that boys usually weigh more than girls. This may be because of extra muscles and mass. Also older people are much heavier probably because of more muscles appearing as they mature*"

(19). Ron in discussion with Andy started to interpret his graph, "*Hey, this is so easy to tell – boys eat way more than girls … fast food. If you look here* (points to graph) *one of the tallest* [bars] *is this* (points to first bar – David Jones with 7 fast foods) *he's a boy. Boy, boy, boy, boy and boy*" (points to the 5 tallest bars).

(20). After further discussion they noticed that the spread or range of fast food eaten was much smaller for girls than boys. Suddenly Andy noticed something about his graph: "*It appears that here (points near the middle of his graph) - as people get about 12 (age) they gradually get … they expand … they sort of weigh much more. Because it's quite even along here (points to the first 8-9 bars) until it gets to KR and MM. Then it goes right up and like keeps going above like another 20kgs (points to remainder of bars) and it keeps sort of staying the same. So it might be something there".*

(21). The interviewer asked for some possible reasons. "*Ah I don't know … it could be something to do with changing schools maybe. … intermediate to high school*".

(22). The interviewer reflected on the fact that they were about 12 years old and whether they would expect to have an increase in weight within the next few years. Andy agreed: *"Yeah we are 12. But … Actually it might be something to do with hormones too …"*

(23). Ron then finally interpreted his graph: "*Boys eat more fast foods than girls except for some exceptions. Girls don't eat more than 4 fast food meals a week. Also the bigger [older] they [boys] are the more fast food they eat".*

4.2. THE DIALOGUE ANALYSIS

Transnumeration

Transnumeration of the data could involve thinking about reclassifying the data, calculating averages for data grouped on a particular variable, or representing the raw or transformed data in tables or graphs. It is thinking about how to change the current representation to another representation.

Cards

Andy and Ron suggested that favourite activities played a role in weight (2) but they were unable to use this information since they did not think to transnumerate or reclassify the activities as sedentary and non-sedentary. Andy thought that gender might be a factor (7). This insight allowed transnumeration to occur since the names of people could now be reclassified as male or female (9). The reclassification of data into new variable groupings was not a prevalent feature of students' work. For instance Andy and Ron recognised age as a factor then sorted the data on age. Other students went further and classified data into age subgroups but only one pair of students recognised and dealt with age group as a variable. Before they moved on to recording the data Andy thought they could do a graph representation involving two factors (9). The identified factors determined the information to be recorded in the tables. The names, however, were still included, since these acted as an identifying marker to the original data which is equivalent to normal statistical practice. Andy transnumerated his data into a recognisable table format (Fig. 2) whereas Ron did not (Fig. 3). Similar practices were observed in the other students.

Table

From the cards to the table the dialogue should shift from one that is rich in context to one that is more reliant on statistical knowledge. There should be a focus on abstracting quantitative and qualitative variables from the context to produce a table. Most students were observed attempting this process. The tabulation of data is a skill that requires determining how to present the data with clarity and unambiguity. It means losing some of the information on the cards and using variable descriptors. This was a skill that was being developed by all the students although it appeared to be a more sophisticated skill than we had realised.

Ideally the dialogue with the table should include both noticing patterns in the table data, and determining graph construction. Both dialogues were observed but such a dual dialogue was not, perhaps indicating that this might be beyond the students at this stage. For most of the students the table was used as an organising tool for graphing (10). It was seen as a method of transferring data to the graph not as a method of also allowing a fuller inspection of the data for possible relationships and trends that might have been missed by looking at the cards. For some students the table was engaged with at the level of deciding scales on axes and/or what factors to graph together. One pair of students engaged with their tables and data fully to the extent that they did not see a need for a graph until they were prompted by the interviewer. Some students were aware that colour coding of factors would enable them to 'see' differences (13) whereas other students used colour for presentation purposes.

Graph

The dialogue with the graph could consider whether another representation might convey more information. This was not part of Andy and Ron's conversation. The graph-representation transnumeration for both Andy and Ron's graphs were in the form of a series comparison of case values (Figs. 4 & 5). Half the students represented multivariate data in series comparison graphs. For Andy and Ron it was only when the graph was being constructed that the idea emerged on how to delineate the two factors (13). Basically students seemed to be actively building on their current experience and knowledge. All students continued the dialogue with cards during this phase (11). It might be a necessary part of the conversation that needed to be continued for convincing each other that the investigation was on track. Even though some students constructed multiple representations, transnumeration thinking was not a feature of the dialogue with the graph(s) perhaps because of the pencil-and-paper environment that was provided and time constraints.

The communication of findings involves visually describing and/or writing a text to convey the information that has been gleaned from the data. Students demonstrated a range of methods from text only, to graph and text, and graph only. Andy and Ron were the best communicators in this respect.

Consideration of Variation

Cards

Initially there were two ways of noticing variation. The first way was noticing anomalies in the data which were beyond what they would normally expect (1). The second way was noticing differences between two people on one variable, say age, then finding another variable where there was a difference such as weight (3). They took datum as absolutes and did not seem to be aware of or did not argue about variation of weight within an age group, nor, when they compared different-aged people, did they argue from a perspective of variation in weights between age groups. An interviewer prompt caused Andy to think of another difference between two cases, that is gender (7). Ron thought of an anomaly outside the dataset (8) to demonstrate that gender and weight might not be related. Anomalies and noticing differences between two individual cases appeared to help shape many of the students' arguments about possible interconnections of variables.

All of the students were aware of what 'norms' to expect in data from their contextual knowledge to which they compared the card data (2, 4). The discussion was based around these contextual 'norms' but as yet the discussion did not include argumentation about variation about those 'norms' or statistical 'norms'. The type of argumentation was to find out whether the data on the cards was representative of the population or would confirm or refute what they already knew about the population (2).

Table

The variation patterns in the table did not appear to be noticed or engaged with, instead the variation dialogue continued with the cards.

Graph

When considering a graph representation the noticing of variation between two individual cases is no longer appropriate in the argumentation; it is now about noticing the variation between groups. There also has to be a distinction made between possibly real (may be explained) and random (cannot be explained) variation. Andy was able to make this distinction in when he 'saw' the jump between the 12 and 13 year olds (possibly real variation) (20) and 'saw through' the random variation for those under 12 (13). Both Andy and Ron 'saw through' the random variation with comments such as "just a few exceptions" (16, 23) but mainly the trend was increasing. Ron reverted to using two individual cases, however, to argue for a trend (17). Differences or variations among the heights of bars were noticed (19) which seemed to help Andy and Ron in picking up the pattern between the girls and boys. These comparisons suggested that they were also *informally quantifying variation* within a group since they noticed that the range for the amount of fast food eaten by the boys was larger than the range for the girls (20). Only some of the students could be said to be 'quantifying' or 'seeing through' the variation.

Noticing variation and arguing for a particular interpretation of the graphs was demonstrated on one or more of these levels by the students through: (a) using the cards or graph and comparing two individual cases, (b) checking a preconceived relationship which determined what variation to notice between two groups of data, and (c) the variation noticed in the graph between two groups of data determining a new possible relationship.

Reasoning with Statistical Models

Cards

Reasoning with raw data means making sense of the information and thinking of possible relationships that could exist. A feature of Andy's thinking was the use of the words 'compare' and 'connection' (3, 5). He seemed to realise that to reason statistically two variables were compared to see whether there was a connection between them. Ron used the word 'compare' on only two occasions – it was not a feature of his thinking – rather he was in a transition between thinking of descriptive questions such as how many people eat 0, 1, 2, etc. fast food meals and thinking of comparison questions such as "fast food and weight". But he did have a contextual group propensity perspective (6). On the one hand, Andy and Ron's argumentation followed the pattern that if something general about the population was proposed then they would confirm or refute it based on one specific example from their experience or from a card (4, 8). On the other hand, if they knew something general about the population, then they would give a specific example from the cards in order for the other person to understand (3). Only Andy and Ron demonstrated an argumentation that determined what questions were feasible or not feasible with the data.

Table

Andy and Ron's dialogue with the table involved the type of graph to be drawn and determining what was possible to graph based on their statistical knowledge (11, 12). Their statistical knowledge actually determined the factors to be compared. It was noteworthy that Andy did not know beforehand how he would display the factor gender. It was only in the action of graphing that he understood how he could communicate that factor in his representation (13). The placing of the names along the x-axis could be due to (a) not wanting to lose that information, (b) not knowing about scatterplots, (c) knowing that the x-scale should quantitatively increase and that it was not possible to have 8, 8, 9, 9, etc. along the axis. A similar feature was present in all the students' work apart from one pair of students who, for example, grouped the data into age groups and plotted average weights. Whatever the explanation might be, the decision to draw such a graph, was based on their current statistical knowledge.

Graph

When students reason with a graph the dialogue should change to focus on 'seeing' patterns in the whole of the dataset, and to talking about central tendencies, the trend (15), the way the data are distributed, and the spread of the data. The discussion should no longer be centred on reading individual cases, which could still be done, but on the whole group. Andy started reasoning with his model during construction by noticing the horizontal trend of the first six bars (13). He did not, however, allude to this trend in his conclusion. It was not until much later when he continued to interact with his graph that he articulated this possible pattern (20). Perhaps the action of graphing helped him to 'see' this particular pattern or perhaps he knew that he should be 'noticing' trends. This has highlighted the need for students to engage with the graph in two ways. The first way should be, based on their question, deliberately looking for a relationship to 'see' if there was a pattern (18). The second way should be to notice (20) a pattern in the data, and see what relationship if any, the pattern might convey.

This type of reasoning was not prevalent in all the students. Some continued to read values from the graph.

For example, Ron reverted to using the cards to argue that boys weigh more than the girls which was based on noticing differences between two people (17). This was a typical dialogue for the other students when interpreting their graphs and for some this was their conclusion. Ron, however, seemed to be in a transition since he seemed to be at the beginning of noticing a trend when constructing his graph (14). Then, with the help of Andy, he wrote a conclusion that communicated two patterns or trends and an upper limit (23). When Ron was reasoning about a trend he did so by comparing the heights of the bars (19) and since more male bars were higher he was arguing that "boys eat more fast food than girls' (23). This aspect was observed in some of the other students. One pair of students actually compared two bars at a time. Therefore a value comparison among bars was the method of argumentation for a trend.

Three pairs of students reasoned about the models they had produced. Their decisions included determining the model was valid, not valid or could not be validated with such a small sample. They then accepted the model, or rejected the model on the grounds that it was incorrect, or decided they could not draw valid conclusions from the model about the data (Rubick, 2000; Yoon, 2001).

The students' argumentation with a bar graph or other statistical model showed they were possibly operating on one or several of these four levels: (a) comparing two individual cases to argue that one individual variable value was higher than the other; (b) comparing two or several groups of data, individual case or averages, to argue that one group had higher values on the whole than the other; (c) 'seeing' and comparing the slope trends of one group with another group to argue that one group was different from the other; and (d) questioning, challenging or dismissing the model itself.

Even though their bar graphs were not easy to reason with for comparison of variables, a third of the students demonstrated a remarkable facility to reason with such 'poor' tools or models. By concentrating on the information contained in the data, these students formed quite sophisticated interpretations. The other students did not tend to do this, but when the interviewer asked two students to look at their three bar graphs and to explain to her if there was anything that was meaningful, one of the students was able to detect a non-preconceived pattern. This suggested that students could 'see' patterns if their attention was drawn to notice what graphs were telling them rather than looking for preconceived patterns or using the graph as an illustration.

Integrating the statistical and the contextual

Cards

Story telling about the observations seemed to be a precursor for thinking about variable relationships, for making sense of and rationalising the information on the cards. Causal thinking was prominent and involved the students using their imagination as well as being sceptical, whereas their statistical knowledge appeared to have little influence at this stage. Contextual knowledge influenced the positing of alternative explanations for causes-and-effects and what variables would be considered or not considered. Causes outside the data gathered, however, tended not to be considered; rather the thinking was mostly within the confines of the given data. From the very beginning, in the case of Andy and Ron, a causal relation was posited between fast food and weight (1). This causality was brought up a number of times (5) yet it was never challenged. Causal relationships between eye colour and name, and gender and weight, were challenged (5, 8). Their own contextual knowledge about population "norms" was drawn upon often to support their statements about whether the datum was an extreme value (2, 4) or to support cause and effect arguments such as weight and favourite activity (2, 3, 5, 6, 7).

Most of the students focussed on the weight variable. This required them to think what they knew about weight and the other variables from their general knowledge. From their dialogue with the cards Andy and Ron suggested the following about weight for the age group 8 to 18 years: as you got older you weighed more (3); eating fast food made you weigh more (6); sedentary activities made you weigh more (2); boys were heavier than girls (7); eye colour did not affect weight (5). They did not evaluate which explanations were more likely but did know that several variables influenced weight.

Table and graph

From the table to the graph the dialogue should ideally shift from one that is focussed on statistical knowledge to one of statistical and contextual interpretation. Hence the integration of statistical and contextual ideas, and conceptions should shape the main dialogue. This dual dialogue was present in the students' thinking in varying degrees. Causal inference was intertwined with the information divulged by the graph and with the contextual knowledge of the students. When Andy noticed the first six bars of his graph were 'even', he stated that this was to be expected (13). This suggested that he was using some contextual knowledge to reassure himself that the pattern was not unusual. When he was asked for possible reasons for the pattern (21) he thought of differences between younger and older people but did not evaluate whether they were 'sensible' reasons. Such a response was typical for the other students. The interviewer then asked him to think whether he would have such a weight gain within the next few years. This prompt produced another reason (22). The use of contextual knowledge to explain the variation or patterns is considered vital for determining whether the patterns could be 'real' and for the interpretation of the patterns for the real situation. Andy justified his written conclusion (18) by suggesting reasons for the trends he had noticed whereas Ron did not (23). Nor in their conversation were any reasons posited for the relationships discovered by Ron. Andy thought of explanations outside the realm of the data given (21, 22) for his graph. Ron tended not to think of explanations outside the given data which was typical of the responses from the other students. Andy, however, suggested that the factors were probable (18).

Even though the students knew there could be several possible explanations for a relationship between variables they displayed a range of interpretations: (a) did not think of an explanation, (b) gave one definitive explanation, (c) gave one probable explanation, (d) thought of several explanations within the confines of the given data, and much less evident (e) thought of alternative explanations outside the realm of the data. Evaluating which explanations would be more likely, likely or not possible was not a feature in the dialogue with the graph. Causal inference seemed to be a natural way of thinking for the students.

5. DISCUSSION AND CONCLUSIONS

From the analysis of the students' dialogue, we have identified five issues within the four statistical thinking elements that should be considered when determining how students construct meanings from data. They are: prior contextual and statistical knowledge; thinking at a higher level than constructed representations; actively representing and construing; the intertwinement of local and global thinking; and the changing statistical thinking dialogue across the representations.

5.1. PRIOR CONTEXTUAL AND STATISTICAL KNOWLEDGE

Firstly, lack of prior statistical knowledge was not an insurmountable obstacle for these students when dealing with multivariate data. Prior contextual knowledge was essential for conjecturing possible relationships within the set of data. Both contextual and statistical knowledge influenced students' understanding and interpretation of the data. Students were building on their own knowledge or folklore of relationships among variables to incorporate new learning about the dataset. Their interaction with the raw data suggested they reasoned from the population to the sample to help in the identification of factors to consider; that is, they used their existing contextual knowledge or folklore about expected norms, patterns or group propensities in the population. The strength of these preconceptions about cause-and-effect affected the way the data were focussed upon and the variables that were compared, and infiltrated the conclusions drawn. It would appear that such preconceived ideas should not be underestimated since they have a strong effect on the meaning that is obtained from data (Watson et al., 1995). This prior contextual knowledge was essential, however, for making sense of the data before posing a question and again when interpreting patterns in the variation (Cobb & Moore, 1997).

Statistical knowledge influenced the types of investigations undertaken, how the data were transnumerated, and how the representations were engaged with. For some students the first barrier was an inability to pose a question with the given data, for others it was an inability to engage with a representation other than at a reading level. Therefore, on the one hand, statistical knowledge was a limiting factor in their ability to construct meaning from the data, but on the other hand, having only a very basic knowledge did not prevent some students from gleaning quite sophisticated insights into the meaning of the data.

5.2. THINKING AT A HIGHER LEVEL THAN CONSTRUCTED REPRESENTATIONS

Secondly, the tables and graphs produced by the students did not indicate the level and depth of their thinking. Students are able to think at a higher level than their constructed representations might suggest. Most representations that the students constructed were not the conventional statistical representations (cf. Moritz, 2000); rather they were their own creation or way of dealing with data in table and graph form. Most students' graphs, however, retained identification of the individual case, which has been highlighted by other research (Friel & Bright, 1996). Their graphical inventions for dealing with bivariate or multivariate data were based on the individual-case bar graph. They spent a lot of time thinking, conjecturing, and learning about the data before and during the construction of representations which, according to Tukey (1977), is a necessary prerequisite for data interpretation. That is, they seemed to be concentrating on the information contained in the data or in the variation. Perhaps this concentration, coupled with understanding their own invented graphs, enabled them to recognise the relationships within the dataset. They had the capacity to think at a considerably higher level than their tools would indicate. At least half of these students had this capacity (cf. Chick & Watson, 2001).

5.3. ACTIVELY REPRESENTING AND CONSTRUING

Thirdly, the physical actions of tabulating and graphing together with the activation of a constant dialogue between the data and themselves helped the students to gain information from the data. Their continuous imaginings and their problems with representations seemed to place them into an enquiry mode of thinking. Thus, the strong feature of all the students' actions and dialogue was that they were actively representing and construing and that interpretation was an ongoing process. As soon as they picked up a card their imaginations were prompted as they interpreted what they were seeing or experiencing. For the two students focussed upon in this paper the table and the graph were not fully visualised before proceeding. Their method was to start their representation and then to solve problems of representation during construction. At the same time as they were actively constructing their graphs they were actively interpreting any perceived patterns or anomalies through focussing on the height of the bars. The physical action of graphing data is highly prized in quality management (Hare, Hoerl, Hromi & Snee, 1995). It is also interesting to note that Ainley et al. (2000) have developed the notion of active graphing when students use technology. It would seem that the actions of graphing and tabulating coupled with ongoing interpretation helped these two students gain insights into the data. Half of the students did not articulate such thinking *during* the construction of the tables and graphs and therefore we could not make any inferences about whether or not they were actively construing during this stage.

5.4. THE INTERTWINEMENT OF LOCAL AND GLOBAL THINKING

Fourthly, the intertwinement of local and global thinking was observed as a feature of the ways in which one thinks with and constructs meanings from data. This observation led us to consider that this continual shuttling between local and global thinking underpins the development of the four statistical thinking elements. Research on teaching statistics suggested that students first focussed on individual cases and then made the *statistical* transition to focus on group propensities (Konold et al., 1997; Ben-Zvi & Arcarvi, 1998; Cobb, 1999; Ainley et al., 2000). We observed, however, that when the students first focussed on individual cases they interpreted them by referring to 'known' group propensities or expected 'norms' and by conjecturing possibilities about deviations from the norm for that person. They compared individual cases and thus thought locally at the statistical level, but at the same time, at the contextual level, they thought locally and globally. The interchange of perspectives occurred within all the elements. For example, for the transnumeration element, when students are presented with raw data, one individual-case entry, say age, is considered at a local level. From this starting point: data are sorted on age and age becomes one global entity; data are split into local age subgroups; and then data are globally redefined as age-group and perceived as one entity again and so forth (Pfannkuch, Rubick, & Yoon, 2002a). For the variation element, at a statistical level, students notice variation by comparing two quantities and eventually shift to comparing patterns qualitatively, whereas, at the contextual level, they deal with variation by hypothesising relationships between individual cases and then shift to hypothesising relationships between variables (Pfannkuch, Rubick, & Yoon, 2002b). The focus on individual cases at the beginning of the investigation seemed to help in drawing out possible global relationships while at the conclusion stage a global relationship could be confirmed or discovered and the explanations for the relationship could range from none to one or several possibilities. Therefore the students seemed to build up their recognition of the relationships in the data through the dual use of local and global thinking, both statistically and contextually.

5.5. CHANGING STATISTICAL THINKING DIALOGUE ACROSS THE REPRESENTATIONS

Fifthly, for the development of statistical thinking across the elements it was observed that each element requires different ways of thinking with different representations. For example, variation thinking with raw data is different from variation thinking with graphs. The students' narratives highlighted awareness of how statistical thinking changed across the representations. Unless prompted, the thinking did not change for some of the students and continued to be the same for the graph as it was for the raw data. Even when the students engaged with the graph, their previous thinking with the other representations continued. It seemed that as the students moved from one representation to another their statistical thinking moved on a continuum back and forth among the representations as they gradually built up their understanding of the data. The continuum started with attention on one individual case and shifted towards noticing patterns in the whole data set. This raises questions about how to shift and draw students' attention to noticing that their thinking should change with changing representations. Cobb (1999) and Ainley et al. (2000) in their teaching experiments seemed to accomplish this subtle shift in students' focus with informed teacher questioning and by concentrating students' attention on looking within representations.

This statistical thinking continuum could be described in terms of the fundamental elements. Reasoning with statistical models changed from attending to isolated pieces of information, to answering the questions and discovering new relationships about global features within the graphical patterns (cf. Lesh et al., 1997). The integration of the contextual and the statistical changed from making up causal stories to fit one or two people, to causal inferences about perceived regularities in relationships between two or three variables. The variation dialogue changed from noticing exact differences between two individual people or noticing extreme values, to noticing systematic differences between grouped data. Transnumeration thinking changed from how to reclassify or order data, to how to represent the variables and data graphically and textually for communication purposes.

5.6. IMPLICATIONS FOR RESEARCH AND TEACHING

This small study suggested there was a big transition required in thinking from anomaly and pairwise variation to global variation. Transnumeration thinking was prevalent but at this stage seemed haphazard which is not surprising as these students were untutored. The students' structuring of data into different representations, be it changing classifications of data, tabulation of data, average, or graph showed different degrees of success. Representational fluency should mean not only interacting with a representation but also connecting that representation to other representations, and being able to change that representation to another representation. The integration of contextual and statistical information, knowledge and conceptions is necessary for making sense of complex systems. Students were arguing from a limited contextual and statistical knowledge base but if prompted sufficiently revealed a rich store on which to build and develop reasoning about data.

Furthermore, this study has led us to consider that reasoning with a statistical model seemed to have two distinct parts. The first part involved reasoning about a group propensity (Konold et al., 1997; Biehler, 1997) (and individual cases sometimes) either by looking for preconceived patterns or by seeing new patterns through a confirmatory or discovery approach. The second part involved decision-making under uncertainty. This decision-making was observed in the evaluation of the statistical model itself (cf. Lesh et al., 1997) whereby a judgement was made on the validity of the model and on whether another model was necessary. Other aspects of decision-making that were observed in the very initial stages of development were: probabilistic judgements which were quantitative (cf. Cobb, 1999) or qualitative (cf. Ainley et al., 2000); causal judgement on whether the variation was real or random and on possible contextual explanations; inference space judgement for generalisation of findings; evaluation of whether the results corroborated with the real situation; and conclusion judgement about what information could be communicated.

Through concentrating the four different lenses of "transnumeration", "consideration of variation", "reasoning with statistical models", and "integrating the statistical with the contextual", on students' thinking we believe it is possible to identify and describe the ways students are thinking within each element and identify the points where their thinking could be scaffolded to another perspective. For transnumeration, a scaffolding point may be the reclassification of raw data into new or transformed quantitative or qualitative variables. For variation, teaching strategies could be developed to turn students' attention from reading a datum for an individual case to thinking of that datum as belonging to a set of theoretical data; and from comparing values on a graph to looking at the range of the data for groups and informally quantifying the variation. When students reason with models, teaching strategies could involve predicting the pattern they might see including the type of variation expected, and drawing their attention to noticing other possible patterns with full awareness about the types of decision-making under uncertainty that are required. Teaching experiments should also consider how to shift students' thinking from their intuitive series comparison graphs for multivariate data to the conventional scatterplot graphs. The integration of the statistical with the contextual may mean scaffolding students' argumentation through attention to evaluation of causal reasoning when responding to statistical data. Overall, teaching strategies may need to focus first on comparing two individual cases and then gradually shift the focus to more and more cases.

This initial exploration into identifying and describing students' thinking from a theoretical framework has now opened up possibilities to do further research at a micro-level on students' thinking and to develop more explicit models for each element of thinking. Further exploration, however, should be continued at the macro-level since such research raises more awareness of the elements of thinking that should be considered overall in developing students' thinking. The study has also highlighted how the use of the Wild and Pfannkuch (1999) framework, helped dissect the student dialogue and reveal insights about the ways in which students are thinking. The model suggested to us to look for certain aspects which we probably would not have noticed without such a prompt. Educators awareness of these elements might give them a greater understanding of how statistical thinking could be developed in students. Reasoning with data is complex and requires fostering students' imagination and producing a web of connections between contextual and statistical knowledge. This integration enables students to construct meanings from data through a constant dialogue with a chain of statistical representations. Questions remain, however, about how such findings could be utilised to effectively develop learning from data in the classroom setting.

ACKNOWLEDGEMENTS

The authors acknowledge helpful comments and discussions on this paper with Jane Watson, Chris Wild, and Caroline Yoon.

REFERENCES

Ainley, J., Nardi, E. & Pratt, D. (2000). The construction of meanings for trend in active graphing. *International Journal of Computers for Mathematical Learning,* 5, 85-114.

Ben-Zvi, D. (2000). Towards understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning,* 2 (1&2), 127-155.

Ben-Zvi, D. & Arcavi, A. (1998). Towards a characterization and understanding of students' learning in an interactive statistics environment. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee & W. K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics* (pp. 647-654). Singapore: International Statistical Institute.

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. 45-56). Voorburg, The Netherlands: International Statistical Institute.

Biehler, R. (1997). Students' difficulties in practicing computer-supported data analysis: Some hypothetical generalizations from results of two exploratory studies. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 169-190). Voorburg, The Netherlands: International Statistical Institute.

Bright, G. & Friel, S. (1998). Graphical representations; helping students interpret data. In S. Lajoie (Ed.), *Reflections on statistics: learning, teaching, and assessment in grades K-12* (pp. 63-88). Mahwah, NJ: Lawrence Erlbaum Associates.

Chick, H. (1999). Jumping to conclusions: Data interpretation by young adults. In J. Truran & K. Truran (Eds.), *Making the difference. Proceedings of the 22nd annual conference of the Mathematics Education Research Group of Australasia* (pp. 151-157). Sydney: MERGA Inc.

Chick, H. (2000). Young adults making sense of data. 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. 157-164). Sydney: MERGA Inc.

Chick, H. & Watson, J. (2001). Data representation and interpretation by primary school students working in groups. *Mathematics Education Research Journal*, 13(2), 91-111.

Cobb, G., & Moore, D. (1997). Mathematics, statistics and teaching. *The American Mathematical Monthly*, 104(9), 801-823.

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning,* 1(1), 5-43.

Curcio, F. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, 18(5), 382-393.

Friel, S. & Bright, G. (1996, March). Building a theory of graphicacy: How do students read graphs? Paper presented at the *Annual Meeting of American Education Research Association*, New York.

Hancock, C., Kaput, J., & Goldsmith, I. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist,* 27(3), 337-364.

Hare, L., Hoerl, R., Hromi, J. & Snee, R. (1995). The role of statistical thinking in management. *Quality Progress*, February 1995, 53-60.

Kaput, J. (2000). Implications of the shift from isolated expensive technology to connected, inexpensive, ubiquitous and diverse technologies. In M. Thomas (Ed.), *TIME 2000 international conference on technology in mathematics education* (pp. 1-24). Auckland, New Zealand: The University of Auckland and Auckland University of Technology.

Konold, C., Pollatsek, A., Well, A. & Gagnon, A. (1997). Students analyzing data: Research of critical barriers. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 151-168). Voorburg, The Netherlands: International Statistical Institute.

Lehrer, R. & Schauble, L. (2000). Inventing data structures for representational purpose: Elementary grade students' classification models. *Mathematical Thinking and Learning*, 2 (1&2), 51-74.

Lesh, R. (2000). What mathematical abilities are most needed for success beyond school in a technology based age of information? In M. Thomas (Ed.), *TIME 2000 international conference on technology in mathematics education* (pp. 72-82). Auckland, New Zealand: The University of Auckland and Auckland University of Technology.

Lesh, R., Amit, M. & Schorr, R. (1997). Using "real-life" problems to prompt students to construct conceptual models for statistical reasoning. In I. Gal & J. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 65-84). Amsterdam: IOS Press.

Mevarech, Z. & 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. (2000). Graphical representations of statistical associations by upper primary students. *Mathematics education beyond 2000. Proceedings of the 23rd annual conference of the Mathematics Education Research Group of Australasia* (pp. 44-0-447). Sydney: MERGA Inc.

Pfannkuch, M., Rubick, A. & Yoon, C. (2002a). Statistical thinking: An exploration into students' variation-type thinking. *New England Mathematics Journal*, 34(2), 82-98.

Pfannkuch, M., Rubick, A. & Yoon, C. (2002b). Statistical thinking and transnumeration. In B. Barton, K. Irwin, M. Pfannkuch, & M. Thomas (Eds.), *Mathematics Education in the South Pacific. Proceedings of the 25th Annual Conference of the Mathematics Education Research Group of Australasia*, (pp. 567-574). Sydney: MERGA Inc.

Rubick, A. (2000). *The statistical thinking of twelve Year 7 and Year 8 students.* Unpublished masters thesis, The University of Auckland, New Zealand.

Shaughnessy, J. M. (1997). Missed opportunities in research on the teaching and learning of data and chance. In F. Biddulph & K. Carr (Eds.), *People in mathematics education. Proceedings of the 20th annual conference of the Mathematics Education Research Group of Australasia*, (pp. 6-22). Sydney: MERGA Inc.

Shaughnessy, J. M., Garfield, J. & Greer, B. (1996). Data handling. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 205-238). Dordrecht: Kluwer Academic Publishers.

Shaughnessy, J. M., Watson, J., Moritz, J. & Reading, C. (1999). School mathematics students' acknowledgment of statistical variation. Paper presented at the *77th Annual National Council of Teachers of Mathematics Conference,* San Francisco, CA.

Torok, R. & Watson, J. (2000). Development of the concept of statistical variation; an exploratory study. *Mathematics Education Research Journal*, 12(2), 147-169.

Tukey, J. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

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 masters dissertation, The University of Auckland, New Zealand.

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

TEACHING STUDENTS THE STOCHASTIC NATURE OF STATISTICAL CONCEPTS IN AN INTRODUCTORY STATISTICS COURSE[1]

MARIA MELETIOU-MAVROTHERIS
Cyprus Ministry of Education
meletiu@spidernet.com.cy

CARL LEE
Department of Mathematics,
Central Michigan University, MI 48859, USA
lee1c@cmich.edu

SUMMARY

The article argues that the persistence of student difficulties in reasoning about the stochastic, despite significant reform efforts, might be the result of the continuing impact of the formalist mathematical tradition, affecting instructional approaches and curricula and acting as a barrier to instruction that provides students with the skills necessary to recognize uncertainty and variability in the real world. It describes a study driven by the conjecture that the reform movement would have been more successful in achieving its objectives if it were to put more emphasis on helping students build sound intuitions about variation. It provides an overview of how the conjecture guiding the study was developed and linked to classroom practice, and briefly discusses the experiences and insights gained from a teaching experiment in a college level, introductory statistics classroom, which adopted a nontraditional approach to statistics instruction with variation at its core. By contrasting students' intuitions about the stochastic prior to instruction to their stochastical reasoning at the completion of the course, it illustrates the potential of the instructional approach as an alternative to more conventional instruction.

***Keywords*:** Statistics education research; Stochastic; Variation; Introductory statistics; Classroom experiment

1. INTRODUCTION

Significant reform efforts have been lately witnessed in statistics education, which have led to a movement away from statistics instruction emphasizing the abstract and the memorization of formulas and procedures. More importantly, these efforts have led to a general acknowledgment that learning occurs most effectively when students engage in authentic activities. Although many students are still being taught in traditional classrooms, there is already a large number of statistics instructors who have adopted alternative approaches to their teaching and many statistics classrooms are experiencing wide incorporation of technology. But, as Hawkins (1997) points out, for reform efforts to be successful, it is also necessary to change attitudes and expectations about statistics education. Changing long-held beliefs and attitudes towards statistics is proving to be quite difficult (Wilensky, 1993). Many people still view statistics as *"a branch of the older discipline of mathematics*" (Glencross and Binyavanga, 1997, p. 303). This affects statistics instruction and hampers the reform efforts.

The linear and hierarchical approach adopted by statistical courses and syllabuses, is testimony to the profound and continuing effect of the formalist mathematics culture on statistics education. The structure of almost every introductory statistics course is to first start with descriptive and exploratory data analysis, then move into probability, and finally go to statistical inference. It is assumed that this simplifies the process of learning by gradually leading students from more basic to more complex connections (Steinbring, 1990). However, presenting statistical content as a sequenced list of curricular topics might lead to compartmentalization of knowledge and fail to communicate to students the interconnectedness of the different statistical ideas they encounter in the course.

[1] *Statistics Education Research Journal, 1(2), 22-37, http:/fehps.une.edu.au/serj*
International Association for Statistical Education

Statistical methods were developed to help us filter out any "signals" in data from surrounding "noise". "Signals" are the messages, the meanings we find in *explained variation*, the patterns that we have not discounted as being transient. "Noise" is the *unexplained variation* that remains after we have "removed" all patterns. Randomness and probability theory are "human constructs" created to deal with unexplained variation. We use probability theory to model and describe phenomena in the world for which no patterns can be discerned, assuming that they had been randomly generated. Thus, "*what probability is can only be explained by randomness, and what randomness is can only be modeled by means of probability*" (Steinbring, 1990, p. 4). Stochastical knowledge is created as "*a relational form or linkage mechanism between formal, calculatory aspects on the one hand, and interpretative contexts on the other*"(Steinbring, 1990, p.5). However, the classroom culture often comes in sharp contrast with this conception of stochastical knowledge as being developed through a "self-organized" process that balances the objective aspects of a situation and the formal means employed to model and describe it. The linear, completely elaborated and hierarchical structure of knowledge presentation encourages the development of the chance concept as a concrete, totally clear and unambiguous generalization defined by methodological conventions. Steinbring (1990), who analyzed teaching episodes from several different classrooms in order to see how the concept of chance was introduced, found that in all of those episodes "chance" was first introduced through performing and discussing a chance experiment. An attempt was then made to describe the experimental outcomes using a rule or a simple stochastical model. Naturally, there was always variation observed between the theoretical predictions and the empirical data. The pattern of justification for the observed variation, regardless of its size, always was that the difference between the empirical result and the theoretical prediction was produced by "chance" (Steinbring, 1990). The difference between theory and experiment was thus neutralized, with chance degenerating into "*a substitute for justification, which serve[d] to deny the importance of the difference between theory and empirical facts in probability*" (Steinbring, 1990, p. 14).

The assumptions posed in the statistics classroom are often too simplistic. Although not necessarily denying underlying causal explanations in case of chance events, a probabilistic approach views them as impractical and adopts a 'blackbox' model (Biehler, 1994, p. 10). However, as Biehler (1994) indicates, the assumption of independence is not plausible in many real-world contexts: "*Even coin flipping can be done in a way that independence has to be rejected in favor of serial correlation, and physical theories can be developed to explain some aspects of coin flipping*" (p.10). Borovcnik and Peard (1996) warn that instruction has traditionally underrated the complexity and dangers of using pseudo-real examples that conflict with students' emotions or with their common sense.

The over-emphasis of the traditional mathematics curriculum on determinism and its "*orientation towards exact numbers"* (Biehler, 1997, p. 187) affects statistics instruction, becoming an obstacle to the adequate judgment of stochastic settings. The law of large numbers is often presented as a canon in the statistics classroom, giving students the false impression that the stabilization of the relative frequency of repeated sampling to the ideal value is *guaranteed*. Similarly, instruction leaves students with the impression that a larger random sample *guarantees* a more representative sample. There is a deterministic mindset and an over-reliance on rules and theorems, forgetting that the uncertainty and variability accompanying all finite statistical processes implies that a sample is almost never totally representative of the population from which it was selected. People have difficulties in distinguishing between the real-world problem and the statistical model. At one extreme are many people who use statistical methods for solving real-world problems in the same way they would use an artificial mathematics problem coming out of a textbook. On the other extreme, we find people who distrust statistics completely because, unlike mathematics, it deals with uncertainty. Both of these two extreme attitudes suggest inadequate understanding of statistics as a decision-support system (Biehler, 1997).

Current practices in statistics education have evolved from a background quite different from today's needs and possibilities. Hoerl, Hahn, and Doganaksoy (1997), argue that we must completely rethink the sequence of topics in order to achieve the objectives for introductory education, which for them should be to "*help students unlearn their deterministic view of the world"* (p. 152) and recognize uncertainty as a characteristic of reality. Chance and Rossman (2001) discuss four perspectives on the sequencing of topics in an introductory statistics course and lay out the merits and drawbacks of each perspective. In this article, we describe the experiences obtained from a teaching experiment that implemented an alternative path to statistics instruction with variation at its core, conjecturing that this path would lead to stronger and deeper understandings by helping students see the "big picture" of statistics (Moore, 1997). By contrasting students' intuitions about the stochastic prior to instruction to their stochastical reasoning at the completion of the course, we attempt to illustrate how the instructional approach employed in the study proved a promising alternative to more conventional instruction.

2. DESIGN OF THE STUDY

2.1. BACKGROUND

The motivation for the study was provided by the results of a previous study of twenty-two students who had just completed an introductory statistics course. The results of that study (Meletiou, Lee, and Myers, 1999), agreed with the main findings of research in the area of stochastics education. Similarly to the research literature, we found in that previous study that the students we interviewed, regardless of whether they came from a lecture-based classroom or from a course following the PACE (Projects-Activities-Cooperative Learning-Exercises) model, had poor intuitions about the stochastic and tended to think deterministically. PACE is an approach developed by Lee (1997, 2000), that attempts to provide a structured framework for integrating projects and hands-on activities conducted cooperatively in a computer-based classroom environment.

After witnessing students' poor statistical reasoning and their deterministic mindset, we concluded that student difficulties might stem from inadequate emphasis paid by instruction to the notion of variation, and to the connections among statistical ideas. This led to a decision to modify the PACE model. The course described in the next sections, which was taught by Lee, although still having a format similar in many ways to that followed in previous semesters, differed significantly in both curricular emphases and structure. A nontraditional approach to statistics instruction that had variation as its central tenet, and perceived learning as a dynamic process subject to development for a long period of time and through a variety of contexts and tools, was adopted.

2.2. VARIATION AS THE CENTRAL TENET OF STATISTICS INSTRUCTION

Statistical thinking is concerned with learning and decision-making under uncertainty. Variation is a critical source of uncertainty. It is the fact that all processes vary which creates the need for statistics. It is the need to deal with variation through measurements that provides a (numerical) basis for comparison that produces data (Snee, 1999). We use statistical tools to analyze this data and observe the pattern that exists despite (or because of) the variation. Thus, according to Snee (1999), the elements of statistical methods are variation, data, and statistical tools. Understanding of variation and using this understanding to improve the performance of processes is the core competency and it should be the focus of statistical education, research, and practice (Snee, 1999). Understanding what data is relevant and how to construct proper methods of data collection and analysis enhances successful application of this core competency (Snee, 1999).

The central element of statistical thinking is variation, and instruction should aim at providing students with the skills necessary to be able to notice and acknowledge it, to explain and deal with it. But, if variation is indeed to be "the standard about which the statistical troops are to rally" (Wild and Pfannkuch, 1999, p. 235), we have to arrive at a common conceptualization of statistics instruction in terms of variation. Wild and Pfannkuch (1999) offer the following three "variation" messages as a starting point: (1) variation is omnipresent; (2) variation can have serious practical consequences; and (3) statistics give us a means of understanding in "a variation-beset world" (see also Cobb, 1992). The subsequent messages of the statistics classroom provide information about tools and methods statistics offers us to make sense of the omnipresent variation.

Pfannkuch's (1997) epistemological triangle, views variation as the broader construct underlying statistics instruction (Figure 1). In encouraging students to develop their understanding of the concept of variation, Pfannkuch's epistemological triangle aims at the same time at promoting richer understanding of all the other main statistical ideas. The epistemological triangle indicates that for conceptualization of variation, a combination of subject and context knowledge is essential (Pfannkuch, 1997). The inter-linked arrows indicate the strong linkage that has to be created between statistical tools and the context of the problem. The assumption underlying the epistemological triangle is that the concept of variation would be subject to development over a long period of time, through a variety of tools and contexts (Pfannkuch, 1997).

Pfannkuch's model bases instruction on contexts directly connected to students' experience, it recognizes that adequate statistical reasoning requires more than understanding of different ideas in isolation. It demands "*integration between students' skills, knowledge and dispositions and ability to manage meaningful, realistic questions, problems, or situations*" (Gal and Garfield, 1997, p. 7). Content is no longer a sequenced list of curricular topics taught in isolation, but "*an interrelated repertoire of conceptual tools that can assist one in making sense of, and gaining insight and prediction over interesting phenomena*" (Confrey, 1996).

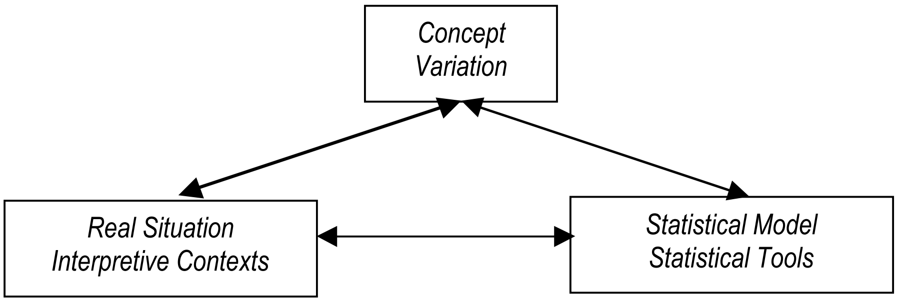

Figure 1. Pfannkuch's Epistemological Triangle

The conjecture driving our study was that the reform movement would be more successful in achieving its objectives if it were to put more emphasis on helping students build sound intuitions about variation and its relevance to statistics (Ballman, 1997). Pfannkuch's epistemological triangle, which calls for the re-structuring of statistics instruction by offering a nontraditional path with variation at its core, seemed well suited for meeting our research aspirations, and was employed in the study to guide curriculum development and instruction.

2.3. CONTEXT OF STUDY AND PARTICIPANTS

The site for the study was an introductory statistics course in a mid-size Midwestern university in the United States. One of the authors, Lee, was the instructor of the course. The study lasted over the span of five weeks. The course began on the last week of June and ended on the first week of August. Class met four times a week, for two hours each time. The number of students in the class was thirty-three (nineteen males and fourteen females). Most of the students in the class (twenty-two students) were majoring in a business-related field of study. Only few had studied mathematics at the pre-calculus level or higher.

2.4. DATA GENERATION

In examining students' learning progress and outcomes, a variety of both qualitative and quantitative data gathering techniques were employed. By assessing students' understanding prior to instruction, and then monitoring changes in their thinking throughout the course, the study attempted to develop a detailed description of the processes students go through in order to become able to intelligently deal with variability and uncertainty.

Given the number of students involved, it was impossible to observe closely every single student in the course. Therefore, we chose to study two groups of students. The primary group consisted of a subset of eight students (five males and three females) and the secondary group encompassed the whole class. Although data from both groups were used in the analysis, our focus was on investigating and describing the learning experience of students in the primary group. The selection criterion for the primary group was willingness to participate in the study.

Next, we explain the specific data that was generated. We have found it useful to describe the data generation process separately for each of three phases of the course: (a) beginning of course, (b) duration of course, and (c) end of course.

Beginning of course

In order to be able to follow the students' conceptual development process, a good understanding of their thinking prior to instruction is required. A diagnostic questionnaire with ten, mostly open-ended, questions was given to all students on the first day of class to assess their intuitive understanding of variability prior to instruction. Additionally, we conducted individual interviews of students in the primary group. The interviews, which were audio-taped, were semi-structured. In the first part of the interview, we posed students in the primary group a set of open-ended questions, which aimed at further investigating their intuitive understandings of variability. The second part of the interview was a follow-up of the questionnaire on variability completed by the whole class. We went over the questionnaire with the students in order to clarify the reasons for the different responses they had provided. Despite the open-ended nature of the diagnostic questionnaire, one-to-one communication with students in the primary group allowed a more thorough investigation of student reasoning.

Duration of course

The data gathering techniques employed during the course included: (1) direct and participant observations, (2) interviews of students, (3) video-taping of group activities, (4) pre- and post-activity assessments, (5) field notes, and (6) samples of student work. Drawing data from several different sources permitted cross-checking of data and interpretations.

Although in this article we do not present findings from the analysis of the data collected during the course, the insights drawn from this analysis did play an important role in shaping the interpretation of results from the end-of-course assessment presented in Section 5. The continuous monitoring throughout the course of the effect of instruction on student learning was constantly supplying us with valuable information on their levels of concept attainment. The assessment strategies used to support and evaluate students' conceptual development helped students further clarify their reasoning strategies.

End of course

In order to assess students' understanding at the completion of the course, a diagnostic questionnaire with fifteen open-ended questions was administered to the whole class. Additionally, we again conducted individual interviews with each of the students in the primary group. During the interviews, we went over the diagnostic questionnaires students had taken at the beginning and the end of the course, as well some assessment tasks they had completed during the course, and also reminded students of some of the responses that they had provided during the interview at the beginning of the course. We prompted students to express their agreement or disagreement with responses they had given earlier in the course, so that we could check whether their reasoning had changed in any way since then.

3. FINDINGS FROM BEGINNING-OF-COURSE ASSESSMENT

Findings from the assessment at the outset of instruction further supported the conjecture that variation is neglected, and its critical role in statistical reasoning is under-recognized. In this section, we present findings from two tasks included in the diagnostic questionnaire given to the whole class, and three tasks completed by students in the primary group during the follow-up interview, which are indicative of the tendency we observed at the beginning of the course for students to think deterministically and to have difficulties in differentiating between chance variation in the data and variation due to some form of underlying causality.

Students' performance in a question in the pre-assessment adapted from Rubin et al (1990), which examined how they balanced the ideas of sampling variability and sampling representativeness, illustrates their propensity to underestimate the actual level of random variability. In this question (see *Gummy Bears Question* in the Appendix), students were told that the Easter Bunny was distributing many packets of 6 Gummy Bears at the Easter Parade that he had made up by grabbing handfuls of Gummy Bears out of a large barrel containing two million green and one million red Gummy Bears that had been randomly mixed, and were asked to estimate the number of green Gummy Bears in a packet. Everyone gave *"4 green, 2 red"* as the estimate, but also everyone realized that *"not every student got exactly 4 green every time because there's variability".* They intuitively understood that probability is the limiting relative frequency, which only approximately holds for real data: *"That is just the mathematical way of figuring it, that number will fluctuate"; "Expected ratios are a general rule, not a formula for each individual occurrence*"; *"It is nearly impossible for the ratio to hold perfectly, unless the Easter Bunny uses his Easter magic."* However, although students did recognize that *"there will be a variation on the pattern of green bears in each bag, because of the random grabbing of the bears when they were placed in the bags*", when asked to estimate the proportion of packets with 4 greens, almost all of them underestimated the effect of sampling variability and greatly overestimated this proportion. Only two students gave estimates that came close to 33%, the actual probability of 4 greens (found by modeling the situation as a Binomial distribution). The estimates that the rest of the students gave ranged from 50%-92%. Several students wrote that they expected 66% of the packets to have 4 greens in them. Rubin et al. (1990) who gave this question to high school seniors also noticed that students answered this question by focusing on samples that mirrored the population proportion of 2G:1R. They over-relied on sample representativeness, underestimating the frequency of samples near the tails of the distribution and overestimating the frequency of the modal sample.

The tendency to underestimate the effect of sampling variability and expect small samples to match population properties was also witnessed in a question in the pre-assessment taken from Garfield and delMas (1990). The question (see *College Students' Interview Question* in the Appendix) described how a worker of a student organization went about conducting a survey at a certain college where half the students were women and half were men and the several measures he took to ensure good representation of all students. Students were told that out of the last 20 students interviewed, 13 were women and 7 were men and were asked whether they thought there would be more women or more men in the next 20 students interviewed. Only 35% of the students correctly stated that one should expect about an equal number of men and women, as who has been selected so far does not affect who will be next selected. Thirty-two percent argued that since more women than men were selected so far, they expected the opposite trend to start happening, 16% tried to find causes behind a difference that – given the small number of people interviewed this far – could be easily explained by chance variation while the rest, employing the "law of small numbers", thought that the trend of selecting more females than males should continue.

The tendency to underestimate the role of chance variation was more pronounced in real world-contexts. Although students did seem aware of the dangers involved when drawing conclusions from small samples, when asked to make their own judgments based on data, they often ignored these dangers and, exaggerating the reliability of the information provided, did not hesitate to use small samples as a basis for inferences. The responses of students in the primary group during the follow-up interview to the *Birth Defects* task (taken from Pfannkuch and Brown, 1996) are indicative. In this task (see Appendix), students were told that each year approximately seven children are born in New Zealand with a missing limb. They were shown a map of New Zealand divided into five regions, with the number of children born with this abnormality in each region during the previous year located on the map. According to the map, there were no children born with the abnormality in either the top or bottom region, while the number of children born with a missing limb in each of the other three regions were three, three, and two respectively. They were asked to comment on this, given the information that in New Zealand one-third of the population lives in the top region and one-sixth of the population in each of the other regions.

Pfannkuch and Brown (1996) found that students' understanding of variation in small samples was poor in this context. Whereas an analysis combining both probabilistic and deterministic thinking would have been more appropriate, all of the students they interviewed gave deterministic explanations, and it was only after repeated probing that some suggested the need for more data. Our findings were very similar. We observed very strong deterministic reasoning in all of the students. George, for example, *"wouldn't want to live in the middle of New Zealand"*, and Julie was convinced that there must be an outside factor causing the difference: *"There is always a chance that anything can happen but, 3 and 0 in the other…there must be a reason for that.."*

Pfannkuch and Brown (1996) conjectured that students' neglect of probabilistic thinking might be the rich experience they have with similar controversial data often appearing in the media and seldom being explored from a statistical perspective. When asked what they think of the possibility of obtaining the outcome {3,3,3,4,4,5,5} (order unimportant) when rolling a fair die 7 times, no student found such an outcome surprising. They approached this problem very differently from the New Zealand one although it is analogous – obtaining 1 or 2 on the die corresponds to the top region of the map where one-third of the population lives, and obtaining a 3,4,5, or 6 corresponds to each of the other regions. Similarly, in our study, students found such a result quite likely due to the small sample size that allows extreme outcomes: *"I think nothing of the results. After a thousand throws each number will be picked around 1/6 of the total throws."*

The different way in which students approach the two problems indicates how much more prone we are to look from a stochastical perspective at standard probability tasks than problems situated in certain real-life contexts (Pfannkuch and Brown, 1996). Students are at fault in searching for causes behind data of such nature, where a lot of other factors besides chance might influence the occurrence of birth defects. Still they should realize that a sample of only 7 children has too little information in it to help us find causes since small samples often have large natural variation even when no causes are operating. They should have shown the same sensitivity to the effects of sample size they showed in the *Child Psychologist* question (see Appendix) taken from Garfield and delMas (1994). In that question, students interviewed were asked to judge the validity of conclusions drawn by a child psychologist who, after studying 5 infants and finding that 4 showed preference for the one toy, concluded that most infants would show a preference for this toy. Every single student interviewed challenged the psychologist's conclusions. Tim for example said: *"4 out of 5, I know it's good for like 4 out of 5 dentists prefer this* *kind of toothpaste, whatever on the commercials, but I would say you need at least a 100 kids...I could get my 5 sons and persuade 4 of them."* His response comes in sharp contrast to how he responded to the *Birth Defects* question:

Int.: *Just by looking at the map, do you see any connection between where one lives and how many kids are born with a missing limb*?

Tim: *Oh, yeah. They correlate because the 1/3 that lives there has 0 because probably there are more doctors and more hospitals and only 1/6 lives there, so there must be something going wrong there. So yes, there has to be a reason*.

Int.: *Do you see that the numbers are small? Do you think this is something you should take into account*?
Tim: *Why?*

4. COURSE DESCRIPTION

4.1 CURRICULUM

The design of the intervention was guided by our conjecture that if statistics curricula were to put more emphasis on helping students improve their intuitions about variation and its relevance to statistics, we would be able to witness improved comprehension of statistical concepts. At the same time, the time constraints and confines of the curriculum were also taken into account. Instruction included the set curriculum typically covered in an introductory statistics course, but was expanded in such a way as to include throughout the course activities that aimed at raising students' awareness of variation. The different topics were approached through the lens of the conjecture. Adjustments to the curriculum were also guided by the following two principles (adapted from Wild and Pfannkuch, 1999):

1. *Complementarity of theory and experience*: Statistical thinking necessitates a synthesis of statistical knowledge, context knowledge, and information in the data in order to produce implications, insights and conjectures. If the statistics classroom is to be an authentic model of the statistical culture, it should model realistic statistical investigations rather than teaching methods and procedures in a sequential manner and in isolation. The teaching of the different statistical tools should be achieved through putting students in authentic contexts where they need those tools to make sense of the situation. Students should come to value statistical tools as a means to describe and quantify the variation inherent in almost any real-world process.

2. *Balance between stochastic and deterministic reasoning*: Instruction should view as an important precursor of statistical reasoning students' intuitive tendency to come up with causal explanations for any situation they have contextual knowledge about. It should present statistical thinking as a balance between stochastic and deterministic reasoning and should stress that statistical strategies, based on probabilistic modeling, are the best way to counteract our natural tendency to view patterns even when none exists, to distinguish between real causes and ephemeral patterns that are part of our imagination.

4.2 CLASSROOM SETTING

The typical setting during a class session was such that it encouraged "statistical enculturation". The instructor's knowledge and behavior contributed towards the creation of an authentic model of the "statistical culture" (Biehler, 1999). It was a setting that modeled realistic statistical investigations and in which statistical dispositions such as appreciation of data were valued and nurtured. The instructor was trying to increase students' awareness of variation, to help them realize that it is the existence of variation, which creates the need for statistical investigations. He would keep on emphasizing that the reason we use statistical tools is to describe trends and patterns and deviations from those patterns existing in the data because of the variation inherent in every process.

No method or procedure was taught in isolation. In contrast to more typical approaches where reference to problems is made to demonstrate statistical content, reference to statistical content in this class was made (in students' mind at least) to help understand a situation, to assist a statistical investigation. The emphasis was on statistical process and along the way students got to learn different statistical methods and procedures. The hope was that by putting students in situations where they needed tools such as the standard deviation, they would realize their usefulness and not wonder why anyone would ever bother to invent them (Erickson, 2000).

Unlike conventional instruction that tends to focus on mechanical application of methods and procedures, the emphasis of this course was on recognizing applicability and interpreting results in context (Wild and Pfannkuch, 1999). In introducing binomial distributions, for example, the main goal was not to teach the formal properties of the binomial distribution but to help students recognize a binomial setting and understand why we can apply this population distribution to model a certain variable and in what ways this is useful. Students were first given a description of five different situations that could be modeled using the binomial distribution, and were asked to work with their group in order to figure out their common properties. Group work was followed by a whole class discussion during which students laid out the main properties of the binomial distribution. Introduction of the probability formula describing binomial distributions was done only after students had brought up several examples of situations in the real world that could be modeled using the binomial distribution.

In order to simplify mathematical relations and build links to students' intuitions, the course emphasized the use of analogies from students' everyday experience in contexts familiar to students. Instruction stressed the complexity of real-life situations rather than making simplistic assumptions that would conflict with students' common sense. When, for example, discussing independent events, and after students had given typical examples of independent events such as coin tossing and die rolling, the instructor asked the class whether the success of a *"free throw"* of a basketball player is independent from the success of his previous *"free throw"*. Students argued that it depends on how the player responds to pressure, on how well he did on the previous throw etc. The instructor agreed remarking that, *"in real life it's hard to say with a straight yes or no"*. He did not reject students' causal explanations although "hot hand" is an example often used by many statistics educators in their pleading for probabilistic reasoning. Tversky and Gilovich (1989) showed, using empirical data, that a binomial model well explains runs (streaks) in basketball player failures. According to this model, the chance of success in a shot is independent from the previous shot, and Tversky and Gilovich, and subsequently many teachers and researchers, concluded that people's tendency to detect patterns (hot hands) is often unwarranted. One need not look for specific causes like nervousness since chance patterns produced by a completely random process well explain the data. However, as Biehler (1994) has pointed out, even when the binomial model well explains the variation in a dataset, one should not exclude the possibility of alternative models that give better prediction and that suggest causal dependence of individual throws. The instructional approach in this class clearly made students aware of this perspective. Similarly, when talking about slot machines in a casino, the instructor noted: *"Although in theory when you put a coin and you pull it down and then you put another coin and you pull it down, although those two events should be independent, mechanically they might not be."*

4.3 EMPHASIS ON THE PROCESS OF CONJECTURING AND DISCOVERY

The idea of making conjectures ran throughout the course. Students would state what they believed might or might not be true, and then looked critically at the data to evaluate their statements. While the instructor encouraged students to make conjectures he, at the same time, also tried to help them understand that conjecturing is not enough – one has to evaluate one's predictions by looking closely at the data and making comparisons (Erickson, 2000). Evaluation of conjectures would typically begin informally by using one or more graphical displays. Students would be encouraged to describe the main features of the distribution displayed by the graph(s), always emphasizing the need to take into account not only the center, but also the spread. Students would look at the displays and try to give explanations for the patterns observed and for the departures from those patterns. Sometimes these explanations would be proposals for a possible model to summarize the dataset.

The evaluation of conjectures would then become more quantitative. An analysis using appropriate numerical summaries would be made to support or refute the conjectures originally made by students. At the start of the course, the analysis was made using simple numerical summaries. Eventually, more tools were added to students' repertoire and the mathematization of the data gradually became more formal. Even when the data agreed with their initial conjecture, the instructor would encourage students to also come up with alternative explanations and see that there can be multiple explanations for a phenomenon, in the hope that this would make them "*less likely to assume that their data 'proves' the obvious cause*" (Erickson, 2000, p.2). The issue of unrealistic or unrelated conjectures that went beyond the information provided by the data was often raised during the process of conjecturing and discovery.

4.4 EMPHASIS ON DATA PRODUCTION AND VALIDITY OF MEASUREMENTS

A special emphasis of the course was on data production issues. Unlike many other statistics courses where study design issues are discussed as a separate topic and almost never appear again, they were continuously brought up in this course. Throughout the course, instruction was stressing the fact that data are numbers collected in a particular context that are studied for a purpose (Rossman, 1996), and that the quality of the conclusions we draw depends on how the data were obtained. When, for example, students were examining graphs, the instructor would point to them that observed patterns in the data depend to a great extent on how the data were obtained and that they might be misleading if data collection had not been properly done. When discussing inferential methods, he would stress that the validity of inferences drawn is based on the assumption of probability-based data production. With regards to the inferential advantages of a larger sample size, students were given several examples of situations in which there was bias in the sample selection process and/or the measurement system in order to realize that, in such situations, increasing the sample size would probably not lead to more valid conclusions. The issue of data production has been stressed by statistics educators as a critical part of the needed reforms in introductory statistics instruction (e.g. Moore, 1997; Garfield, 1995; Hogg, 1992). It was emphasized at each possible chance during the course.

5. FINDINGS OF END-OF-COURSE ASSESSMENT

In our previous study of PACE and other statistics students (Meletiou et al., 1999), we had witnessed superficial knowledge of statistical concepts and a tendency to think deterministically and seek causes behind ephemeral patterns in the data. In contrast, students in the current study were found to have much better understanding of the relationship between chance and regularity, to reason much more effectively about the stochastic.

The results of the assessment at the end of the course were very encouraging. They indicated that students in the current study recognized that, in addition to knowing about the center of a distribution, one also needs information about its spread. All of the students acknowledged that whenever comparing measures of center one always ought to also take "the overall spread" into account. Most of them had good understanding of the meaning and purpose of the different numerical summaries they had learned in class. It was for example very impressive that, in contrast to our previous research findings where almost no student really understood what standard deviation means, the majority of students in this study had a pretty good grasp of the meaning and use of standard deviation. They explained that one calculates standard deviation to get *"information about the distribution between the scores, the distance…outside the center*", *"to figure out the deviation, the average deviation of the scores from the mean."* They also all knew that, in addition to standard deviation, measures such as *"the range of the box"* (interquartile range), also give us information about the spread of a dataset. They understood that mean and standard deviation are not the only two measures that define the shape of the distribution. And although still having some difficulties with constructing and interpreting graphs, these students' understanding was much more sophisticated than that of students in the previous studies we had conducted.

Several tasks were given to students at the end of the course to investigate their understanding of the relationship between sampling variation and sampling representativeness. One of the tasks was one of three versions of an assessment item used by Shaughnessy, Watson, Moritz, & Reading (1999) in an exploratory study on student understanding of variation. A total of 235 primary students (grades 4 to 6) and 89 secondary students (grades 9 and 12) from the US and Australia had participated in that study. This version, which the authors called the CHOICE version, was given to a total of 105 students. Students had to choose, among five possible lists, the one that is most likely to represent the number of reds drawn by five students who each drew 10 candies out of a bowl of 100 wrapped candies that had 50 reds (see *Candies Question* in the Appendix).

In analyzing student responses, the same procedure as that of Shaughnessy et al. (1999) was followed. Responses were scaled both on the basis of their use of centers and of their use of spreads. For the "centering" scale, student responses were categorized as LOW, FIVE or HIGH. Responses for which the mean number x of reds was $4 < x < 6$, were classified as FIVE, otherwise they were classified as either LOW or HIGH. For the spread scale, the following categories were used: NARROW, REASONABLE, and WIDE. Responses in which the range was 7 or more are pretty unlikely to occur and were classified as WIDE, and so are those with ranges less than or equal to 1, which were classified as NARROW. Ranges between 2 and 7 were considered REASONABLE. According to the scale, responses can be classified as follows:

Table 1. Classification of Student Responses

Response	*Center Classification*	*Spread Classification*
A: {8,9,7,10,9}	HIGH	REASONABLE
B: {3,7,5,8,5}	FIVE	REASONABLE
C: {5,5,5,5,5}	FIVE	NARROW
D:{2,4,3,4,3}	LOW	REASONABLE
E: {3,0,9,2,8}	FIVE	WIDE

The expected number of reds drawn by each student is 5. Under an ideal situation, response C would have the highest probability of occuring. However, in answering this question, students ought to take into account not only the center of the distribution, but sampling variation also. Based on a binomial model of the problem, one standard deviation away from the expected number of reds in a sample of ten candies is 1.581. Although this question considers only five students, a fairly small sample, a reasonable spread to expect in the outcomes would be between 0 and 3, but, very unlikely to always be exactly zero. The best response is therefore B, which is centered on 5 and is also a reasonable response in terms of spread.

Table 2 compares the performance of students in the Shaughnessy et al. (1999) study, with that of students in this study.

Table 2. Results of Current Study vs. Results of Shaughnessy et al. (1999) study

Classification	*Shaughnessy et al. Study* %	*Current Study* %
Center		
Low	13	0
Five	56	100
High	27	0
Unclear	4	0
Spread		
Narrow	16	12
Reasonable	76	88
Wide	4	0
Unclear	4	0
Correct		
Five, Reasonable	35	88

Students in the current study did better in estimating both center and spread. Instruction seems to have been particularly effective in helping them take both spread and center into account. Whereas in the Shaughnessy et al. (1999) study, only 35% of the students belonged to the FIVE, REASONABLE category (i.e. chose response B), in this study the percentage of students belonging to this category was 88%. Since helping students move away from "uni-dimensional" thinking and be able to integrate center and variability into their analyses and predictions should be one of the main goals of statistics instruction, the results are encouraging. It is an important accomplishment of instruction given that in the Shaughnessy et al. (1999) study, although most students' measures of spread were reasonable, they predicted values that were either too high or too low on the centering scale. Also, in that study, the use of words explicitly referring to variation was quite rare. In contrast, students in the current study gave explanations that indicated they were integrating ideas of spread and center:

"Because 50% of the candies are red, the handfuls should be close to 5 reds each time so B. Not C because it's random, there is a margin of error".

"Because they all range around 5 per pick, as would a sample with 50% reds. The others seem too far away or impossible, like C".

"Because the average that would be expected should be 5 with some variation above and below the expected value".

"It's unlikely with .50 probability of reds that anyone got 0 or 10 or straight 5's. There are .50 reds and so we would expect to see more of those but this is a random sample and thus there should be some variability. We expect to have some below 5 and some above. B shows that".

" It's all about variance, but "central-tendency" must always be counted".

Of course, one should take into account the fact that students in the Shaughnessy et al. (1999) study were primary and high school students, whereas the present study deals with college students having completed a statistics course. Nonetheless, in that study, while a steady growth across grade levels on the "centering" criterion from 34% at Grade 4 to 83% at Grade 12 was observed, there was "an apparent oscillation on the variability criterion across grade levels." The researchers noted "a high spike occurring in our Grade 9 students, and then a drop off at Grade 12, for both the REASONABLE and the FIVE, REASONABLE categories." They speculated that the steady growth in the FIVE category is an indication of the considerable focus of school curricula on "center". A possible explanation they saw for the oscillation at Grades 9 and 12 is that Grade 9 students participating in their study were spending more time on data analysis than the higher level mathematics students, whose school work on probabilities might have interfered with their reasoning about this problem. The exposure to probability did not seem to interfere with the reasoning of the students in the current study.

Table 3. Results of end-of-course assessment compared on "College Interviewer" Question

Response	*Beginning %*	*End %*
A: The worker seems to interview more women than men. There could be several reasons for this. Perhaps women are more willing to talk about their opinions. Or, maybe the worker goes to areas of campus where there are more women than men. Either way, the worker is likely to interview more women than men out of the next 20 students.	16	6
B: Since half of the students on this campus are men and half are women, you would expect a 50/50 split between the number of men and women the worker interviewed. Since there tended to be more women than men so far, I expect the opposite trend to start happening. Out of the next 20 students the worker interviews, there will probably be more men than women so that things start to balance out.	32	24
C: Half the students on this campus are men and half are women. That means that the worker has a 50/50 chance of interviewing a man or a women. It should not matter how many men or women the worker has interviewed so far. Out of the next 20 students interviewed, about half should be men and half women.	35	64
D: So far, the trend seems to be more women to be interviewed than men. Out of the next 20 students the worker interviews, I would expect the same thing to happen. The worker will probably interview more women than men out of the next 20 students.	17	6

At the completion of the course, students were given again the same question given at the beginning which was describing how a worker of a student organization went about conducting a survey at a certain college where half the student population were women and half were men, and was asking students to predict whether there would be more women or more men in the next 20 students the worker interviews given that out of the last 20 students interviewed 13 were women and 7 were men (see *College Students' Interview Question* in the Appendix). Student performance at the end of the course compared to their performance on the same question prior to instruction, is another example of the positive effect of instruction in helping improve students' probabilistic reasoning (see Table 3).

Sixty-four percent of the students at the end of the course, compared to 35% of them in the pre-assessment, realized that due to the independence of random samples, one should still expect that, out of the next 20 students interviewed, about half should be men and half women (choice C). There were still a considerable proportion of students (24%) employing the balancing strategy and arguing that they expected the opposite trend to start happening, but in general, students' performance was much improved.

Students were, at the completion of the course, much less prompt compared to the beginning, to assume that short-term fluctuations in the data must be causal and develop causal explanations. When for example we reminded students in the primary group, during the end-of-course interview, of the *Birth Defects* question (see

Appendix), they all had a different opinion about this situation compared to the beginning of the course. They stressed that the number of children is so small that one cannot give causal explanations. Tim, for example, who had argued in the beginning of the course that the probability of giving birth to a child with a missing limb correlates with where one lives, now pointed out: *"There is not enough information to…it's not a big enough…it's only 7 people. It's not enough number of subjects to understand what's going on."* Lucas remembered that when he first saw this question he was thinking, *"there might be something wrong with the sanitation or the water, something like that."* Now though, he realized that *"this is only one year so, last year or the year before, they could have had 3 down here and 2 up there and 2 over here. You have to look at many years to see what's happening."* He added: *"That's why I liked this class. I learned to look at the big picture of things."*

The efforts of instruction to present statistical thinking as a balance between deterministic and stochastical reasoning succeeded in helping students move away from "uni-dimensional" thinking and integrate center and variation into their analyses and predictions. Although not totally letting go of their deterministic mindset, students were much more willing to interpret situations using a combination of stochastic and deterministic reasoning. The course increased significantly student awareness of sampling variation and its effects. Instruction managed to get across to them the idea that "thinking about variability is the main message of statistics" (Smith, 1999, p. 249).

6. IMPLICATIONS FOR INSTRUCTION AND RESEARCH

The expectation that students will transfer the understanding obtained through coins, dice, and games of chance to everyday contexts seems to be a naïve assumption, as previous research studies (Pfannkuch and Brown, 1996), as well as student assessment at the outset of this study, have indicated. The skills required to understand variation in random devices are very different from the skills required to understand variation in real-life contexts. Instruction needs to take into account the great variety of prior beliefs, conceptions, and interpretations that students bring to each situation.

This study investigated how a teaching pedagogy focusing on data and variation centered on students' experiences could promote understanding of the stochastic nature of statistical concepts. Unlike more typical approaches, which attempt to develop probabilistic reasoning through standard probability tasks, the model employed in the study based instruction on realistic contexts directly connected to students' experience. Although the conclusions drawn from the study focus on a single group of students for a short duration of time and thus could not be used to make general inferences about the population of students taking introductory statistics courses, assessment of student learning at the end of the course suggests that the teaching pedagogy implemented in this study did help students improve their reasoning about the stochastic and might deserve further investigation. The simultaneous focus of this model on variation and on the process of statistical investigation seems to be a promising alternative to more conventional instruction, where the linear and consecutive structure of the course comes in sharp contrast with the complex nature of stochastical knowledge.

There is a lot still to be learned regarding students' reasoning about variation. More research needs to be carried out to investigate intuitions about variation of students of different age groups and different backgrounds. Through conducting this study, we have come to realize that assessment of thinking about variation is heavily reliant upon both the types of assessment tasks employed and the context in which the tasks are situated. Students come to a situation with a wide range of skills and knowledge and offer responses that are difficult to anticipate (Cohen and Checile, 1997). We, similarly to Pfannkuch and Brown (1996), documented students' neglect of probabilistic thinking when interpreting certain real world phenomena and their tendency to come up with causal explanations for short-term fluctuations in data that could be easily explained by natural variation. Conversely, we found that in certain occasions a students' response might be erroneous due to poor knowledge not of statistical content but of the context of the situation, or due to misunderstandings about what the question is asking (Jolliffe, 1994). The wealth of information that emerged out of this study shows the advantages of using a variety of assessment tasks when investigating students' reasoning about variation. To get a more complete profile of student intuitions, future research as well as instruction on variation ought to also use a variety of assessment tasks and multiple-forms of assessment that complement each other. Additionally, the instructional materials and assessment items used by educators and researchers need to be opened up for scrutiny so that they can be gradually improved by the statistics education community.

Interested readers could, by contacting us, be provided with all the teaching materials and assessment tasks employed in the study. In addition, by keeping all the data collected in a well-organized and retrievable form, we can easily make them available to any researchers challenging the findings and seeking to reanalyze the data.

APPENDIX

Gummy Bears Question

Suppose you took your little nephew on an Easter parade. At the parade, the "Easter Bunny" handed out packets of Gummy Bears to all of the students. Each packet had 6 Gummy Bears in it. To make up the packets, the Easter Bunny took 2 million green Gummy Bears and 1 million red Gummy Bears, put them in a very big barrel and mixed them up from night until morning. Then he spent the next few hours making up the packets of six Gummy Bears. He did this by grabbing a handful of Gummy Bears and filling as many packets as he could. Then he reached into the barrel and took another handful, and so on, until all the packets were filled with 6 Gummy Bears.

a) When you get home from the parade, you open up your packet. How many green Gummy Bears do you think might be in your packet? Can you explain how you got that?
b) Do you think all the students got *n* greens, where *n* is the number of Gummy Bears you gave in part (a)? Can you explain why?
c) If you could look at the packets of 100 students, how many students do you think got n greens?
d) Remember that the Easter Bunny was starting with 2 million greens and one million reds. Did he run out of one color long before the other when he was filling the bags or did they both last until near the end? Why?

College Students' Interview Question

Circle the best answer to the following problem:

At a nearby college, half the students are women and half are men. A worker for a student organization wants to interview students on their views about recent changes in the federal government's funding of financial aid. The worker wants to get a good representation of the students, and goes to as many different areas on campus as possible. Three or four students are interviewed at each place the worker visits. Out of the last 20 students interviewed, 13 were women and 7 were men. Now, you do not know what time of day it is, to which part of campus the worker has already gone, or where the worker is going next. Out of the next 20 students the worker interviews, do you think more will be women or men?

a) The worker seems to interview more women than men. There could be several reasons for this. Perhaps women are more willing to talk about their opinions. Or, maybe the worker goes to areas of campus where there are more women than men. Either way, the worker is likely to interview more women than men out of the next 20 students.
b) Since half of the students on this campus are men and half are women, you would expect a 50/50 split between the number of men and women the worker interviewed. Since there tended to be more women than men so far, I expect the opposite trend to start happening. Out of the next 20 students the worker interviews, there will probably be more men than women so that things start to balance out.
a) Half the students on this campus are men and half are women. That means that the worker has a 50/50 chance of interviewing a man or a women. It should not matter how many men or women the worker has interviewed so far. Out of the next 20 students interviewed, about half should be men and half women.
b) So far, the trend seems to be more women to be interviewed than men. Out of the next 20 students the worker interviews, I would expect the same thing to happen. The worker will probably interview more women than men out of the next 20 students.

Birth Defects Question

Every year in New Zealand approximately seven children are born with a limb missing. Last year the children born with this abnormality were located in New Zealand as shown on the map.

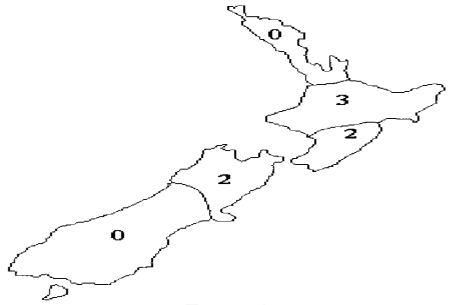

Figure 1

What do you think? (In New Zealand, it is common knowledge that one-third of the population lives in the top region and one-sixth of the population in each of the other regions.)

Die Toss Question
A fair die is tossed 7 times resulting in the outcome 3,3,3,4,4,5,5 (order is unimportant). What do you think of these results?

Child Psychologist Question
A child psychologist is engaged in studying which of two toys infants will prefer to play with. Of the first five infants studied, four have shown a preference for this toy. The psychologist concludes that most infants will show a preference for this toy. Do you think the psychologist has drawn a valid conclusion?

Candies Question
A bowl has 100 wrapped hard candies in it. 20 are yellow, 50 are red, and 30 are blue. They are well mixed up in the bowl. Jenny pulls out a handful of 10 candies, counts the number of reds, and tells her teacher. The teacher writes the number of red candies on a list. Then, Jenny puts the candies back into the bowl, and mixes them all up again. Four of Jenny's classmates, Jack, Julie, Jason, and Jerry do the same thing. They each pick ten candies, count the reds, and the teacher writes down the number of reds. Then they put the candies back and mix them up again each time.

I think the teacher's list for the number of reds is most likely to be (please circle one):
a) 8,9,7,10,9
b) 3,7,5,8,5
c) 5,5,5,5,5
d) 2,4,3,4,3
e) 3,0,9,2,8

Explain your reasoning.

REFERENCES

Ballman, K. (1997). Greater emphasis on variation in an introductory statistics course. *Journal of Statistics Education, 5*(2).

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.

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.

Biehler, R. (1999). Discussion: Learning to think statistically and to cope with variation. *International Statistical Review, 67*(3), 259-262.

Borovcnik, M. & Peard, R. (1996). Probability. In A.J. Bishop (Ed.), *international handbook of mathematics education* (pp. 239-287). Netherlands: Kluwer Academic Publishers.

Chance, B. L. & Rossman, A. J. (2001). Sequencing topics in introductory statistics: A debate on what to teach when. *The American Statistician*, *55* (2), 140-144.

Cobb, G. (1992). Teaching statistics. In L.A. Steen (Ed.), *Heeding the call for change: suggestions for curricular action* [MAA Notes No. 22] (pp. 3-33). Washington, DC: The Mathematical Association of America.

Cohen, S. & Chechile, R. A. (1997). Overview of ConStatS and the ConStatS assessment. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 110-119). Voorburg, The Netherlands: International Statistical Institute.

Confrey, J. (1996, April). Strengthening elementary education through a splitting approach as preparation for reform algebra. Presented at the *annual meeting of the American Educational Research Association*, New York, NY.

Erickson, T. (2000). *Data in depth. exploring mathematics with Fathom*. Emeryville, CA: Key Curriculum Press.

Gal, I. & Garfield, J. (1997). Curricular goals and assessment challenges in statistics education. In I. Gal and J. B. Garfield (Eds.), *The assessment challenge in statistics education.* Burke, VA: IOS Press.

Garfield, J. & delMas, R. (1990). Exploring the stability of students' conceptions of probability. In J. B. Garfield (Ed.), *Research papers from the Third International Conference on Teaching Statistics.* University of Otago, Dunedin, New Zealand.

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: The International Study Group for Research on Learning Probability and Statistics.

Garfield, J. B. (1995). How students learn statistics. *International Statistical Review, 63*(1), 25-34.

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.

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.

Hoerl, R., Hahn, G. & Doganaksoy, N. (1997). Discussion: Let's stop squandering our most strategic weapon. *International Statistical Review, 65*(2), 147-153.

Hogg, R.V. (1992). Towards lean and lively courses in statistics. In F. Gordon and S. Gordon (Eds.), *Statistics in the twenty-first century* [MAA Notes No. 26] (pp. 3-13). Washington, DC: Mathematical Association of America.

Jolliffe, F. (1994). Why ask why? In J. Garfield (Ed.), *Research papers from the Fourth International Conference on Teaching Statistics*. Minneapolis: The International Study Group for Research on Learning Probability and Statistics.

Lee, C. (1997). Promoting active learning in an introductory statistics course using the PACE strategy. *Proceedings of the 6th International Symposium on Mathematics Education* (pp. 199-206). Mexico City.

Lee, C. (2000). Computer-assisted approach for teaching statistical concepts. *Journal of Computers in the Schools, 16* (1), 193-208.

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.* San Antonio, Texas.

Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review, 65*(2), 123-165.

Pfannkuch, M. (1997). Statistical thinking: One statistician's perspective. In J. Garfield & J. Truran (Eds.), *Research papers on stochastics education* (pp. 171-178).

Pfannkuch, M. and Brown, C. M. (1996). Building on and challenging students' intuitions about probability: Can we improve undergraduate learning? *Journal of Statistics Education, 4*(1).

Rossman, A. J. (1996). *Workshop statistics: Discovery with data.* New York: Springer-Verlag.

Rubin, A., Bruce, B. & Tenney, Y. (1990). Learning about sampling: Trouble at the core of statistics. In J. B. Garfield (Ed.), *Research papers from the Third International Conference on Teaching Statistics.* University of Otago, Dunedin, New Zealand.

Shaughnessy, J. M., Watson, J., Moritz, J. & Reading, C. (1999). School mathematics students' acknowledgment of statistical variation. For the NCTM Research Precession Symposium: *There's More to Life than Centers*. Paper presented at the 77th Annual NCTM Conference, San Francisco, California.

Smith, T. M. F. (1999). Discussion. *International Statistical Review, 67*(3), 248-250.
Snee, R. D. (1999). Discussion: Development and use of statistical thinking: A new era. *International Statistical Review*, *67*(3), 255-258.
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.
Tversky, A. & Gilovich, T. (1989). The cold facts about the "hot hand" in basketball. *Chance, 2(1),* 16-21.
Wild, C. J. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223-265.
Wilensky, U. (1993). *Connected mathematics - Building concrete relationships with mathematical knowledge.* Ph.D. Thesis, Massachusetts Institute of Technology.

MARIA MELETIOU-MAVROTHERIS
11 Achilleos St.
2370 Ayios Dometios
Nicosia, Cyprus

NEW APPROACHES TO GATHERING DATA ON STUDENT LEARNING FOR RESEARCH IN STATISTICS EDUCATION[1]

BETH L. CHANCE
Cal Poly – San Luis Obispo
bchance@calpoly.edu

JOAN B. GARFIELD
University of Minnesota
jbg@umn.edu

SUMMARY

Over the last fifteen years there has been a strong emphasis on active learning, use of real data in the classroom, and innovative uses of technology for helping students learn statistics. A recent survey in the United States (Garfield, 2001) documents that many tertiary teachers of statistics courses have made changes toward these recommendations. Now more than ever, more research is needed on the effects of these instructional methods and materials on student learning, retention, and motivation. This research need first requires the determination of effective research methodology in statistics education. In assessing students' conceptual understanding, reasoning abilities, and attitudes, and their development, alternative methods of gathering student data are needed that supplement comparative experiments and improve on traditional assessment items that focus on calculations, definition, and rote manipulations. This article will present and critique additional methods for obtaining research data on how students develop an understanding of statistics, including classroom-based research and videotaped student interviews/observations.

Keywords: Statistics education research; Assessment; Classroom-based research; Clinical interviews

1. INTRODUCTION

During the last fifteen years, educators have witnessed a movement in statistics education aimed at shifting the focus of instruction away from theory and recipes toward statistical thinking, genuine data, conceptual understanding, and active learning. Much of this movement has been motivated by research in educational psychology, psychology, mathematics education, and science education (see Garfield, 1995). However, there is still need for documented evidence of whether such changes enhance student learning, retention, and appreciation of statistics. This evidence is also needed to determine the most effective instructional techniques and to develop models of how students shape their statistical understanding. Without more published research, we will not be able to continue to move these changes forward and expand their impact.

While there is much overlap with research questions in mathematics and science education, statistics education poses a unique set of challenges. For example, problem *context* plays a role in statistics that is not paralleled in mathematics (Cobb and Moore, 1997). Furthermore, probabilistic reasoning and randomness appear to require distinct teaching and learning strategies (see e.g., Falk and Konold, 1992). While there has been a strong increase of activity and publications in statistics education in recent years (e.g., *Journal of Statistics Education*, International Conference on Teaching Statistics (ICOTS), Statistical Education Research Section of the International Association for Statistical Education (IASE)), more discussion and reflection are needed to clarify "what should be considered as research in statistics education, how we establish the validity of our research findings, what priority questions need to be studied, and what theoretical frameworks and research methods might be recommended to carry out this research" (Batanero, et al., 2000). These questions need to be answered

[1] *Statistics Education Research Journal, 1(2), 38-41, http:/fehps.une.edu.au/serj*
International Association for Statistical Education

in order to establish visibility, legitimacy, and understanding of research results, as well as to form a basis for on-going research and training of future researchers. In order to advance the field of statistics education and achieve academic recognition, it is essential to have a well-developed research literature and research agenda (Batanero, et al., 2000). Moreover, little is known about, or has been published on, the methodology of statistics education research (Jolliffe, 1998). Statistics education research will not be able to sufficiently impact policy or the practice of teaching until individual studies become grounded in a broader program of study.

2. CURRENT TRENDS IN STATISTICS EDUCATION RESARCH

The teaching of statistics is unique in that instructors typically come from a variety of backgrounds and fields of application. Similarly, statistics education research has been conducted by a variety of individuals who represent different disciplines, educational programs and training in research methods. Research in psychology has revealed ways people reason about statistical or probabilistic information, paying particular attention to faulty reasoning and misconceptions (e.g., Garfield, 1998). However, this research has not been expanded to all age groups or closely tied to teaching practice. Similarly, studies conducted by mathematics education researchers have focused on how children reason about particular topics such as averages or graphs. There has also been work on how K-12 teachers understand and reason about statistics. Recent research efforts in statistics education have focused on comparisons of two types of instruction (e.g., laboratory environment vs. traditional lecture) or prediction of achievement based on mathematical ability, attitudes, and other variables. Attention has also been directed at evaluating student ratings of new implementations. While all these types of research studies offer interesting results, they are often limited in their generalizability and validity. This focus needs to be expanded (to other types of questions, across all age groups, towards research on teachers of statistics) and to tie the research more closely to classroom practice. In particular, little of this research has focused specifically on statistical reasoning or other issues unique to statistics education.

A concern regarding many of these studies is that they lack visibility as well as cross-disciplinary or cross-institutional collaboration. While there are some large conferences, such as the ICOTS series, there appears to be little connection among researchers in the years between these conferences. In an effort to coordinate these research efforts, there are several research study groups aimed at statistics education including the Statistical Reasoning, Thinking and Literacy International Research Forum, and the Statistics Education Research Group based at the University of Massachusetts at Amherst, as well as the RSS Centre for Statistical Education. While these organizations have been conducting and supporting statistics education research, they are relatively new (so support is not extensive) and they have not been as focused on establishing standards of research, or exploring effective methods. It can also be difficult for a new researcher to find an obvious focal point of these efforts.

Consequently, the need to explore the role and future of statistics education research has recently become a theme of sessions at international conferences. For example, a session of the Sixth ICOTS was held in July 2002 was entitled “Research in Statistics Education” and included talks on developing statistics education research; theoretical models of statistical knowledge, thinking, reasoning, and learning; and a roundtable discussion of major problems and directions in statistics education research. Thus, discussions are beginning on how to legitimize statistics education as a research domain and how to train future researchers in statistics education.

We see the pressing needs as falling into two categories. One is documenting evidence of the effects on students of these instructional changes, trying to identify the most effective instructional techniques, while also developing models of how students come to understand statistics which will help foster additional reform and be closely tied to practice. The other is generating more discussion and reflection on acceptable research methods and a research agenda. If we hope to establish the validity and legitimacy of statistics education as an area of research then we need a well-developed research literature that we can point people to, as well as accountability in our methods. These discussions will also help inform the training of future researchers.

3. LIMITATIONS OF CURRENT RESEARCH ACTIVITIES

As statistics education research becomes a more visible discipline, it is vital to consider appropriate research techniques. When formulating a research methodology, we must meet several criteria. The research must be valid so that the data correspond to what the research purports to measure and so that the information can be generalized beyond the study at hand. The knowledge gained must be reliable, consistent, and replicable, while providing sufficient documentation of the activities and observations involved. There must be objectivity and belief that the research process is not affecting the outcomes being measured and that the evaluation is fair and equitable.

In designing studies to meet these criteria, the obvious place to start is with the gold standard, the randomized comparative experiment. For example in clinical trials, we can impose a treatment, control the environment, and draw clear cause and effect conclusions. However, when trying to apply this to the education world, things become more complicated. For example, the first step – randomization – is not really possible in the educational setting, especially if we are looking for long-term, even semester long, effects. It is not feasible for us to manipulate which courses students will take and self-selection into sections is not a sufficient substitute for randomization. Student mobility between sections and drop rates are also serious difficulties. Thus, we will not be able to maintain the independence of observations that is required by our traditional statistical techniques. Variables interacting with the instructional environment such as instructor attitude, time of day, resources available, and classroom culture may have dramatic effects on student achievement and attitude and cannot be controlled or even measured as in a laboratory setting.

It also may not be feasible to create the classroom implementation that is of interest. For example, in deciding what factors most directly affect student achievement, to be an authentic reflection of student performance, the assessment most likely needs to be tied to a grade and there are ethical issues involved in manipulating variables that could adversely affect student grades. In the standard research model, we can try to design a more controlled environment but this leads to separation of the research from classroom practice and to significant time delays before the research can be applied to the classroom environment. Often the results from a controlled laboratory setting are not immediately relevant.

Furthermore, this type of research too often ignores the integral role played by the classroom instructor. Trying to achieve objectivity through an outside observer ignores that the presence and demeanour of the observer can also affect the results. An external researcher may not sufficiently understand the details of the classroom environment and culture, or even have the subject matter knowledge to adequately appreciate and document students' experiences. Directly involving the classroom participants (teachers, curriculum designers, students) in the research and immersing the researcher into the classroom environment allows for more in-depth study, more reflection, and better interplay between theory and practice.

The above is not an exhaustive discussion of the issues involved but does begin to support the view that the classroom environment is a sufficiently complex and dynamic world that is not always well described by traditional research techniques. Similarly, many traditional measurement techniques, such as standardized exams, final course grades, and student ratings, are not sufficient, especially when trying to measure student reasoning. Studies have shown that students can do well on final exams but then still demonstrate poor statistical reasoning on other tasks. As Lesh and Lovitts (2000) caution, "Most existing high-impact standardized tests are poorly aligned with national standards for instruction and assessment." Current assessment methods typically are not dynamic in nature and fail to inform the researcher of the learning processes involved. One-time measures of achievement also fail to explore the developmental nature of learning or provide concurrent feedback to the study, nor allow an iterative research approach. In particular, traditional assessment strategies do not tell us enough as to why a particular teaching method or activity works, how students' understanding and reasoning are affected or unaffected by the learning experience, nor provide direction for how teaching practice should be changed. Advances in technology also enable new methods of measurement and the collection of different types of student data, from tracking student focus on a computer screen to extensive videotaping and internet-based records. These new techniques allow movement beyond multiple choice, standardized paper-and-pencil exams. This movement is crucial in order to expand the types of information we obtain and the types of research questions we can ask.

Thus, effective research requires a careful combination of appropriate research methods and assessment techniques. Alternative approaches include classroom-based research, teaching experiments, naturalistic observation, and videotaped interviews. These techniques have been used in other disciplines such as social science and anthropology, and a recent handbook (Kelly & Lesh, 2000) acquaints mathematics and science education researchers with these approaches. However, many statistics education researchers do not have background in these different areas. A statistician in particular is typically schooled in controlled experiments and may not be familiar with theories of learning, assessment, or education.

4. NEW DIRECTIONS

While the standards of research are constant, how they are met is not. While traditional research methods also aim for "hard data" that allow cause-and-effect conclusions, alternative methods of research aim for systematic, scientific investigations from which inferences can be drawn. More and more investigators are replacing purely statistical procedures with the collection of rich, diverse data from multiple sources that document the situation being investigated and provide a scholarly account of the situation and/or the intervention.

For example, Moschkovich and Brenner (2000) outline how the above research standards are met in naturalistic research: Validity can be obtained through prolonged investigation, immersion into the environment, triangulation of multiple data sources, and frequent member checks of interpretation. Generalizability can be achieved through extensive description of the classroom situation being investigated and multi-site designs. Such documentation allows the reader to determine if the results are relevant to their environment, shifting the responsibility from the researcher to the reader. Reliability is enhanced by combining efforts and perspectives of multiple researchers and by direct involvement of participants in the research program. Objectivity can be assessed through extensive documentation and dissemination. However, naturalistic researchers also admit that pure objectivity may not be feasible or even desirable and instead aim for acknowledged and controlled subjectivity through researcher-participant immersion, while being explicit about how prior assumptions and beliefs could be influencing the research observations.

Classroom-based research, or Action research, has been defined as "ongoing and cumulative intellectual inquiry by classroom teachers into the nature of teaching and learning in their own classrooms" (Cross and Steadman, 1996; Feldman & Minstrell, 2000). Thus, instead of ignoring the integral role the teacher plays in the learning process, the classroom teacher becomes a key partner in the research team, helping to develop the questions to be investigated and assisting in data collection. By directly involving the teachers in the research process, classroom research aims to incorporate their perspectives, insight, and understanding of the classroom culture into the analysis. This approach also allows for further probing into the student and instructor experience, and adjustments in the evaluation process can be immediately implemented. The gap between theory and practice is narrowed and the evaluation becomes a dynamic process that changes in response to results and feedback, while simultaneously focusing on curricular development, instruction, and assessment. Below we offer an example of a classroom based research project we have been conducting to investigate how interaction with a conceptually based interactive software program helps students develop statistical reasoning skills.

Videotaped clinical interviews build on techniques used by cognitive psychologists. These methods are used to study the form of knowledge structures and reasoning processes (Clement, 2000). Researchers in statistics education are utilizing these techniques in studies that explore student understanding of data, relationships, tendency, and inference. These studies have been helpful in generating models, and allow for independent viewers/ coders to compare their interpretations and see if there is convergent validity in their findings. For example, participants at the International Research Forum on Statistical Reasoning, Thinking, and Literacy bring videotapes with them to share with colleagues in small working groups to discuss and validate their results. (See http://www.beeri.org.il/srtl/.)

5. EXAMPLE

We now provide an example of a classroom research project used to investigate how student interaction with a simulation program affects their statistical reasoning. Our goals were to understand student thinking and to inform other instructors about the use of such a simulation program to teach the topic of sampling distributions. This is a notoriously difficult topic for statistics students to understand, but is also the gateway to understanding statistical inference. We hoped to provide insight into *how* to best integrate technology into instruction, *why* particular implementations appeared to be more effective, and *how* student understanding evolved through use of the program. Thus we were more focused on understanding students' knowledge structure and reasoning process than on establishing a simple cause and effect relationship.

We began gathering data in three diverse college settings: an introductory statistics course for non-traditional students, an introductory statistics course for business and science majors, and a graduate level course in education. This allowed us to work with students from a wide variety of programs of study, educational backgrounds, and ages. As researchers, we also had a wide variety of backgrounds from cognitive science to educational psychology to statistics. We were able to create the learning environment that we wanted to investigate and fully integrate this into the existing course (students used a dynamic, interactive computer program developed by delMas, see delMas, 2001). This allowed us to continually relate what we were observing in the classroom to existing theory while also generating new models of student learning.

We have used a wide variety of measurement techniques to try to capture student reasoning (Chance, Garfield, & delMas, in press). In the first stage of the research, we utilized graphics-based test items to determine whether students could demonstrate a visual understanding of the implications of the Central Limit Theorem by choosing the appropriate pictures that corresponded to an empirical sampling distribution for different sample sizes. Initially, students were asked to justify their choice of graphs and explain their reasoning. These responses were then categorized so that future instruments asked students to select which statement best represented their own reasoning. Students were given these test instruments before using the program and after using the program the next day in order to isolate the change in understanding from interacting with the program. We also developed some open-ended questions where students had to provide their own justification and reasoning for their answers and some post-test application problems that could be given later in the course to see if they could still apply and use their knowledge.

In an effort to improve student performance further, we incorporated a model of conceptual change. Students were asked to select their responses and then use the program to check their predictions. This forced students to more directly confront the misconceptions in their understanding and improvements in their performance on the post-test were significant (delMas, Garfield, and Chance, 1999). This led to additional investigations in the role of pre-requisite knowledge, development of tools for identifying prevalent misconceptions, and further refinements to the activity.

Currently, our research is focused on formulating a model of student development of statistical reasoning. Through videotape analysis we are documenting student explanations of their reasoning using students at several levels of development to validate this model. For example, this has allowed us to further document student choices, and how correct choices are often still accompanied by faulty reasoning. These methods have allowed us to gain much more in depth understanding of individual students and what they know.

This brief summary showcases how collaborative, classroom-based research can be used in statistics education. Does this type of research meet the standards we discussed before? We feel that through prolonged investigation and immersion into the learning environment and by gathering multiple sources of data and looking at the consistency in information from these multiple sources, then we do have validity in what we are measuring. We don’t claim to have used all instructional settings but that they were diverse enough that something that works in these three situations has a good chance of working in other settings as well. By extensively describing our situation then, the readers can decide if the results are relevant to their environment. We also feel it is important to have our multiple perspectives to be able to check either other’s observations in what we are seeing in these data as well as to directly involve the teacher in the study (instead of the top-down model where the teacher is told to try something based on theory where they may not have the belief or skills). Still, it is important to have the non-participant viewpoint as well and through documentation and public dissemination we allow others to evaluate whether we were able to achieve that outsider perspective.

We also see additional benefits. This approach allows the classroom results to immediately provide feedback to the theory. It also provides direct access to the students and instructor, focusing more on the process than the end result, which allows for more probing and follow-up. It is also a dynamic process that allows for immediate adjustment in the research process and creates a much more iterative approach. This allows the student response to drive the investigation more than our prior beliefs. We feel we have learned a lot more about students’ understanding than we would have ever anticipated.

Using our own students as “subjects” allowed direct access to the students and the ability to specifically create the desired learning environment, as well as additional insight into the students’ experiences. We have used our diverse perspectives in the development of the learning environments and the cross-checking of interpretations and evaluations. The investigations were tied to existing theory and are helping to generate new theory. The project has led to the development of new assessment instruments that are now available to other instructors. While no part of our research could be considered a traditional experiment, we feel we have contributed insight into why an activity works, have demonstrated transferability in the learning gains to other instructional settings, and are continuing to employ a variety of research methods to evaluate the progress of this project.

6. SUMMARY

This paper has examined new approaches to gathering data on student learning. Many of these techniques can be applied equally well to statistics education research. Perhaps the two most important lessons to take from this discussion are that the research methodology must match the research question and that a variety of tools should be employed: "Different techniques generate different types of information, and it is often the case that a single technique will not provide the breadth of information necessary to answer unequivocally the research questions under investigation" (Mestre, 2000). These techniques should combine qualitative and quantitative data: "There is no single correct approach to evaluation problems. The message is this: some will need a quantitative approach; some will need a qualitative approach; probably most will benefit from a combination of the two" (Herman, Morris & Fitz-Gibbon, 1987). We are not trying to say one should never do experiments in an educational setting. There are certainly (very focused) questions and situations where such methods are appropriate. Instead we want to advocate using a variety of tools, using different techniques to answer different kinds of questions, gathering both qualitative and quantitative data. Even with clinical trials, pharmaceutical companies spend millions of dollars on research before the clinical trial stage. Relaxing the strict adherence to classical experimental methods will allow richer sources of information through complementary techniques and new research questions. Statistics education research in particular is still developing a set of coherent research questions. By gathering information from a variety of sources we will be able to develop more informed research questions. This development will be greatly aided through additional naturalistic observation and documentation of students prior to more systematic investigations.

The gathering of data is closely connected to theoretical frameworks, prior research, and type of methods and design used. In considering ways to improve the information gathered in educational research studies in statistics, we need to keep in mind the context of the growing and developing discipline of statistics education. Toward this end, statistics education needs to establish standards for preparing researchers in statistics education. This includes recommended coursework in statistics, education, learning theory, measurement, and qualitative and quantitative research methods. Clearly, future researchers in statistics education need to have training and cross-disciplinary collaboration in psychology, education, math and science education, and alternative research methodologies. The body of statistics education research also needs to become more visible and accessible across disciplines. This should include publication of examples of high quality research, literature reviews, bibliographies on certain topics, and a research handbook specific to statistics education research. Such tools will enable researchers to combine traditional methods with alternative approaches in order to best answer a wider array of research questions.

ACKNOWLEDGEMENTS

This discussion was originally presented at the 53rd Session of the International Statistical Institute, Seoul Korea, August, 2001. A shorter version (Chance and Garfield 2001) appears in the *Proceedings of the 53rd Session of the International Statistical Institute*, 2001.

REFERENCES

Batanero, C., Garfield, J., Ottaviani, M. G., & Truran, J. (2000). Research in statistical education: Some priority questions. *Statistical Education Research Newsletter*, 1(2), 2-6.

Chance, B. & Garfield, J. (2001). New approaches to gathering data on student learning for research in statistics education. *Proceedings of the 53rd Session of the International Statistical Institute. Bulletin of the International Statistical Institute* (Tome LIX, Book 2, pp. 205-208). Seoul, Korea: International Statistical Institute.

Chance, B., Garfield, J., & delMas, R. (In press). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.) *The challenge of developing statistical literacy, reasoning, and thinking*, The Netherlands: Kluwer Publishers.

Clement, J. (2000). Analysis of clinical interviews. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 547-589). Mahwah, NJ: Lawrence Erlbaum Associates.

Cobb, G. & Moore, D. S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly*, 104(9), 801-824.

Cross, K. P. & Steadman, M.H. (1996). *Classroom research: Implementing the scholarship of teaching*. San Francisco: Jossey-Bass Publishers.

delMas, R. (2001). *Sampling SIM* (version 5). http://www.gen.umn.edu/faculty_staff/delmas/stat_tools/

delMas, R., Garfield, J. & Chance, B. (1999). Assessing the effects of a computer microworld on statistical reasoning, *Journal of Statistical Education*, 7(3).

Falk, R. & Konold, C. (1992). The psychology of learning probability. In *Statistics for the twenty-first century* (pp. 151-164). MAA Notes, Number 26, The Mathematical Association of America.

Feldman, A. & Minstrell, J. (2000). Action research as a research methodology for the study of the teaching and learning of science. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 429-456). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Garfield, J. (1995). How students learn statistics. *International Statistical Review*, 63, 25-34.

Garfield, J. (1998). The challenge of assessing statistical reasoning. Paper presented at *American Educational Research Association Annual Meeting,* San Diego.

Garfield, J. (2001). *Evaluating the impact of educational reform in statistics: A survey of introductory statistics courses.* Final Report for NSF Grant REC-9732404. Minneapolis, MN.

Herman, J.L., Morris, L.L., & Fitz-Gibbon, C.T. (1987). *Evaluator's handbook, second edition*, Newbury Park, CA: Sage Publications.

Jolliffe, F. (1998). What is research in statistical education? In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W. K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics* (pp. 801-806). Singapore: International Statistical Institute.

Kelly, A. & Lesh, R. (Eds.), (2000). *Handbook of research design in mathematics and science education* (pp. 361-512). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lesh, R., & Lovitts, B. (2000). Research agendas: Identifying priority problems and developing useful theoretical perspectives. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 45-71). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Mestre, J. (2000). Progress in research: The interplay among theory, research questions, and measurement techniques. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 151-168). Mahwah, NJ: Lawrence Erlbaum Associates.

Moschkovich, J. N. & Brenner, M. (2000). Integrating a naturalistic paradigm into research on mathematics and science cognition and learning. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 361-512). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

BETH L. CHANCE
Department of Statistics
California Polytechnic State University
San Luis Obispo, CA 93407

UNDERGRADUATE COURSES IN DENTAL STATISTICS IN BRITAIN AND IRELAND[1]

NIGEL SMEETON
King's College London
nigel.smeeton@kcl.ac.uk

SUMMARY

Despite the recommendation of the General Dental Council that statistical methods and data analysis should form part of the curriculum of undergraduate dental degrees, little is known about the teaching of statistics in dental schools. This informal study was carried out to obtain information on the methods of teaching and assessment used in dental schools in Britain and Ireland.

***Keywords*:** Teaching; Assessment; Dentistry; Undergraduates; Survey; Statistics education research

1. INTRODUCTION

Historically, the teaching of statistics in dental courses in Britain and Ireland, if conducted at all, has had a very minor role. However, in 1990 the General Dental Council published recommendations for dental training, one of which sought to rectify this (General Dental Council, 1990, p.10):

> *'The teaching should introduce the student to the principles of scientific thought and argument including the evaluation of scientifically established facts, experimental design, statistics and the analysis of data, and place the clinical instruction in the scientific context'*.

Although dental schools have sought to address the lack of statistical teaching, apart from a paper on the teaching of statistics to dental students at the former United Medical and Dental Schools of Guy's and St. Thomas's Hospitals, London (Smeeton, 1996) little is known about the content of the British and Irish courses that have been established. In April 2001 I attended a committee meeting at which one of the dental staff made an unsubstantiated remark about the increasing importance of computer aided self-directed learning in the teaching of statistics in UK dental schools. I had no idea whether or not this remark was correct but was curious to find out the answer, hence this investigation.

2. METHOD OF INVESTIGATION

A listing of the dental schools in Britain and Ireland was obtained from the Dental Education Resources on the Web site (http://www.derweb.co.uk/main/dentsch.html). Information on the teaching of dental statistics, including timetables, lectures, exercises and assessment where possible, was obtained from dental school web sites. A request for information was sent through the international dental education e-mail group. Personal acquaintances involved in statistics teaching within medical and dental schools were contacted by e-mail. This led to a dental statistics teaching e-mail group (DENTSTAT) being established and publicized in the Royal Statistical Society (RSS) News. These sources yielded detailed information on around two-thirds of the relevant dental schools.

In addition, a series of questions was sent by e-mail to members of the organizers' circulation list for the 2002 annual conference for UK teachers of medical statistics at Burwalls, Bristol (held from 10th-12th April) asking about:

(1) The number of dental students admitted per year.
(2) The years in which dental statistics teaching takes place.
(3) Whether they are taught statistics together with medical students or not.
(4) Approximate number of hours of teaching/ revision, for lectures, tutorials and practicals separately
(5) Details of in-course assessments, if any.

[1] *Statistics Education Research Journal, 1(2), 45-48, http:/fehps.une.edu.au/serj*
International Association for Statistical Education

(6) Details of written examinations, if any.
(7) Use of statistical packages and computer aided / internet learning.

I emphasized the importance of receiving feedback from all dental schools including those institutions about which I had no information at that time. To promote interest further, I agreed to lead a session on the topic at the 2002 medical statistics teachers' meeting. Prior to giving the talk, I revisited the dental school web sites for updated information.

3. RESULTS

Detailed information about statistics teaching for the academic year 2001/02 was collected from all fourteen dental schools in Britain and Ireland (Belfast, Birmingham, Bristol, Cardiff, Dublin, Dundee, Glasgow, Leeds, Liverpool, London – King's College, London – Queen Mary's, Manchester, Newcastle and Sheffield). Apart from one telephone conversation all contact was by e-mail. The annual intake of students ranged from 35 (Dublin) to 140 (King's College, London) with a median of 60. Teaching was organized by statistics staff at 9 (64%) of the dental schools, with dental departments taking responsibility at the other institutions (Table 1).

Table 1. Specialties responsible for the teaching of statistics to dental students

Specialties	Number
Statistics	9
Clinical dental sciences	2
Oral biology	1
Physiology	1
Restorative dentistry	1

The Bachelor of Dental Science (BDS) degree entails a five-year course with an optional extra year for the in-depth study of a particular aspect of dentistry leading to a BSc degree. The additional qualification is commonly referred to as an 'intercalated' degree. Traditionally the first two years have emphasized basic dental sciences whereas the later years have focused on clinical practice. Vertical integration is now being introduced with themes such as dental public health extending throughout the whole five years. From this investigation, statistics teaching took place in years 1 and/or 2 in seven dental schools. In three schools the teaching came in the later years of the course with teaching in both early and later years at four institutions. In all but two dental schools students were taught statistics separately from medical students. Separate teaching took place at all schools where dental staff were responsible.

Table 2 shows the range of face-to-face teaching methods in use. The lecture was an important component of teaching at most dental schools. All but one of the courses involved lectures, with 10 or more hours of lectures being timetabled at four. Tutorials were arranged in four schools and in one they were the only face-to-face classroom teaching. Small group sessions were commonly based around discussion of recent articles taken from high profile dental journals.

Table 2. Range of face-to-face teaching in use

Lectures only	7
Lectures and computing sessions	3
Lectures and tutorials	1
Lectures, tutorials and computing sessions	2
Tutorials only	1

Textbooks were mentioned by staff from four institutions: Bulman and Osborn (1989) (2), Bland (2000) (2) and Dunn and Everitt (1995) (1). Two schools specifically recommended a text. The use of statistical packages was mentioned at nine schools, SPSS (5) and Minitab (3) being the most commonly used. Statistics practical classes involving computing featured in five of the courses.

Statistics revision classes, made available to all students prior to formal assessments, were offered at five dental schools, with a range from 1 to 4 hours of teaching on offer. These involved the discussion of sample examination questions, in which the tutor explained the main points with students asking for clarification as required.

Project work was set on three dental courses and two courses involved in-course assessment exercises. Four dental schools indicated the use of computer aided self-directed learning. One school used QUERCUS (McCloskey et al., 1997), another Statistics for the Terrified (Morris et al., 1997) and two others had developed in-house computer aided learning material. A further dental school classified one-to-one statistical consulting for student projects as self-directed learning. The institutional intranet featured in four courses; two used it for lecture notes, three for course exercises, three for past examination questions and one for discussion between students and staff. Only two of the four institutions provided access to outside users enabling me to check information given by staff.

Formal assessment of statistical understanding was conducted at ten of the schools. Institutions used a range of assessment methods including multiple-choice questions, short answer questions, longer structured problems and the analysis of a provided data set. Examination papers devoted to statistics were uncommon; questions usually featured as a section in multi-disciplinary papers.

A comparison of the courses organized by statisticians and those run by dental staff showed that five of the former provided revision classes. In contrast, revision classes did not feature in any of the statistics courses arranged by dental departments. The courses featuring use of the internet were all run by statistics departments.

One of the more substantial courses, at Belfast, involved four hours of lectures with a one-hour tutorial in Year 1 and eight hours of lectures with two hours of tutorials in Year 2. There were no computing sessions as such. Nevertheless, project work was a central aspect of the course; there were two projects to be completed in each of the first two years. These were designed, among other things, to encourage students to gain skills in the use of the statistical packages Minitab and SPSS. In-course assessment involved use of the local intranet, where the course notes and past examination papers could also be found (restricted access). Four one-hour revision classes were available although most students only attended one of these as they covered similar material.

4. DISCUSSION

These findings reveal that, at most establishments, use of the traditional methods of course delivery such as lectures and tutorials is common. The lack of revision classes on courses run by dental staff could cause some concern. In my experience, statistics teachers tend to view the teaching of dental students as arming the students with skills for the interpretation of dental papers and their own analysis of data from laboratory practicals. Some dental staff, on the other hand, prefer an ‘on the job’ approach to the teaching of statistics, where the data are confronted first and only then are the statistical tools provided for the analysis. There is certainly evidence of these alternative views from the e-mail comments that I received. The merits and problems of these approaches deserve frank discussion between the two sides.

Computer based activities mainly took place in statistics practical sessions using well-known packages. In answer to the question that initiated this investigation, only four courses used specific self-directed computer assisted learning tasks. However, the self-directed seeking out of ad hoc statistical advice for projects may have been underestimated if the time spent with tutors in this way went unrecorded. As far as the internet is concerned, again only around one-third of courses had begun to incorporate such material. Some of the staff indicated that intranet teaching was either in preparation or under active consideration. The development of such material is very time intensive and may require dedicated staff to be recruited for these tasks. However, some of the variation between institutions could well be due to a difference in teaching philosophies, where the development of computer based activities is regarded by some, but not all, as a high priority. It would be interesting to repeat this exercise in three or four years time to see whether these proposed advances have materialized.

Information about dental statistics teaching in other parts of the world is sparse. Harraway and Sharples (2001) give a detailed description of a biostatistics course at the University of Otago, Dunedin, New Zealand that is designed for health sciences students before they commence their professional courses. This course is based around the provision of skills for laboratory practicals, including statistical computing, and the critical appraisal of research papers. Some of these students then proceed to dental school.

The discussion following my conference talk considered the benefits and pitfalls of increasing the proportion of computer-based teaching and the effectiveness of 'on the job' statistical teaching. There was also some debate on the role of statistics staff in the teaching of dental statistics. In the U.K. there is increasing pressure on academic staff to concentrate on their research output at the expense of teaching. The impact of this has not always been negative; at Guy's Dental School (now part of King's College London) an increase in the involvement of clinically trained dental staff as tutors has enabled statistics to be seen as more relevant by students. However, at some dental schools there has been a total transfer of responsibility for dental statistics teaching from statistics to dental staff. Reassuringly, in general the previous organizers remain available to advise those who have taken on this task.

Inter-professional teaching, along the lines developed by Harraway and Sharples (2001) has not as yet received serious consideration in the U.K. Some academics believe strongly that the teaching of statistics should be delivered with the future professional roles of the students in mind, on the basis that this will make courses more relevant and interesting. However, with the constraints on staff teaching time, inter-professional education may need to be examined carefully. The greater efficiency in the use of teaching resources would result in larger student groups but this would be preferable to relinquishing the overall responsibility for teaching statistics altogether.

The alternatives for the future can only be examined meaningfully in the light of findings from carefully conducted research into teaching effectiveness. Informal surveys based on personal contacts, such as the one described here, run the risk of being incomplete and biased. An increase in adequately funded formal collaboration between academic institutions is required.

ACKNOWLEDGEMENTS

I am grateful to all those who supplied information about dental statistics teaching, without whom this report would not have been possible, especially Mike Stevenson, Queen's University Belfast, who explained the Belfast course to me in detail.

REFERENCES

Bland, M. (2000). *An introduction to medical statistics*. 3rd Edition. Oxford: Oxford Medical Publications.

Bulman, J.S. & Osborn, J.F. (1989). *Statistics in dentistry*. London: BDA.

Dunn, G. & Everitt, B. (1995). *Clinical biostatistics*. London: Arnold.

General Dental Council. (1990). *Recommendations concerning the dental curriculum.* London: General Dental Council.

Harraway, J. A. & Sharples, K. J. (2001). A first course in biostatistics for health sciences students. *International Journal of Mathematical Education in Science and Technology 32*, 873-886.

McCloskey, M. M., Blythe, S. & Robertson, C. (1997). *QUERCUS: Statistics for bioscientists*. London: Arnold.

Morris, S., Szuscikiewicz, J. & Preston, M. (1997). *Statistics for the terrified.* Abingdon: Radcliffe.

Smeeton, N. (1996). Open wide: teaching statistics to dental students. *The Statistician 45,* 517-520.

NIGEL SMEETON
Department of Public Health Sciences
King's College London
Guy's, King's and St. Thomas' Hospital Medical School
Fifth Floor, Capital House
42 Weston Street
London SE1 3QD
UK

SOME BASIC REFERENCES FOR THE TEACHING OF UNDERGRADUATE STATISTICS[1]

PETER HOLMES
RSS Centre for Statistical Education Nottingham Trent University
Peter_Holmes@btinternet.com

***Note**: This bibliography has first been published in MSOR Connections Vol1 No. 4, November 2001, page 6 and is now reproduced in this journal with permission of the MSOR editors. The full list of references is available from the web page: http://www.ltsn.gla.ac.uk/articles/teachingresources.asp*

1. INTRODUCTION

Over the past 25 years or so there has been a growing interest and amount of research work into the teaching of probability and statistics. This interest and research has been reflected in the five International Conferences on Teaching Statistics, the establishment of journals such as *Teaching Statistics* and the *Journal for Statistics Education* as well as an increasing number of articles in other journals and papers at other conferences. Initially the emphasis was on school pupils but, increasingly, there has been an emphasis on teaching undergraduates.

In their bibliography, Sahai, et al (1996) list 2367 references up until the year 1994. With so much published work it is difficult for newcomers to the field to know where to start. The following list of basic references attempts to pull together the various strands of research about undergraduate teaching so that new lecturers will be able to get a quick overview of current thinking and where it has come from. The many older references are to give an historical context and reflect the influences on today's practice.

As in all such summary bibliographies there is a lot of subjectivity in the choice of what to include. It was difficult to decide whether or not to include textbooks. In the end I decided to include a few that had been particularly influential on the way statistics is taught at undergraduate level. I have not included any of the very interesting references that are specific to the school level because this would have made what was intended to be a short list even longer than it has become. The list has been circulated amongst a lot of people working in the field of statistical education and I have benefited from their advice. In the final analysis, though, the final decision was mine and any errors and omissions are mine. I would welcome correspondence about any important contributions that are missing and any references that I have included that you think should not be.

2. CONFERENCES

2.1. ICOTS-1

References included in D. R. Grey, P. Holmes, V. Barnett, & G. M. Constable (Eds.) (1982), *Proceedings of the First International Conference on Teaching Statistics*. Sheffield, U.K.*:* Teaching Statistics Trust. *Particularly:*

Barnett, V. Why teach statistics, 3 – 15.

Ehrenberg, A. S. C. We must preach what is practised, 215 – 218.

Joiner, B. L. The case for using computers in teaching statistics, 307 – 312.

Snee, R. D. Cooperation between University and industry statisticians in the United States, 491 – 502.

Mead, R. Teaching experimental design in the computer age, 591-602.

Shauhgnessy, J. M. Misconceptions or probability, systematic and otherwise; teaching probability and statistics so as to overcome some misconceptions, 784 – 801.

[1] *Statistics Education Research Journal, 1(2), 49-53, http:/fehps.une.edu.au/serj*
International Association for Statistical Education

2.2. ICOTS-2

References included in R. Davidson & J. Swift (Eds.) (1986), *Proceedings of the Second International Conference on Teaching Statistics,* Victoria, Canada: University of Victoria Conference Services. *Particularly:*

Speed, T. Questions, answers and statistics, 18 – 28.

Barnett, V. Statistical consultancy, a basis for teaching and research, 303 – 307.

Anderson, C. W., & Loynes, R. M. University statistics – what are we trying to teach and how? 317 – 321.

Vännman, K. Statistics in industry and implications for teaching, 355 – 356.

Hunter, J. S. Bayesian approach to teaching engineering statistics, 380 – 384.

2.3. ICOTS-3

References included in D. Vere-Jones (Ed.) (1991), *Proceedings of the Third International Conference on Teaching Statistics*. Voorburg, The Netherlands: International Statistical Institute. *Particularly:*

Garfield, J. B. & delMas, R. Students' conceptions of probability, 1, 340 – 349.

Jolliffe, F. The loss of intuition – a lesson for the teacher, 1, 350 – 356.

Konold, C., Pollatsek, A., Well, A., Hendrickson, J. & Lipson, A. The origin of inconsistencies in probabilistic reasoning of novices, 1, 357 – 362.

McConway, K. Sixteen years of statistics at a distance, 1, 390 – 399.

Martin, P. Teaching introductory statistics at tertiary level: a shift in emphasis, 2, 56 – 60.

Oosthuizen, J. H. Teaching and learning statistics, 2, 61 – 66.

McGillivray, H. Teaching large classes of science and engineering students, 2, 67 – 72.

Lock, R. H. Alternative introductions to applied statistics for mathematics students, 2, 91 – 94.

Lock, R. H. Some favourite data sets: using the computer on real data in class, 2, 222 – 227.

Vännman, K. Encouraging the engineering student to feel the importance of statistics – some ideas, 2, 245 – 254.

Schuenemeyer, J. H. Training statisticians to be consultants, 2, 439 – 445.

2.4. ICOTS-4

References included in Y. Escoufier & A. El-Ghazali (Eds.) (1994), *Proceedings of the Fourth International Conference on Teaching Statistics*, Vols. I and II. Voorburg, The Netherlands: International Statistical Institute. These proceedings were produced *before* the conference and only a few entries have the full presentation; most have only the abstracts. Full reports are given separately for the research papers in J. B. Garfield (Ed.), (1994), *Research Papers from the Fourth International Conference on Teaching Statistics*, Minneapolis, MN, University of Minnesota.

2.5. ICOTS-5

References included in L. Pereira Mendoza (Chief Editor) (1998), *Proceedings of the Fifth International Conference on Teaching of Statistics* Volumes 1, 2, and 3. Voorburg, The Netherlands: International Statistical Institute. *Particularly:*

Scheaffer, R. L. Statistics education – bridging the gaps among school, college and the workplace, 1, 19 – 26.

Section on Statistical Education at Post-Secondary Level, 1, 121 – 280.

Section on Research on Teaching Students at Post-Secondary Levels, 2, 693 – 750.

2.6. ICOTS-6

References included in B. Phillips (Ed.) (2002), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town: International Association for Statistics Education (CD ROM).). *Particularly:*

Cobb, G. W. The American Statistical Association's undergraduate statistics education initiative.

Nolan, D. Case studies in the mathematical statistics course.

Diamond, N. T. and Sztender, E. M. Simplifying consulting problems for use in introductory statistics lectures.

2.7. OTHER CONFERENCES

Hawkins, A. (Ed.) (1990). *Training teachers to teach statistics*, *ISI Round Table.* Voorburg, The Netherlands, International Statistical Institute.

Garfield, J. B. & Burrill, G. (Eds.) (1996). *Research on the role of technology in teaching and learning statistics*, Voorburg, The Netherlands: International Statistical Institute.

Batanero, C. (Ed.) (2000). *Training Researchers in the use of statistics*, Voorburg, The Netherlands: International Statistical Institute.

Brunelli, L. & Cicchitelli, G. (Eds.) (1994). *Proceedings of the First Scientific Meeting of the International Association for Statistical Education*, Perugia, Italy, (IASE): University of Perugia.

3. BOOKS

Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Erlbaum.

Anderson, C. W. & Loynes, R. M. (1987). *The teaching of practical statistics*. New York: John Wiley and Sons.

Bishop, A. J. et al. (Eds.) (1996). *International handbook of mathematical education*. The Netherlands: Kluwer. *Particularly:*

Borovcnik, M. & Peard, R. *Chapter 7: Probability,* 239 – 287.

Clarke, D. *Chapter 9: Assessment,* 327 – 370.

Shaughnessy, J. M., Garfield, J. B. & Greer, B. *Chapter 6: Data handling,* 205 – 237.

Bowman & McColl, J. (1999). *Statistics and problem solving: Computer based studies from the STEPS project*, London: Arnold.

Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh: Oliver & Boyd.

Gal, I. & Garfield, J. B. (Eds.) (1997). *The assessment challenge in statistics*, Voorburg, The Netherlands: International Statistical Institute.

Gordon, F. & Gordon, S. (Eds.) (1992). *Statistics for the 21st Century (MAA Notes No 26)*, Washington D.C.: Mathematical Association of America.

Green, D. R. (1994). *Teaching statistics at its best*. Nottingham: Teaching Statistics Trust.

Grouws, D. (Ed.) (1992). *Handbook of research on mathematics teaching and learning*, New York*:* Simon & Schuster Macmillan. *Particularly:*

Shaughnessy, J. M. Chapter 19: Research in probability and statistics: Reflections and directions*,* 465 – 494.

Hand, D. J., Daly, F., Lunn, A. D., McConway, K. J. & Ostrowski, E. (Eds.) (1994), *A handbook of small data sets*, London, Chapman and Hall.

Hand, D. J. & Everitt, B. S. (1987). *The statistical consultant in action*. Cambridge: Cambridge University Press.

Hawkins, A., Jolliffe, F. & Glickman, L. (1992). *Teaching statistical concepts*. London: Longman.

Hoaglin, D. C. & Moore, D. S. (Eds.) (1992). *Perspectives on contemporary statistics (MAA Notes No 21)*, Washington D.C.*:* Mathematical Association of America.

Holmes, P. (Ed.) (1986). *The best of teaching statistics.* Nottingham, Teaching Statistics Trust. (also at http://science.ntu.ac.uk/rsscse/ts/bts/contents.html).

Kahneman, D., Slovic, P. & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*, Cambridge, Cambridge University Press.

Kapadia, R. & Borovcnik, M. (Eds.) (1991). *Chance encounters: Probability in education*. Dordrecht: Kluwer.

Laurillard, D. (1993). *Rethinking university education*. London: Routledge. (subtitled: A framework for the effective use of educational technology).

Moore, D. S. (2001). *Statistics: concepts and controversies (5th edition)*. New York: Freeman.

Moore, T. (2001). *Teaching resources for undergraduate statistics (MAA Notes 52)*. Washington D.C.: Mathematical Association of America.

Romberg, T. A. (Ed) (1992). *Mathematics assessment and evaluation: Imperatives for mathematics educators.* Albany, New York: SUNY Press.

Scheaffer, R. L., Gnanadesikan, M., Watkins, A. & Witmer, J. (1996). *Activity-based statistics: Student guide.* New York: Springer.

Steen, L. A. (1992). *Heeding the call for change (MAA Notes 22),* Washington D.C.: Mathematical Association of America.

Tanur, J. M., Mosteller, F., Kruskal W. H., Pieters, R. S. & Rising, G. R. (Eds.) (1989). *Statistics: A guide to the unknown (*3rd edition). Pacific Grove, CA: Wadsworth.

Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.

Tukey, J. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley. *(This was the original important text. A popular interpretation is the book by Velleman et al below).*

Velleman, P. F. & Hoaglin, D. C. (1981). *Applications, basics and computing of exploratory data analysis.* Boston, MA: Duxbury Press.

4. SOME INFLUENTIAL TEXTBOOKS

Barnett, V. (1999). *Comparative statistical inference* (3rd edition). New York: Wiley.

Berry, D. A. (1996). *Statistics: A Bayesian perspective*. Belmont: Duxbury Press.

Cobb, G. W. (1998). *Introduction to design and analysis of experiments*. New York: Springer Verlag.

Freedman, D., Pisani, R. & Purves, R. (1997). *Statistics* (3rd edition). New York: WW Norton

Moore, D. S. & McCabe, G. P. (1999). *Introduction to the practice of statistics* (3rd edition). New York: Freeman.

Rossman, A. J. & Chance, B. L. (2000). *Workshop statistics: Discovery with data* (2nd edition). New York: Springer Verlag.

5. JOURNALS AND REGULAR PUBLICATIONS

Journal of Statistics Education. An online journal of the American Statistical Association at http://www.amstat.org/publications/jse

Teaching Statistics. Although primarily aimed at the teaching of statistics in schools this journal does contain articles that are relevant for teaching at first year university level. http://science.ntu.ac.uk/rsscse/ts

American Statistical Association: Proceedings of the Section on Statistical Education. These are produced each year after the ASA Annual conference. http://www.amstat.org/publications/

Newsletter of the IASE Statistical Education Research Group. Available online from http://www.ugr.es/~batanero/sergroup.htm

The American Statistician, published by the American Statistical Association, has a fairly regular feature called *Teacher's Corner* and publishes other articles on teaching statistics at the undergraduate level.

6. ARTICLES

American Statistical Association (1980). Preparing statisticians for careers in industry: Report of the ad Hoc Committee on Training Statisticians for Industry (with discussion). *American Statistician*, 34, 65-80.

Batanero, C. (2000). Controversies around the role of statistical tests in experimental research. *Mathematical Thinking and Learning,* 2(1&2), 75-98.

Berry, D. (1997). Teaching elementary Bayesian statistics with real applications in science. *The American Statistician 51,* 241 – 274. (This is the first of three papers - the others are by J. Albert and D. S. Moore - on teaching Bayesian statistics, with discussion from D. V. Lindley, J. A. Witmer, T. H. Short, D. A. Freedman and R. L. Scheaffer.)

Biehler, R. (1997). Software for learning and for doing statistics. *International Statistical Review,* 65, 167 – 189.

Ehrenberg, A. S. C. (1976). We must preach what is practised: A radical review of statistical teaching. *The Statistician*, 25, 195-208.

Friel, S. N., Curcio, F. R., &. Bright, J. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education,* 32(2), 124-158.

Garfield, J. B. (1994). Beyond testing and grading: Using assessment to improve student learning. *Journal of Statistics Education* 2(1).

Garfield, J. B. (1995). How students learn statistics. *International Statistical Review,* 63*,* 25-34.

Gelman, A. (1998). Some class-participation demonstrations for decision theory and Bayesian statistics. *The American Statistician,* 52, 167 – 174.

Greenfield, A. A. (1979). Statisticians in industrial research: The role and training of the industrial consultant. *The Statistician,* 28, 19 – 27.

Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics*, 6, 187-205.

Holmes, P. (1988). Matching Education and Assessment with Employment Needs in Statistics (MeaNs Project): final report, (http://science.ntu.ac.uk/rsscse/means/means.html) and recommendations, (http://science.ntu.ac.uk/rsscse/means/reports/finrep.html)

Hunter, W. G. (1981). The practise of statistics: The real world is an idea whose time has come. *American Statistician, 35, 71 – 76.*

Kanji, G. K. (1979). The role of projects in statistical education. *The Statistician,* 28, 19 – 27.

Moore, D. S. (1988). Should mathematicians teach statistics? (with discussion). *The College Mathematics Journa,l* 19, 3 – 34.

Moore, D. S. (1997). New pedagogy and new content: the case of statistics (with discussion). *International Statistical Review,* 65, 123-165.

Nolan, D. & Speed, T. P. (1999). Teaching statistical theory through applications. *The American Statistician,* 53, 370 – 375.

Snell, J. L. & Finn, J. (1992). A course called 'Chance'. *Chance* 5, 12-16.

Vere-Jones, D. (1995). The coming of age of statistical education. *International Statistical Review,* 63, 3-23.

Wild, C. J. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry (with discussion). *International Statistical Review,* 67, 223 – 266.

7. GENERAL BIBLIOGRAPHY

A very full bibliography of references in statistical education up to 1994 is given in:

H. Sahai, A. Khurshid & S. C. Misra (1996). A second bibliography on the teaching of probability and statistics. *Journal of Statistics Education* 4(3) (http://www.amstat.org/publications/jse/v4n3/sahai.html)

PETER HOLMES
RSS Centre for Statistical Education
Nottingham Trent University
UK

ICOTS-6. DEVELOPPING A STATISTICAL LITERATE SOCIETY[1]

Cape Town, 7th to 12th July 2002

1. REPORT FROM THE IPC EXECUTIVE

IPC Executive in Cape Town

ICOTS–6 held in Cape Town (South Africa) from 7 to 12 July 2002, was undoubtedly a great success, both from the organisational and the scientific point of view. The 472 attendants, in fact a large majority of the IASE membership, met for six days, during which they presented their papers. The superb accommodation of the Holiday Inn Cape Town, gave the delegates the possibility to exchange views and to dwell upon their past, present and future projects and activities in a pleasant place. All of this allowed participants the opportunity to strengthen old friendships and to start new ones.

The atmosphere was of an active, scientific community open to new fellow members, happy to share their experiences, to receive information and input coming from different sources and suggesting new ideas and projects for the future. Preparing for this Conference has been a challenge, and has required a long period of work.

The success of ICOTS–6 is based on well organised team work. In September 1998, the IASE Executive Committee began the process by choosing the Conference theme "*Developing a Statistically Literate Society*" and appointing a team of people to act as the Executive Committee of the International Programme Committee (IPC): M. Gabriella Ottaviani (Chair), Brian Phillips (International Organiser and Editor) and Dani Ben-Zvi (Scientific Secretary). After this a group of people were coopted as members of the International Programme Committee (IPC). Within the IPC, 11 topics were proposed with one, or sometimes two people appointed as Topic Convenors as indicated later in this report. In addition Linda Haines (South Africa) was included on the IPC as the Local Organising Committee (LOC) representative.

After presenting an abstract of their topic, each convenor(s) proposed a number of session organisers, each of whom, in turn, presented an abstract of the session and put forward at least three invited paper authors. The International Programme Committee members have co-operated worldwide to gradually construct an interesting and useful program showing the best of the activities, studies and research in Statistics Education, in order to work towards "Developing a statistically literate society". Through the interest of their proposals and the quality of the authors they invited, they ensured the standard of ICOTS–6 to be very high. This activity required a continual effort for about four years, particularly of the Executive Committee, that had to plan and supervise all the process and to maintain contacts with the Local Organising Committee. Communications were kept and maintained by e-mail.

Of enormous assistance were the IPC and the LOC Websites. A very special thanks goes to Dani and Dagan Ben-Zvi, for the wonderful IPC website at http://www.beeri.org.il/icots6 which they designed and maintained in very trying circumstances. It constantly showed where the Conference preparations were at, what the next steps were and the corresponding deadlines. This site has unfortunately been terminated but post-conference information will be available at the IASE homepage: http://www.cbs.nl/isi/iase.htm. The LOC Website, designed and maintained by the University of Natal Public Section at the address: http://icots.itikzn.co.za/ gave all important logistic local information and it proved to be most important in providing information in the unforeseeable relocation of the venue from Durban to Cape Town. We recommend a visit to the ICOTS6 IPC website now available at http://icots6.haifa.ac.il/icots6.html.

ICOTS–6 is the first ICOTS where a refereeing system has been proposed to the authors. This was due not

[1] *Statistics Education Research Journal 1(2),54-74, http:/fehps.une.edu.au/serj*
International Association for Statistical Education

only to the request coming from some colleagues who needed their papers refereed for funding purposes, but also from a scientific desire of the Exec and IPC who thought that the time had come to improve the quality of the papers presented at the ICOTS Conferences. This, no doubt increased the task of the Exec, particularly of Brian Phillips who had to follow the refereeing process, but as Brian Phillips noticed, "The larger than expected response for authors to have their paper refereed was also most heartening". All papers whether refereed or not, underwent an editing process to ensure a quality product. The papers presented at this Conference were numerous, nearly 300 papers form the Proceedings of ICOTS–6. In fact the IPC Executive Committee decision to produce a CD of the ICOTS–6 Proceedings, rather than hard copy, was very well received, and the work done by Brian Phillips and his Editorial Board to edit the CD has been epic. In addition some 25 posters were on display throughout the conference.

There were several other scientific products of the Conference: the ICOTS–6 Abstract Book is a useful printed guide to the CD. It contains titles and authors of the plenary sessions as well as sessions, titles and authors of the invited and contributed topics. Besides this, it contains the abstracts of all papers listed and an e-mail list of ICOTS–6 authors and organisers. The ICOTS–6 Papers for School Teachers is a peculiarity of ICOTS–6 that put in evidence a further achievement of this successful Conference. In fact, following the suggestions of the IPC Executive Committee, the Local Organising Committee and in particular Jacky Galpin, Delia North and Jacky Scheiber, succeeded in organising a series of events to reach out to local school teachers (see report from local organisers). ICOTS–6 Papers for School Teachers contains a collection of papers which were selected from all papers presented at the Conference by the ICOTS–6 Local Organising Committee as of special interest to South African Teachers. The 232 pages of the booklet contain the plenary sessions papers, the Statistics Literacy papers and 31 papers selected from the other sessions.

Many persons have committed themselves to design and organise this conference during the last four years, but no doubt all of this has been worthwhile. A conference like ICOTS only happens because of the commitment of a large number of people from around the world who are prepared to freely give much time and effort. We would like to pay tribute to the great support we received from so many people who helped in the making the conference such a success. This includes three IASE Executives from 1997 to 2002, an International Program Committee of 18 people, many who also worked as Topic Convenors, a Local Organising committee of 11 people, 76 session organisers and a professional events organiser and staff, Sue Bumpsteed Conferences (Pty) Ltd, Lynn Selby, the AV Coordinator for her commitment to ensuring the audio visual aspects of the conference worked to perfection and Lynne du Toit of Safari Tours for making sure everyone got there safely and had a great social program. We greatly appreciated the excellent co-operation of well over 300 authors and give special thanks to more than 70 referees who so generously gave of their time and expertise to do such a professional job. We certainly could not have finished with such a quality product without the work of the sub-editors and the CD designers. Furthermore, we wish to thank the Local Organising Committee, especially Jacky Galpin, who were extremely helpful in getting the program together as well as the many other aspects of hosting the conference. We also express sincere thanks to the 18 sponsors for easing the financial and logistic problems of running such a conference. In closing, we are very happy to say that the many dramatic events which have occurred during the planning of this conference did not deter this brilliant and devoted group of statistics educators from providing all the valuable material and getting to Cape Town for this wonderful occasion. Any statistics educator who did not attend really missed out on a very special occasion.

We think that the IASE should be proud of this event that has contributed to better understand its task and its potentialities in divulging statistics, its teaching/learning, and its usefulness in everyday life. We will be willing to assist those who are now starting the long task of planning for ICOTS-7 in Brazil in 2006.

Acknowledgements

The IASE greatly appreciates the support given by many funding and academic institutions as well as individuals who have helped make ICOTS-6 possible and contributed to the conference works and funding:

- **ICOTS International Programme Committee**: Maria Gabriella Ottaviani (Chair, Italy), Brian Phillips (International Organiser and editor of Proceedings, Australia), Dani Ben-Zvi (Scientific Secretary, Israel, also developed the IPC web page http://www.beeri.org.il/icots6), Carmen Batanero, Philip J. Boland, Carol Joyce Blumberg, Andrew I. Dale, Iddo Gal, Joan Garfield, Helen MacGillivray, Vitalis Muba, Rene H.M. Smulders, Gilberte Schuyten, Allan J. Rossman, Susan Starkings, Jane Watson, and Lawrence Weldon.
- **ICOTS Local Organising Committee**: Delia North (Chair), Jacky Galpin (Treasurer), Linda Haines (IPC representative), Sue Bumpsteed, Andrew Dale, Riaan de Jongh, Iain McDonald, Vishnu Naidoo, Mbulaheni Nthangeni, Jackie Scheiber, and Gwen Williams.

- **Funding and Academic Institutions**: ISI Teaching Statistics Trust Fund, Standard Bank of South Africa, South African Statistical Association, South African Department of Education, American Statistical Association, Education Section of the American Statistical Association, Teaching Statistics in the Health Sciences Section and Statistics in Epidemiology Section of the American Statistical Association, Vredestein Rubber Resources, Victorian Branch of Australian Statistical Association, Kibbutz Be'eri, Israel, Instituto Balear de Estadística, Instituto Catalan de Estadística, The Augsburg College, W. M. Keck Statistical Literacy Project, USA and the School of Mathematical Sciences, Swinburne University of Technology, Australia. We are also very grateful to the International Statistical Institute and its officers, in particular Marcel Van den Broecke and Daniel Berze, for its professional and technical support, as well as for providing assistance to some delegates to attend the conference through the ISI Development Fund Programme.

Finally we thank those who helped in the refereeing process and contributed to increase the quality of the papers presented in these proceedings.

2. REPORT FROM THE LOCAL ORGANISERS

A total of 472 delegates from 54 countries attended the sixth international conference on teaching statistics, held in Cape Town (South Africa) from 7 to 12 July 2002. The conference was a great success despite the relocation of the host city which was necessitated just four months prior to the event! The many months (actually years!) of hard work by both the LOC and IPC paid rich dividends as all indications are that delegates found the conference organization to be excellent - accommodation, transport, speakers audio-visual requirements, tours and social functions were of the highest standard.

ICOTS-6 LOC

The conference was supported by the City of Cape Town and the SA government, evident from the fact that, at the Mayoral reception on the Sunday night, the guest speaker was Tami Mseleku, Director General of Education, while the Master of Ceremonies was Pali Lehohla, the Statistician General of SA. In addition, the conference was opened by the Trevor Manuel, Minister of Finance of South Africa, after the flags of the countries of the participants were carried in, and the national anthem sung by a group of children, led by the famous Cape Minstrels.

Most of the delegates had accommodation in the Holiday Inn Cape Town, Strand Street, the conference venue. This cut down on transport time and gave delegates more time to network and socialize. The delegates were well catered for on the social front - a mayoral reception on the Sunday night, a “happy hour” around the posters on Monday night, wine tasting around the posters on Tuesday night, local tours on Wednesday afternoon and the conference banquet on Thursday night ensured that delegates relaxed after the academic demands of the day. Delegates booked many pre- and post-conference tours at the conference tour desk which was available throughout the conference. Visits to the various game parks in South Africa (particularly Kruger National Park) were the most popular choice, while a tour to a local township was a very popular choice for the Wednesday afternoon local tour.

A total of 25 posters were on show during the conference - amongst them three posters by South African school children. The children presented their posters on the Tuesday and thoroughly enjoyed the excitement of discussing their posters with the many interested delegates. The posters presented by the children were the winning entries in local statistics poster competitions run at a few schools.

A major local thrust for ICOTS6 was a series of events put in place to reach out to local school teachers. The South African Statistics Association (SASA), Association of Mathematics Educators of South Africa (AMESA), Statistics South Africa and the Department of Education united to present a wonderful program for local school teachers to become acquainted with basic statistics concepts (many local school teachers have had no previous statistics training) which will soon be part of the new school syllabus in South Africa.

On the Saturday a 1-day CensusAtSchool workshop was held in Durban (following on from the national mathematics school teachers conference which had just ended on the previous day). This workshop was attended by disadvantaged teachers selected from all the provinces in South Africa and focussed on the data collection process in the recent CensusAtSchool project in South Africa, as well as demonstrating how to use CensusAtSchool materials in the classroom. This workshop was repeated in Cape Town on Sunday. The attendees of this CensusAtSchool workshop consisted of a some international ICOTS delegates, local school teachers from the Cape Town area and most significantly, a group of teachers, from each province in South Africa, selected by the Department of Education. These teachers include key mathematics co-ordinators from the 9 provinces of South Africa. Aspects of CensusAtSchool from other countries were also presented at both workshops, giving an international perspective to the data sets, which will be made available to all schools in South Africa.

A local teacher session, running for the full duration of the conference, was organised by SASA and AMESA, and ensured that the teachers got sufficient training in statistics to be able to meet the demands of the statistics section of their new school syllabus (to be fully implemented in 2005). The local teacher session was split into two strands, Primary (grades 4,5,6) and Senior (grades 7,8,9). A workshop approach prevailed throughout and this ensured that the teachers would have adequate materials to use in the classroom. Each teacher received a die, plastic cups and various coloured poker chips and in no time groups were merrily simulating their data and arguing the finer points of probability theory! Other sessions focussed on using details of histograms, charts, plots and other aspects of the school syllabus, as well as interpretation of newspaper articles and other material incorporating statistical concepts. The teachers were very excited to discover the relevance of statistics to all aspects of teaching at school, and in fact to all aspects of life. The local teacher session and the CensusAtSchool workshop was captured on video camera in order to be used in follow-up workshops to be held in the various provinces in South Africa. Presentation of these workshops was a requirement for funding received by many of the teachers who attended ICOTS6. Support from SASA and AMESA will assist these teachers in spreading knowledge gained at ICOTS6.

ICOTS6 certainly gave local teachers the training to assist in creating a statistically literate society in South Africa!

3. OPENING ADDRESS

THE HONOURABLE TREVOR MANUEL (MP)
Minister of Finance
South Africa

The Opening Address at ICOTS6 was delivered by South Africa's Minister of Finance, The Honourable Trevor Manuel. It is exceedingly rare that we get the perspective of a government policy maker at the most senior cabinet level on statistics and its role in policy making, and on statistics education. So we are publishing the text of this excellent speech in full in the pages that follow.

3.1. SIXTH INTERNATIONAL CONFERENCE ON TEACHING OF STATISTICS OPENING ADDRESS

I would like to commend ICOTS and the organising committee for the work done in preparation for this conference. In looking at the programme my particular favourite is Session 4H "Educating Managers, Executives, Lawyers, Politicians, Government officials and other Decision Makers". Numeracy for Lawyers? For Politicians? What next – elementary numeracy for Auditors?

My enthusiasm for the work of ICOTS arises from my responsibilities as a policy maker in South Africa. Our young democracy is but 8 years old and was built to improve on the quality of life of all its citizens. This task is undertaken against the history of huge inequalities; not just in material circumstance, but also in access to knowledge and understanding. I believe fundamentally that democracy works when citizens participate, because citizens know, because they've been empowered through education.

Dr. Helmut F. Spinner (1999) (German obviously), at the conference on policies and statistics in the European Union: Challenges and responses, defines a Knowledge Society: as a well informed society in fact, that should become increasingly better informed and he argues that in a complete knowledge society, all the knowledge of the world will be available to everyone, available everywhere, available simultaneously and available freely. This is what we are striving to achieve here in South Africa. Dr. Spinner poses the five preconditions for this to happen as:

- The non-technological infrastructure should first be upgraded
- Literacy should be achieved
- There should be promotion of use
- There should be promotion of access
- Basic freedoms should be guaranteed.

The question to yourselves is what knowledge do we impart in the teaching of statistics? I will hazard it is knowledge as understanding and knowledge as insight. The latter being more important in that we do not see people in our offices churned out of universities performing at this third level of knowledge. That is knowledge as insight.

Larry Gonick and Woollcott Smith (1993), authors of the cartoon guide to statistics say "*we muddle through life making choices based on incomplete information*." In order to make informed decisions in the face of incomplete data we often use statistics and…."*what makes statistics unique is its ability to quantify uncertainty, to make it precise. This allows statisticians to make categorical statements, with complete assurance about their level of uncertainty.*" By quantifying this uncertainty, one begins to know what to do about risk, and what resources to allocate to it. On a daily basis my political portfolio has to deal with uncertainties about the markets, currency behaviour, prices of the bullion, the all share index, inflation rate, GDP growth, savings and investment, sectoral performance, employment, investor confidence, likely direct foreign investment, asset security, their growth, adequacy of their deployment and ultimately derivation and appropriation of value for society. In examining this array of information, the risk is that I may confuse the noise of so much information with reliable statistics.

I do not wish to replace the statisticians in Stats SA, they must be consummate professionals, nor do I wish to undermine the valuable work that the Statistics Council does in providing an external quality assurance to the work of Stats SA. I seek, as policy maker to define the terms for engagement between the statistical agency and Cabinet.

From the list of matters outlined above that concern government and my portfolio, you can realise that I have to deal with statistics matter, that is, measurement of inputs, outputs, outcomes and impact. While the first of the four measures might be a matter of arithmetic and largely financial accounting, the latter three are the subject of statistical collection, collation and derivation of indices for measurement, evaluation and decision making for determining:

- What types and levels of inputs should be made?
- Why these inputs should be made?
- For whom these inputs should be made?
- When these inputs should be made?
- Where these inputs should be made?
- How these inputs should be made to have maximum effect?
- How will I know that the inputs are working for me? Am I measuring what I need to measure? How will I recognise success? How will I recognise failure?

"Three fourths of the mistakes a man makes are made because he does not really know what he thinks he knows" (James Bryce 1838-1922). Am I measuring what I need to measure? Do I have confidence in the indicators I am using? And what is my confidence level?

My responsibility as a Minister of Finance is to firstly ensure that the country has sufficient high quality statistics produced by our statistical agency and secondly to ensure that the populace can receive the statistics. This is where educators come in. Thirdly, Government has to intermediate between the generation and the use of statistics.

Indeed *"when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind. It may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science*". (William Thomson, Lord Kelvin 1824-1907).

“Before Census@Schools I did not know how tall I was”, remarked a young girl of 15 years of age from a remote school. This was in September 2001 when StatsSA conducted Census at Schools. “The census says we are not as many as we are” was exclaimed a resident from the Western Cape Province when the census results of Western Cape were released in October 1998. “How can we be classified as not being poor” was the question raised by a resident of the Province of Mpumalanga when the report on poverty measurement was released in 2000.

Statistics are a matter of life and death too. Where they are applied to amplify knowledge they can save lives and where they are ignored losses can be incurred.

Larry Gonick notes that in 1986, the space shuttle challenger exploded, killing seven astronauts, including a lady teacher. The decision to launch at a temperature of 29 degrees had been made without doing a simple analysis of performance data at low temperature. On the other hand the trials of salk polio vaccine performed in 1954 on a sufficiently large sample of children, 400,000 in number eliminated bias in the results. Robust statistical analysis of the data firmly established the vaccine’s effectiveness, and today polio is a thing of the past.

Why should a young girl of fifteen at school fail to know about herself? She says I did not know how tall I was. A UN Statistics Division Handbook on the Operation and Organization of a Statistical Agency (December 2001) notes that, encouraging schools and high schools in the learning of statistics constitutes best practice. In fact they note that in Poland, there is an active high school competition for the best essay in which extensive use is made of official statistics. I have also realized that Canada recently has introduced the teaching of statistics for teachers and as I surfed the Web I came across a website on statistics in primary schools for Japan.

It will be desirable that statistical institutions avail their stacks of data to schools for use as teaching material. The United Kingdom in outlining their management model for statistical office opine that there has to be a critical mass of key skills. So you are potentially producers of these skills for that reason you have to seek relevance.

I have looked at the programme and I have been struck by a few topics that will be handled, and allow me to reference them without influencing you on which ones you should attend. I am pleased to see that a political angle is brought to the fore through a presentation titled looking at the behaviour of the electorate by Theodore Chadjipaledis. Larry Gonick in his book “a cartoon guide to statistics” argues that all this probability stuff is only good before an election. Your poll star statistician will tell you that I am 95% or 99% confident that you will win the election. This is so because of several things, such as response bias because voters can lie to the interviewer, secondly, the actual voters is what counts although the potential voters constitutes an unbiased sample and thirdly voters may not be home to answer the poll. After the election, the senator is either 100% in or 100% out.

Let me digress by drawing attention to a matter of fundamental importance to policy makers all over, but especially in the developing world, namely HIV and Aids. There is no dispute about the severity of the disease but it is exceedingly hard to deal with this in an environment where there is no reliable statistics available. In South Africa we have results from ante-natal clinics and we have available actuarial models constructed for appreciating risk in the life assurance companies and not for public policy making. In addition the disease remains stigmatised in communities and doctors do not always capture Aids as a cause of death raising concerns about the veracity of parts of a huge mortality study presently underway. Simultaneously, there are a range of interests on this matter who all claim infallible statistics. In an environment of the absence of a cure, policy makers face a huge dilemma on resource allocation. Would money be best spent on research into a cure or a vaccine, on preventive education, on drug therapies, on care for people living with the disease or welfare on those left behind? Each of these choices is relative and dependent both on the reliability of statistics and on the ability to engage with the populace. In many respects the results will be shaped by the extent of exposure of ordinary people to the basic natural sciences; physiology, nutrition and efficacy of drugs which all influence lifestyle choices. None of this is assisted by statistical noise be this on infection rates or life expectancy. Part of my appeal to this conference is to give attention to the interconnectedness between the teaching of numeracy and the links to that which would empower people.

I am pleased that Professor John Volmink is presenting on the issues of statistical literacy for South Africa. My appeal to Stats SA is to engage with these matters and to bring results which would allow other parts of government to take forward the work of Census@Schools. I want to welcome the entry into the discussion of community based learning and wish to express appreciation to Trisha Thorne and Rob Root for their paper. I am encouraged by the paper by Madden and Choi on the duties of statistical agencies to their clients. I am happy that Koffi N’Guesson is advancing discussion on the training of African statisticians – our heads of state gathered in Durban to launch the African Union and take forward the New Partnership for Africa’s development will depend for their efforts on good quality statistics to advance these programmes.

My challenge to all of you is to help us help the 15-year-old who did not know how tall she was. My challenge to all of you is to help us to understand what we should focus on to build an empowered society. My challenge to this conference is to help us to meet Spinner's definition of a knowledge society.

I wish you fruitful deliberations.

THE HONOURABLE TREVOR MANUEL
Minister of Finance
PO Box 29
Capetown
8000
South Africa

3.2 REFERENCES IN TREVOR MANUEL'S ADDRESS

Chadjipaledis, T. (2002). Using statistics as a tool in political research. The case of electoral behaviour. Proceedings, ICOTS6. (See section 5.4)

Gonick, L. and Smith, W. (1994). The cartoon guide to statistics. Harper Collins.

N'Guessan, K. (2002). Training of African statisticians: the experience of the National School for Statistics and Applied Economics (ENSEA) in Abidjan. Proceedings, ICOTS6. (See section 5.4)

Madden, R. (2002). Statistical agencies' duties to their clients: who teaches, who learns? (Presented by Ching Choi) Proceedings, ICOTS6. (See section 5.5)

Thorme, P. and Root, R. (2002). Community-based learning: motivating encounters with real-world statistics. Proceedings, ICOTS6. (See section 5.5)

Volmink, J. (2002). The importance of numeracy and in particular of statistical literacy for South Africa. Proceedings, ICOTS6. (See section 4)

3.3. MINI-BIOGRAPHY OF THE HONOURABLE TREVOR MANUEL

(from the South African National Treasury website http://www.finance.gov.za/people.htm)

Trevor Andrew Manuel was born on 31 January 1956 and grew up in Kensington, Cape Town where his mother Philma still lives. He attended Windermere Primary School and matriculated from Harold Cressy High School in Cape Town.

Mr. Manuel completed a National Diploma in Civil and Structure Engineering at the Peninsula Technicon and studied Law during his many periods of detention. He practised as a technician until 1981, whereupon he entered public life. From 1981 he was the General Secretary of the Cape Areas Housing Action Committee. In 1983 he was elected Regional Secretary and a National Executive Member of the United Democratic Front (UDF), a broad anti-apartheid coalition.

For these activities, he was repeatedly detained without trial or placed under house arrest between 1985 and February 1990, spending a total of thirty-five months in detention.

He was elected to both the National Executive Committee and the National Working Committee of the ANC in 1991 and appointed head of the ANC's Department of Economic Planning. In this capacity, he was responsible for the shaping of ANC Economic Policy.

At the time of the historic elections of 27 April 1994 he was elected as an ANC Member of Parliament and in May 1994 he was appointed Minister of Trade and Industry. On 4 April 1996, he was appointed Minister of Finance, a position he currently holds.

Mr. Manuel was selected by the World Economic Forum as a "Global Leader for Tomorrow" in January 1994, he was appointed to the Advisory Committee of the UN Initiative for Trade Efficiency. In September 1994, he was awarded the Africa Prize by the German Africa Foundation, jointly with Derek Keys, who was the Minister of Finance in South Africa at that time. In March 1996 he was awarded the Rotary International Paul Harris Award for Outstanding Achievement by the Rotary Club of Isipingo-Prospection.

4. SUMMARY OF PLENARY SESSIONS

John Volmink, South Africa. *The importance of numeracy and in particular of statistical literacy for South Africa.*

The development of a Revised National Curriculum Statement is seen as a key project in the transformation of South African Society. The thrust of the project is towards achieving "a prosperous, truly united, democratic and internationally competitive country with literate, creative and critical citizens leading productive, self-fulfilled lives in a country free of violence, discrimination and prejudice." (Curriculum 2005, Learning for the 21st Century 1997, Department of Education, Pretoria.). Curriculum reform in South Africa thus faces a two-fold challenge. The first is the post-apartheid challenge which requires developing the knowledge, values and skills base for South Africa's citizens necessary for greater social justice and development. Secondly, there is the challenge of participating in a global economy. This raises questions about the knowledge, values, skills and competences for innovation and economic growth for the 21st Century. The view taken by the curriculum designers is that the best route to greater social justice and development is through a high-knowledge and high skills curriculum. This paper will explore the meaning and importance of numeracy and in particular of statistical literacy, within this context. The paper will focus largely on the relationship between values and mathematical/statistical literacy within the South African context.

Scott Murray, Canada and Iddo Gal, Israel. *Preparing for diversity in statistics literacy: institutional and educational implications.*

Improving the public's understanding of statistical information requires that producers or reporters of statistical messages are aware of: The nature of people's statistics literacy, The factors that affect the difficulty of statistics-related messages, The existence of individual or group differences in statistics literacy; and The information needs of different target audiences. Implications are discussed regarding the need to prepare different types of communicative products and formulate strategies for dissemination and public education.

Jane M. Watson, Australia. *Doing research in statistics education: more than just data.*

As teachers of statistics we know the fundamental components of statistical enquiry, be it classical or exploratory. When we turn the focus on ourselves as statistics educators, we run the risk of forgetting some of the fundamental principles of good research – principles that are broader than carrying out statistical significance tests. In this talk I want to present some examples of research in statistics education to illustrate the stages and outcomes that contribute to results that have a scholarly impact on the statistics education community. As a single teacher with a good idea on how to teach "confidence intervals," I do not expect anyone to pay much attention to me. If I can, however, place my ideas in the context of others' ideas or research on teaching confidence intervals; conduct a study – maybe a case study or a controlled experimental design – that is valid for considering the issue I want to promote in teaching about confidence intervals; and have my results refereed by peers in the field; then I can expect people to pay attention to me.

Peter Holmes, UK. *Teaching, learning and assessment: complementary or conflicting categories for school statistics.*

Over many years I have been attempting to improve statistical literacy in the population by changing the school curriculum. All such attempts have to be put in the general context of teaching, learning and assessing the subject. Ideally these should complement and reinforce each other. In practice they often conflict - in particular assessment can distort the learning process. In this talk I consider the nature of these conflicts and how they might be overcome in practice, giving examples from a lifetime's experience.

Maria Gabriella Ottaviani, Italy. *1982-2002: From the past towards the future.*

This paper, after considering the reasons and aims that gave origin to the International Conference on Teaching Statistics (ICOTS), traces the line of thought along which the Conference developed from 1982 to 2002. This is done by applying textual data analysis to the titles of the papers published in the Proceedings of the first five Conferences, and to the titles that were on the International Programme Committee Web site on October 27, 2001. Knowing past and present enables one to present suggestions about possible future Conference scientific developments.

George W. Cobb, USA. *What can cheap computing offer statistics literacy?*

For statistics at the research level, cheap computing has pushed aside the old order and ushered in a new one based on algorithmic thinking. This is a much deeper change than just teaching an old dog faster ways to do the standard tricks. The old dog now lies content in the sun; computers have brought us an energetic young puppy who is eagerly exploring the whole field anew. What does all this mean, if anything, for statistical thinking at the introductory level? Is it time for the standard curriculum to lie down in a sunny spot and give some new puppies a chance? My view on this is still evolving, in response to what I read, what I hear from colleagues, and what I learn from my students in the classroom. Thus I can't say for certain where my thinking will be in July 2002, but my general theme will be the opportunities that algorithmic thinking offers us to revisit basic ideas of statistics.

International Forum on Statistics Literacy: Statistical Literacy, Statistical Numeracy and Developing Society's Statistical Health and Richness. Chair: Helen McGillivray, Australia. *Panel Members:* Luigi Biggeri, Italy, Lisbeth Cordani, Brazil, Pali Lehohla, South Africa, Jessica Utts, USA.

During the past decade, many statisticians and statistical educators have discussed or been involved in matters ranging from the accreditation of statisticians to education for statistical thinking to the interaction of statistics and information technology to the problems of society numeracy and mathematics education. Even the ICOTS6 theme itself has produced considerable discussion amongst statisticians as to whether we should talk about statistical literacy or statistical numeracy. What is common in all these topics of discussion, debate and involvement, is the understanding that statistics is integral to an extensive range of functions of society and to many disciplines, and that society's statistical health depends on facilitating development of, and access to, a full and diverse range of statistical capabilities. That is, as with mathematics, the totality of society's statistical health and strength depends on its ongoing development of the continuum of statistical understanding, abilities and skills, from universal basic statistical literacy through the full spectrum to highly sophisticated scientific and management capabilities. With the addition of a suitable description of some form of universal basic statistical literacy, the definitions of types of statistical ability and skills discussed in Bartholomew's Royal Statistical Society's Presidential address (1995), illustrate this continuum. Facilitating the development, resourcing and ongoing nurturing of this continuum is an enormous challenge. The forum speakers, together with other contributers and questions, will comment on just some aspects of this challenge from both international and personal perspectives.

5. SUMMARIES OF TOPIC SESSIONS

Note: Complete data of authors, organisers and presenters as well as full papers, including those not presented orally, can be found in the **ICOTS-6 Proceedings CD**, which are available from the International Statistical Institute Permanent Office, P.O. Box 950, 2270 AZ Voorburg The Netherlands, Fax 31-70-3860025, E-mail: isi@cbs.nl.

5.1. TOPIC 1. STATISTICAL LITERACY. Convenors: Iddo Gal and Brian Phillips

There are many questions involving statistical thinking which confront people when they are at work, handling household affairs, reading a newspaper or watching TV, or in leisure. Some examples are: How does investing on the stock market compare with gambling? How should we interpret news stories about the latest health research findings? Is it wise to bet on the underdog? How can we interpret reports involving tables and graphs? How can we know when statistics are being misused or abused in the media or in advertisements? Such and related questions face people from all walks of life, whether or not they have had any formal education in statistics.

Under this topic many questions were discussed involving statistical thinking which confront people when they are at work, handling household affairs, reading a newspaper or watching TV, or in leisure. The term "statistical literacy" does not have a single accepted meaning, but in general refers to people's ability and propensity to interpret, critically evaluate, and communicate about statistical information, data-related claims, or chance-related phenomena which they may encounter in diverse life contexts. In many societies or communities citizens are increasingly being expected (or declare their right) to be informed and act as critical consumers of statistical and probabilistic information. To that effect, presentations suggested some level of statistics literacy that may be desirable not only from all school or university graduates, but also of all adults, regardless of their educational and personal backgrounds.

Presentations at Session 1A: *Frameworks and Studies in Statistics Literacy*, organised by Iddo Gal (Israel) dealt with 'Three kinds of statistical literacy: what should we teach?', Milo Schield, 'Analysis of data from a nationwide psychological project involving coin-tossing predictions', David Green and 'Profile for statistical understanding', Chris Reading.

Topics for presentations in Session 1B: *Statistics for the Citizen*, organised by Brian Phillips (Australia) run in two sessions, were 'What educated citizens should know about statistics and probability', Jessica Utts, 'Promoting statistics thinking amongst secondary school students in the national context', Philip J. Boland, 'Toward a statistically literate citizenry: what statistics everyone should know', Jerry L. Moreno, 'Experience of dealing with the media on congenital anomaly research', Beverley Botting, 'Dna "fingerprints" and their statistical analysis in human populations', A.Marie Phillips and 'Probability and game shows', Mike Fletcher. All the speakers gave very good presentations and these sessions proved very popular. Presenters also described some research-based findings or concerns based on classroom experiences regarding statistics literacy levels of learners or people in general, and suggested various courses of action that educators can follow, either in terms of specific classroom activities or regarding the needed general approach to instruction and skill development.

5.2. TOPIC 2. STATISTICS EDUCATION AT THE SCHOOL LEVEL. Convenor: Jane Watson

Topic 2 focused on statistics education at the school level and it was encouraging to observe that the response at ICOTS6 was the largest at any ICOTS thus far. Twenty-one papers were presented at the conference under five subgroupings reflecting innovative ideas at the elementary level (organized by Dave Pratt), innovative ideas at the secondary level (organized by Gail Burrill), research (organized by Pat Thompson), curriculum development (organized by Dani Ben-Zvi) and sociocultural aspects of learning (organized by Paul Cobb). Although there was some overlap of interest with other topic groupings within ICOTS6, particularly related to Research (Topic 6) and Technology (Topic 7), the focus in Topic 2 was intended to reflect the context of what occurs in relation to the teaching and learning of statistics at the school level.

Papers in Session 2A, *Statistics in Elementary School*, organised by Dave Pratt (UK), considered three different aspects of statistics learning by young children: variation, the arithmetic mean, and randomness. Grade 3 students were the recipients of 10 lessons on chance and data with an emphasis on variation in the report presented by Jane Watson. Pre and post tests indicated a significant gain in basic understanding of both chance and data, and the part variation plays in relation to the fundamental concepts. José Luis Cortina discussed three different ways that 12-year-old students make sense of the arithmetic mean as a ratio. These were based on in-depth interviews using problems from the initial phases of the curriculum and provided a basis for suggestions for designing instruction. Efthymia Paparistodemou discussed a case study of 6-8-year-old children working in a game-like environment to construct spatial representations of a sample space. With the ability to inspect and reconstruct the rules of the game, children displayed their understanding of random events within a novel medium.

Papers in Session 2B, *Innovative Ideas for Teaching Statistics in Secondary School*, organised by Gail Burrill (USA) reflected perspectives from four different countries: the United Kingdom, the United States, the United Arab Emirates, and Australia. James Nicholson outlined the pedagogical basis for a range of materials, including the use of technology, in particular to allow the exploration of larger data sets and of the effects of random variation. Based on a project in Northern Ireland, examples of materials focusing on some key statistical concepts were presented. The relationship of statistics to the rest of the mathematics curriculum was the feature of Jeffery Witmer's presentation based on the Data Driven Mathematics curriculum. He suggested that many mathematical topics, such as formulas, linearity, centers, inequalities, and matrices can be used to motivate and lay the foundation for the relevance of statistics both to mathematics itself and to the real world. The perspective taken by Hanan Innabi reflected the importance of critical thinking throughout the school curriculum. Examples were given from secondary school statistics to show how statistics provides a practical, interesting, and different way to facilitate critical thinking. Using his experience as an Australian secondary teacher, Anthony Harradine, suggested ways of moving beyond the traditional methods of teaching the Normal Distribution and the Central Limit Theorem. These focused on conceptual understanding, meaningful calculation of probabilities, and using the outcomes in an integrated fashion.

Papers in Session 2C, *Research on Teaching Statistics at the School Level*, organised by Pat Thompson (USA), reflected many diverse approaches to research in statistics education at the school level: research related to students, research related to teachers, research involving both teachers and students, and the overall relationship of research to the classroom. Dani Ben-Zvi described grade 7 students' cooperative work on a data assessment task in a computer-assisted environment following a carefully designed Exploratory Data Analysis course. Of interest were their formulations of research questions and hypotheses, and use and interpretation of data representations. One presentation that focused on research only on teachers' understanding was that by Jinfa Cai. He compared and contrasted twelve inexperienced and eleven experienced teachers' constructions of pedagogical representations for teaching the arithmetic average. Although the experienced teachers were better able to predict representations and errors common for middle school students, these were not more evident in their generation of lesson plans. Two studies reported on projects involving both teachers and their students. Silio Rigatti Luchini presented the results of a study involving 145 teachers and over 2000 students aged 6-10 in five Italian provences, with teachers attending a preliminary training course. Concept maps were used as a method of comparing student understanding before and after the teaching occurred. In the other presentation James Nicholson reported on a project where a number of teachers and students worked iteratively with materials being developed in areas such as Correlation and Regression, Sampling Methods, and Estimation. The aim was to produce materials to address conceptual difficulties occurring at the A-level (in the U.K.). Finally, Gail Burrill looked at the overall picture of the relationship of research to teaching at the school level, in particular in relation to the production of curriculum materials. Although indicating that progress has been made in some areas, she raised several questions that would help build a more coherent story for future use of research outcomes in the classroom.

Of the papers in Session 2D, *Innovative Statistics Curriculum Development and Research Projects at the School Level,* organised by Dani Ben-Zvi (Israel) three focused on the curriculum itself, one on the implementation of a new curriculum, and two on instructional issues. The CensusAtSchool project for students aged 7 to 16 was discussed by Doreen Conner. Based on a website in the UK, the project offers the opportunity for students to gather information about themselves and then become part of national and international (including Australia and South Africa) data bases. Annie Morin addressed the curriculum issues associated with statistics becoming a part of the mathematics curriculum at the school level, interdisciplinary projects being developed, and the increasing availability of computers. She felt the movement of topics downward from the tertiary level required a more precise definition of objectives and the ways to, and limits on, achieving them. Peter Holmes, on the other hand, reviewed the past 40 years of curriculum change with respect to school statistics. Mainly using experiences from the United Kingdom and the United States, he discussed the successes and the lessons to be learned for the future. One paper dealt with curriculum implementation. South Africa is about to introduce a new curriculum, which will include statistics education for the first time, and Delia North presented a plan to assist teachers who have little or no training in statistics. A pivotal aspect of the plan is that statistical training be developed according to the age of the learners, bearing in mind the mathematical tools at their disposal. One of the instructional presentations, by Gianfranco Galmacci and Anna Maria Milito, described the results of an Italian study of 6000 students at every school level and 338 teachers, comparing how different teaching approaches influenced the students' learning processes. The other paper on instruction, by Koeno Gravemeijer, presented an instructional design heuristic called "emergent modeling", with a sequence on data analysis as an example. Emergent modeling focuses on modeling as "organizing" and has a dual meaning, related both to the process by which models emerge, and to the process by which these models support the emergence of more formal knowledge.

The three papers in Session 2E, *Sociocultural Aspects of the Learning of Statistics at the School Level,* organised by Paul Cobb (USA), presented very different aspects of the sociocultural influences on statistical learning and understanding. Paul Cobb reported on a classroom design experiment where 12-year-old students developed identities as those who chose to engage in, saw value in, and viewed themselves as competent at developing data-based arguments. He also discussed the aspects of the design that supported this transition. Collaborative work was the focus of the report by Carolina Carvalho, which considered the encouragement of peer interaction during class. Analysis of excerpts of interactions was the methodology employed. Celia Hoyles and Richard Noss described the findings of a study of the ways paediatric nurses think about the notions of average and variation. Conclusions were drawn about ways that more general mathematical meanings are constructed and "transferred", taking into account both cognitive and sociocultural perspectives.

Overall the 21 papers provided a wide range of views and insights into the teaching and learning of statistics at the school level. The discussions following the presentations were often lively and many contacts were made for the future exchange of outcomes and ideas. Finally I would like to thank Lionel Pereira-Mendoza who began the organization of Topic 2 but had to withdraw part way through the process.

5.3. TOPIC 3. STATISTICS EDUCATION AT THE POST-SECONDARY LEVEL. Convenors: Gilberte Schuyten and Allan J. Rossman

Statistics is a crucial part of the education of students in many disciplines and is used by an ever increasingly number of people in the workplace. Statistics is studied maybe by more students at post secondary level than any other subject is. A wide variety of statistical contents and student audiences is covered at this level. The presentations at ICOTS-6 provided a representative sample of ways statistical contents and skills can be taught taking into account different backgrounds of students. Sessions were categorized by type of course, student audience, pedagogical approach, and statistical topic.

Session 3A, *Statistics as a Service Subject in First Level Courses* was organised by Beth Chance (USA). The last decade has seen a renewed focus on the introductory statistics course for non-statistics majors at the tertiary level. This "service course" has been infused with recommendations for active learning, conceptual understanding, real data, and effective use of technology. Instructors have aimed to make the course more interesting and accessible to students by connecting the material to their own majors and non academic experiences. Speakers in this session discussed recent innovations in statistics instruction aimed at making the course more relevant to a general audience. Particular attention was paid to experiences that allow students to apply their knowledge in a social or humanitarian context, and to how these experiences have impacted students' perception of the utility of statistics.

Session 3B, *Statistics as a Service Subject in Second Level Courses*: Teaching Regression Models was organised by Joachim Engel (Germany). Modeling functional relationships between two or more variables is a central topic of applied statistics. In this session various approaches were discussed to teach regression models. The focus was on teaching understanding and concepts of various regression models, not on mathematical details. The technological aspect of using modern software as a tool to promote understanding (e.g. through simulation) was also covered. Presentations dealt with 'Interpretation of regression output: diagnostics, graphs and the bottom line', 'Understanding regression', 'Accessible methodologies for estimating density functions' and 'Advanced topics for a first service course in statistics'.

Session 3C, *Statistics for Future Statisticians* was organised by Ann Cannon (USA). A major effort has been undertaken recently in the United States to identify what should constitute a major (first degree) and minor in Statistics. Participants in the discussion have included industrial statisticians (end-users) as well as faculty from a broad range of post-secondary institutions. The first speaker discussed the results of the Undergraduate Statistics Education Initiative (USEI) in the United States. The second and third speakers reported on the status of majors (first degrees) and minors on other continents.

Session 3D, *Statistics and Research Designs: An Integrated Approach*, organised by Glenys Bishop (Australia). Much current statistical education at the post-secondary level focuses on methods of data analysis and the use of statistical packages. Subject matter experts often teach research methods to upper undergraduates and beginning postgraduate students but their emphasis is less on the statistical validity of study designs as on their practicalities. The papers showed how the integrated approach can be used in teaching undergraduates or postgraduates, for experimental or survey or observational study designs. Presentations dealt with 'Experimental research in a statistical concepts course', 'Teaching statistics and research methods in a virtual learning environment', 'Improvement of teaching and use of statistics in Africa's Sub-Saharan countries: the example of Benin' and 'Statistics made alive'.

Session 3E, *Statistics Learning with Cases/Projects*, was oganised by Roxy Peck (USA). Speakers in this session described how case studies and/or projects have been integrated into statistics instruction at the post secondary level in innovative ways, and addressed the resulting impact on student motivation and learning. Presentations dealt with 'Case studies in the mathematical statistics course', 'Statistical investigations – Drawing it all together', 'Survey sampling: learning by doing. A twenty years graduate level teaching experience' and 'Projects for advanced undergraduates – Leaving the script behind'.

Session 3F, *Bayesian Statistics* was organised by Dalene Stangl (USA). Due to advances in numerical methods and computation, use of Bayesian methods is rapidly increasing both within the statistics profession as well as in substantive research areas. This is evidenced by a rise in publication of Bayesian-based statistics textbooks and a rise in publication of substantive research articles using Bayesian methods. Teaching of Bayesian methods in undergraduate courses must follow. Speakers discussed the ease and difficulty of teaching the Bayesian perspective and shared teaching resources with those interested in bringing Bayesian methods into their own courses.

Session 3G, *Nonparametric Methods,* was organised by Noel Veraverbeke (Belgium). The so-called nonparametric methods began some fifty years ago. Initially they aimed at distribution-free procedures, not requiring the restriction of a parametric model for the data. These methods are typically based on the ranks of the observations and turn out to be less computational and conceptually simpler than their parametric counterparts. In later years further important classes of nonparametric methods developed. It is also important to note that the ever more powerful computer packages greatly influenced the whole area of nonparametric statistics. Advantages and disadvantages were discussed and attention was given to the way the methods fit in the teaching of our statistics courses. Presentations dealt with 'The teaching and practical implementation of the nonparametric bootstrap', 'A short introduction to nonparametric curve estimation', 'Visual basic applications and spreadsheet for teaching estimation of nonparametric density and regression functions' and 'Tests for interaction in a two-way layout: Should they be included in a nonparametrics course?'

Session 3H, *Teaching Consultancy Skills to Statisticians* was organised by Gabriella M. Belli (USA). Some ways that consultancy skills may be acquired is via formal coursework, mentorships, job training programs, by working jointly with faculty consultants, by informal training or observation, or simply through one's own experience. The papers in this session focused on various aspects of preparing students for statistical consultancy.

Session 3I, *Statistics for Future Teachers* was organised by Graham Jones (USA) and Zakayo Msokwa (Tanzania). In response to the critical role that information and data play in our technological society, there have been international calls for reform in statistics education at all grade levels of the school mathematics curriculum. If this reform is to be successful, prospective teachers of elementary, middle, and high school grades need to be equipped with the content and pedagogical knowledge to teach data handling effectively.. Speakers in this session presented and discussed innovative programs in statistics education for prospective teachers of elementary, middle and secondary school.

Session 3J, *Statistics for Future Health Care Professionals*, was organised by Tom Short (USA). Health care professionals must balance competing demands in addition to their primary concern of patient care. This session brought leading educators from a variety of health care education contexts together to share their views on training future health care professionals. Presentations dealt with 'Making statistics relevant for undergraduate nurses', 'From testing to decision-making: changing how we teach statistics to health care professionals', 'Usage of medical journal articles in biostatistical training for residents' and 'Intensive short-courses in biostatistics for fellows and physicians'.

Session 3K, *Sampling for Surveys*, was organised by Alan H. Welsh (Australia).The image of sampling conveyed by many books and courses is of a dry, turgid subject beloved by pedants. The content often seems like an unconnected set of topics characterised by clever but apparently ad hoc methods. In fact, sample surveys are an exciting, important area of statistics in which all the statistical issues appear – how to collect data, how to analyse data, how to interpret data, appropriate frameworks for inference, conflicting paradigms etc. So how can we teach basic sampling with verve and style? The talks in this session explored different approaches to answering this question. Presentations dealt with 'Training professionals in survey sampling', 'Use of mini-projects in the teaching of survey sampling', and 'Survey training for official statisticians in Brazil'.

Session 3L, *Multivariate Statistics,* was organised by John Harraway (New Zealand). As well as discussing some of the more recently developed multivariate techniques, this session included series of papers highlighting the breadth of application by covering methodology and data from such diverse areas as epidemiology, ecology, environmental science, marketing research and the social sciences. For each of these subjects, interesting recent data sets were presented. Presentations dealt with 'Multivariate methods for ecology and environmental science', 'Hierarchical linear models for the analysis of longitudinal data with applications from HIV/AIDS program evaluation' and 'Making multivariate interesting and fun for students'.

Session 3M, *Hypothesis testing* was organised by Alan McLean (Australia). The title of papers were: 'Hypothesis testing in psychology: throwing the baby out with the bath water?', 'Statistacy: vocabulary and hypothesis testing', 'Hypothesis tests, confidence intervals, and common sense' and 'How significance tests should be presented to avoid the typical misinterpretations'.

Session 3N, *Teaching Categorical Data Analysis* was organised by Michael Campbell (UK). Categorical data are traditionally analysed by the chi-squared test, one of the first statistical tests to be developed in the history of statistics. Since then there has been considerable work on other models to describe categorical data, including ordinal, polytomous, continuation ratio, stereotypical and Row and Column models. However understanding the consequences and usefulness of these models has still some way to go. The following papers were presented:

'Teaching categorical data analysis', 'Teaching statisticians and applied researchers statistical methods for analysis of data from rating scales' and 'Teaching statistics on-line: our experiences and thoughts'.

Session 3O, *Statistics for the Actuarial Syllabus*, organised by Jacky Galpin (South Africa). Much of the content of the actuarial syllabus is made up of statistical materials, and is generally taught by statisticians and not actuaries. Recent changes to the actuarial syllabus have required the teaching, at undergraduate level, of material normally taught at postgraduate level. This session presented some aspects of these issues. Presentations dealt with 'Plan member risk and the defined contribution pension plan', 'Net value and ruin theory by spreadsheet', 'Teaching stochastic calculus to the 3rd year students' and 'Teaching statistics to the modern actuary'.

5.4. TOPIC 4. STATISTICS EDUCATION/TRAINING AND THE WORKPLACE. Convenors: Carol Joyce Blumberg and René H.M. Smulders.

Session 4A, *Making Statistical Consulting and Technical Co-operation More Effective was* organised by Jean-Louis Bodin and chaired by Bart Meganck. Koffi N'Guessan described the various training programs that ENSEA carries out for individuals mainly from French-speaking countries in Sub-Saharan Africa. Carmen Arribas discussed the training courses, seminars, workshops, etc. in Spanish, held both in Spain and in Latin American countries, for professionals working at National Statistical Offices, Central Banks, Ministries of Planning and the other institutions producing the official statistics of the countries in the region.

Session 4B*, The Role of National and International Statistics Organisations in Improving Statistical Knowledge in the Workplace* was organised and chaired by Marcel Van den Broecke. Bart Meganck focused on the learning of statistics in the context of Eurostat's role in developing statistical programmes and the transfer of statistical knowledge for the implementation of monetary union in Europe. The paper by Madge Haven and Madhuri Mulekar presented by Carol Joyce Blumberg described the ASA's publications in the area of statistics education, its activities at the primary and secondary school levels, its role in statistics education at the post-secondary level, and its varied continuing education activities. Daniel Berze gave a historical overview from 1945 until the 1991 creation of the IASE, with particular reference to the ISECs in Calcutta and Beirut and the activities of the ISI Statistical Education Committee.

Session 4C, *Training of Official Statisticians* was organised and chaired by Denis Farrell. Bradley Payne, Peter Holmes and Neville Davies described the development and implementation in Malawi of a course in Key Statistical Skills for clerks parallel to the Ordinary Certificate in Statistics of the UK's Royal Statistical Society. Pilar Martín-Guzmán commented on how the de-centralisation of a country's statistical system, the globalisation and worldwide need of harmonisation of statistical systems, and the increasing number and variety of users influence the training of official statisticians. Ruslan Motoryn described how his University trains students in both the international and national standards of national accounts needed for a market economy.

Session 4D, *Distance Learning* was co-organised by Lea Bregar and Irena Ograjenšek (Slovenia) and chaired by Irena Ograjenšek. The first paper was 'Statistical Education and the Workplace: Present State of Affairs and Future Challenges' by Lea Bregar, Irena Ograjenšek and Mojca Bavdaž Kveder. The paper focused on the question how modern information and telecommunication technology, including distance learning, can increase the quality and efficiency of statistical training at the workplace from the learners' point of view. The second paper was 'Experiencing Statistics at a Distance' by W. Robert Stephenson (Iowa State University, USA). He discussed how a two-semester sequence of distance education courses in Applied Statistics for Industry for managers and engineers in the workplace incorporates the use of videotapes, practical experiences and other activities. The third paper was 'Just-in-Time Network-Based Statistical Learning: Tools Development and Implementation' by Lea Vermeire, An Carbonez, Paul Darius (Katholieke Universiteit Leuven, Belgium) and Jill Fresen (University of Pretoria, South Africa). The paper reported on their work in short-course and in-company training in statistics, with special attention given to a self-study course for a government department and a system of highly interactive applets for visualisation of statistical concepts related to the linear model. The paper 'From Online Learner to Online Teacher' by Sharon Copeland-Smith (Swinburne University of Technology and Multimedia, Australia) was not presented orally, but is in the Proceedings.

Session 4E-1, *The Use of Census Material in Statistics Teaching* and Session 4E-2, *The Interface Between Official Statistics and University Teaching* were organised and chaired by Sharleen Forbes (Statistics New Zealand, New Zealand). Session 4E-1 began with the presentation by Maria-Gabriella Ottaviani (University of Rome "La Sapienza", Italy) of the paper 'The Italian Census at School' by Cristiana Conti (Italian National Statistical Institute (ISTAT), Italy) and Enzo Lombardo (University of Rome "La Sapienza", Italy) and the presentation by Lesley Hooper (Statistics New Zealand, New Zealand) of her paper 'Making Census Count in the Classroom'. Both papers briefly discussed the history and mechanics of the Censuses of school children carried out in their respective countries. They also described the various materials and websites that the national statistics offices developed for use by teachers and students relating to Censuses. The Session continued by having the audience ask questions of a panel chaired by Lesley Hooper and consisting of Maria-Gabriella Ottaviani, Sharleen Forbes, Reinie Cordier (Statistics South Africa, South Africa) and Doreen Connor (The Nottingham Trent University, England), who have all been involved in Censuses involving children.

The first paper in Session 4E-2 was 'The Use of Official Statistics in Teaching University Geography Students in Italy' by Enrica Aureli and Riccardo Russo (University of Rome "La Sapienza", Italy). This paper discussed the statistical methods and tools taught at the upper level in the geographical disciplines at Italian universities and described some selected postgraduate courses. The second paper was 'Teaching Official Statistics in an Irish University Statistics Department' by Patrick Murphy (University College Dublin, Ireland). His paper outlined the development of a new course in Official Statistics for undergraduate students at his University that can easily be adapted for use elsewhere. The final paper in this session was 'Development of Customer Oriented Learning Environment at Statistics Finland' by Reija Helenius (Statistics Finland, Finland). This paper described how training in the use of statistical information for users from various governmental and non-profit agencies is continuously being developed by Statistics Finland and discussed, as an example, Statistics Finland's web-based-learning project.

Session 4F, *Statistical Training and Education of Lawyers, Judges, Doctors, Researchers, and Other Professionals* was organised by Elisabeth Svensson (Örebro University, Sweden). The only paper presented orally from this session was 'Cramming for Court: Teaching Statistics to Litigators' by Mary Gray (American University, USA), whose actual presentation was combined with those in Session 4H. The paper gave guidelines for statisticians, who are preparing lawyers to effectively present and to cross-examine witnesses presenting statistical evidence during court trials, based on her experiences as a lawyer and as a statistician. The papers 'Statistical Training for Doctors in the UK' by Michael Campbell (University of Sheffield, UK) and 'Teaching Statistics to Medical Doctors through Research Methods: A Case of Medical Education Research in Iran' by Abbas Bazargan (University of Tehran, Iran) are available on the Proceedings CD.

Session 4G, *Preparation and Training of Workers in the 21st Century* was organised and chaired by Albert Shulte (recently retired from Oakland County (Michigan) Schools, USA). The first paper was 'Statistical Education and Training for Workers of the Public Administrations: Objectives, Issues, Strategies' by Luigi Biggeri (Italian National Statistical Institute (ISTAT), Italy) and Alberto Zuliani (University of Rome "La Sapienza", Italy). The paper focused on (i) the need for quantitative skills for public administration management and personnel; (ii) how to define the objectives of the education and a plan of training; and (iii) the strategies, issues and evaluation of some specific experiences. The second paper was 'Preparing Workers for the 21st Century: The Importance of Statistical Competencies' by Beverley Carlson (United Nations Economic Commission for Latin America and the Caribbean (ECLAC), Chile). This paper discussed the need to generate more meaningful statistics about the workings of the rapidly changing labour market and the interplay between the supply of skilled manpower and the statistical competencies required by employers, and how to use of this information to improve the teaching of statistics at all levels. The third paper was 'Statistics – Driving Success or Blocking the Road?' by Stephen A. Zayac (Ford Motor Company, USA) and a summary of it was given by Albert Shulte. This paper discussed statistical training needs from the viewpoint of someone in industry. The final paper in this session was 'Would you Allow Your Accountant to Perform Surgery? Implications for Education of Primary Teachers' by Lionel Pereira-Mendoza (National Institute of Education, Singapore). This paper discussed the issues related to statistical knowledge as it applies to primary teachers, since most teacher education programmes for primary teachers include mathematics education courses, but do not specifically address statistical education.

Session 4H, *Educating Managers, Executives, Politicians, Government Officials and Other Decision Makers* was organised by Theodore Chadjipadelis (Aristotle University, Greece) and chaired by Elisabeth Svensson (Örebro University, Sweden). The first paper was 'Using Statistics as a Tool in Political Research: The Case of Electoral Behavior' by Theodore Chadjipadelis (Aristotle University, Greece). This paper discussed the problems of using statistical techniques in the political sciences by following and commenting on the phases of observation of electoral behaviour. The second paper was 'Statistics Education for Future Managers: Needs, Obstacles, Possible Solutions' by Corinne Hahn (ESCP-EAP and NEGOCIA, France) and Patrick Dassonville (ESCP-EAP, France). This paper gave examples of the types of on-the-job problems future managers are likely to meet and discussed the difficulties and some possible solutions when teaching statistics in management schools. The third paper was 'International Statistics for Public and Private Decision Makers: New Tools to Improve the OECD

Communication Policy' by Enrico Giovannini (Organisation for Economic Co-operation and Development (OECD), France). This presentation discussed OECD's "new vision" for its statistical activities, including the development of a new statistical information system and the system's implications for the training of private and public decision makers.

Session 4I, *Statistical Training and Education in Environmental Settings* was co-organised by María Virginia López and María del Carmen Fabrizio (University of Buenos Aires, Argentina) and chaired by María del Carmen Fabrizio. The first paper was 'Factors Affecting Performance in a University Service Course on Biostatistics: An Update' by John A. Harraway (University of Otago, New Zealand). His paper reported that there were no differences in marks received in university level Biostatistics between those students who had mathematics with calculus and those who had mathematics with statistics in secondary school, nor were differences found by gender. The second paper was 'Role of Statistics in the Education of Agricultural Science Students' by Katarina Cobanovic (University of Novi Sad, FR Yugoslavia). This paper described the author's experiences of teaching statistics in the Agricultural Faculty at her University with discussion of problems and dilemmas encountered and some solutions. The third paper was 'The Status of Statistics in Agricultural Studies: An Epistemological Approach' by María Virginia López, María del Carmen Fabrizio, María Cristina Plencovich, and Hernán Giorgini (University of Buenos Aires, Argentina). This paper described the results of a survey of 23 Argentine universities with agriculture programs that obtained information, using various indicators, about the insertion of Statistics into the university programs and explored the epistemological ideas underlying in the teaching of Statistics in agricultural schools. This paper was a replacement at the co-convenors' request for a withdrawn paper. It is not in the *Proceedings* and can be obtained from María Virginia López at mvlopez@mail.agro.uba.ar.

Session 4J, *Practical Training in the Workplace for Tertiary and Postgraduate Students* was organised by Katherine Taylor Halvorsen (Smith College, USA) and chaired by John McKenzie (Babson College, USA). The first presentation was 'Internships for Undergraduate Statistics Majors: The BYU Experience' by Lara Wolfson (Brigham Young University, USA). She discussed how BYU has prepared students for internship experiences, how BYU has created an academic tie-in with students' internship experiences and various partnerships with employers (Note: This paper is not in the *Proceedings;* a copy of the PowerPoint presentation is available from ljwolfson@byu.edu). The second presentation was 'Statistics Education and Bulgarian Management Training Institutions Development Project' by Nadezhda Tsankova (University of Veliko Turnovo 'St. Cyril and St. Methodius', Bulgaria). In the paper she introduced the general aims and objectives, methodology, and specific outcomes of an investigation of students' opinions of the Project, with a focus on the "Statistics in Internet" distance-learning module. The third presentation was 'An Internship Program at a Liberal Arts College' by Katherine Taylor Halvorsen (Smith College, USA). She described the new internship program at Smith College, how students find internships, and how the College prepares students for internships (Note: This paper does not appear in the Proceedings; a copy of the PowerPoint presentation is available from khalvors@email.smith.edu).

Session 4K, *Training of Institutional Research Professionals* was organised and chaired by Amanda Lourens (Technikon Pretoria, South Africa) as President of the South African Statistical Association (SASA) and as Vice-Chairperson of the South African Association of Institutional Research (SAAIR). The first paper was 'Training Institutional Research Professionals' by Gerald W. McLaughlin (DePaul University, USA) and Josetta S. McLaughlin (Roosevelt University, USA). Their paper focussed on training, with specific attention to the various roles of an institutional research professional, the statistical and analytical tools used to perform tasks, and the need to teach others to use and interpret statistical results. The second paper was 'Promoting Statistics Literacy: New Opportunities for the Training of Institutional Research Professionals' by Linda Hewitt (Centre for Interdisciplinary Research and Development, Trinidad and Tobago). This paper examined the existing functions and areas of operations with respect to institutional research, as well as the challenges surrounding the new and emerging demands for statistics and indicators in the 15 member states of the Caribbean Community (CARICOM) region. The third paper was 'Training Institutional Research Professionals: Teaching a Statistics Course in Six Hours or Fewer, The Art of Teaching' by Mary Ann Coughlin (Springfield College, USA). This paper focussed on the difficulties that are associated with teaching statistical content and skills in professional development settings and discussed various pedagogical approaches designed to increase statistical understanding. The final paper was 'Intriguing Facets of Institution Research' by Pieter J. Vermeulen (University of Pretoria, South Africa). This paper described how the strategies and solutions to problems encountered in institutional research differ substantially from those of the past, mainly due to the advancement of technology.

Session 4L, *Statistics Education and the Workplace: A Challenge for All (A Discussion and Informal Gathering)* was co-organised by Carol Joyce Blumberg (Winona State University, USA), Daniel Berze (ISI Permanent Office, The Netherlands), and René Smulders (Statistics Netherlands, The Netherlands) and chaired by Carol Blumberg. The session began with 15 minutes of informal gathering. Sharleen Forbes (Statistics New Zealand, New Zealand) and Enrico Giovannini (OECD, France) then gave brief introductory remarks about the types of relationships needed in the future between statisticians in academic institutions, official statisticians, and statisticians in the private sector. Copies of their power-point presentations are available from Carol Blumberg at cblumberg@winona.edu. The remaining 40 minutes of the session were spent in audience discussion, with many ideas generated for as to how individuals, as well IASE and ISI, can help foster interaction between statisticians in academic institutions and those in the public and private sectors. One suggestion is already being implemented by IASE as part of the International Statistical Literacy Project.

5.5. TOPIC 5. STATISTICS EDUCATION AND THE WIDER SOCIETY. Convenor: Helen MacGillivray

It is reported that H.G. Wells once said that "Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write." The statistical sciences provide the underpinning for the analysis and communication of quantitative information involving variation, across all areas of society. Possibly more than for any other discipline the interaction between statistical developers, users and recipients needs to closely influence and be influenced by, statistical education at all levels, recognising that life-long learning is not just a cliché in statistics and that statistical understanding is a key enabler across modern society.

The sessions in this topic aimed to explore the multi-faceted interactions between statistical education and the roles of statistics in key aspects of society. It is both a strength and weakness of the statistical profession that statisticians work in highly diverse and dispersed areas and workplaces, working with professionals in many and varied other fields. The challenges of this for the professional societies, continuing professional development, statistics education and accreditation, formed the theme of session 5D, organised by Neville Davies. Derek Pike's cogently argued and persuasive paper on the importance and relevance of the continuing professional development process to ensure maintenance of professional standards, also sought to foster ongoing constructive debate between academic and commercial statisticians. Brian Phillips' paper on the roles of the IASE was complemented by the paper (Neville Hunt, Flavia Jolliffe, Neville Davies) on the role of the Royal Statistical Society in shaping statistics education in the UK and beyond, while Pali Lehohla's paper on a South African perspective on promoting statistical literacy illustrated society's need for broad-based statistics awareness raising programs.

The impact of information technology on statistics and statistics education has already been enormous, continues to grow, and takes many forms. Software and visualisation are integral to both the practice of, and education in, statistics. Mike Fuller's paper in Session 5F, organised by Gianfranco Galmacci, moved from this to considering the extent to which statistics curricula can adapt to changing opportunities generated by ICT, and the potential for use of statistical ideas in computer science education. Ewan Crawford and Adrian Bowman's paper (presented by Ewan) gave valuable insights to the extent to which ICT can support resources, networking and information sharing in statistics education, discussing the range of associated organisational challenges and opportunities.

The world of health and life sciences (Session 5G, organised by Petter Laake) also has many challenges and opportunities for statisticians and statistics education, in both traditional and emerging fields. This was reflected by the variety of papers in this session, and their emphasis on close interaction with user areas. Bradley Payne, Nick Merryfield and David Griffiths reported on the first ever survey of the UK medical sales field force, demonstrating a new interrogation tool to facilitate analysis. Nibia Aires considered the advantages and disadvantages of teaching classic methods in the post-genomic era. Ivar Heuch discussed the balances required in statistics training that are taken into the practical challenges of epidemiological data, and Penelope Pekow reviewed experiences in teaching biostatistics within an exchange program between the Medical University of South Africa and the University of Massachusetts.

Like mathematics, statistics has the qualities and duties of transferability and enablement. And mathematics itself across all its levels, is an enabler for statistical understanding, development and hence education. In session 5A, organised by Brian Greer, Jeff Evans discussed how tracing the development of concepts of affect and emotion in mathematics education research is informative for research on teaching statistics. The other three papers considered the interaction between aspects of mathematics and statistics education. Michael Bulmer discussed the development of a computer-based version of concept maps in teaching statistics, giving an interactive concept map with a narrative. Joachim Engel considered activities incorporating statistical concepts and mathematical foundations, and Jerry Moreno presented some insight into the NSF-funded project Data-Driven Mathematics which motivates mathematics topics from a data point of view.

The papers in Session 5B, organised by David Griffiths, illustrated the question of whether statistics education is a discipline in itself or in association with the contexts of user areas. Nye John and David Johnson's paper on teaching managers to think statistically, Dennis Pearl's paper on using health science examples to create statistical understanding for a diverse group of students in a variety of majors, and Brian Greer's paper on teaching the essential rationale of statistical methods in psychology, led to a brief but lively discussion of experiences across the areas.

In the wide worlds of business, government and engineering, statistical education at all levels is not only increasingly important but also benefits greatly and often unexpectedly from constant interaction with statistical usage and real problems. Session 5C, organised by Eric Sowey, considered just some aspects of the business and government sectors. Richard Madden's paper (presented by Ching Choi) used topical examples from official statistical agencies to draw some important messages for teaching statistics in practical and ethical issues. Jonathan Cryer reviewed the effects of the 17 annual US Conferences to date on Making Statistics More Effective in Schools and Business, and Gerald Goodall and Derek Pike focussed on defining relevant continuing professional development objectives and materials in the diverse areas of business and government.

All speakers in Session 5D, organised by Helen MacGillivray, emphasized the importance of statistics in the engineering world and the challenges of both engaging the students and facilitating their development of both statistical literacy and specific statistical skills under the pressures of an engineering course. Richard Wilson's and Stephen Vardeman's papers both focussed on the need to connect with engineering environments and thinking whether in a first or second course. The student perspective from James Moody, Australia's Young Engineer of the Year, emphasized the dual needs of general and specific statistical underpinning for engineering students, and Neil Diamond discussed how to distil valuable teaching tools from real and sometimes complex consulting problems.

Session 5H, organised by Jacky Galpin, after hearing from Eric Sowey on educating statisticians to enhance their future clients' statistical understanding, from Jacky on the integrating social issues such as HIV/AIDS, and from Swapna Mukhopadhyay on connecting with socially sensitive and important issues such as gun violence, incorporated a most interesting and lively audience and panel discussion on social issues and statistics education. In many ways, this discussion epitomised the theme of Topic 5, with all speakers contributing to a rich variety of aspects of the challenges, opportunities and needs in the interaction of statistics education with the wide areas of society it underpins, serves, and is enriched by.

5.6. TOPIC 6. RESEARCH IN STATISTICS EDUCATION. Convenors: *Joan B. Garfield and Carmen Batanero*

Research in statistics education is a priority area for IASE, since a well-developed research literature and research agenda are essential for promoting the field of statistics education. In this particular ICOTS meeting we paid particular attention to the development of statistics education research, examining current research problems, directions for future areas of inquiry, and the use of theoretical models on which our research might be based. In addition to including sessions on teaching and learning statistics and probability we focused sessions on research related to assessing student outcomes, the role of technology in learning statistics as well as in conducting research on student learning, and on the conceptions and beliefs of statistics teachers.

Session 6A, *Developing Statistics Education Research* was organised by Mike Shaughnessy (USA). Presentations dealt with 'Creating cognitive conflict in a controlled research setting: sampling' (Jane Watson), 'Students' individual and collective statistical thinking' (Edward Mooney and Cynthia Langrall), 'Assessing and tracing the development of Basotho elementary students' growth in probabilistic thinking' (Victor Polaki), 'Framework for teacher knowledge and understanding about probability' (Tova Kvatinsky), and 'Students' understanding of variability in a probability environment' (Mike Shaughnessy).

Session 6B, *Theoretical Models of Statistical Knowledge, Thinking, Reasoning and Learning* was organised by Maxine Pfannkuch (NZ) and Chris Wild (NZ) and included the following presentations: 'Studying the median: a framework to analyse instructional processes in statistics education' (Juan Godino), 'The development of a framework characterizing middle school students' statistical thinking' (Cynthia Langrall and Edward S. Mooney), 'How students experience learning statistics and teaching' (Peter Petocz, Anna Reid) and 'Statistical thinking models' (Chris Wild).

Session 6C, *Technology and Research in Teaching and Learning Statistics* was organised by Erica Morris (UK) and chaired by Juan Godino (Spain). Presentations dealt with 'The role of computer based technology in developing understanding of the concept of sampling distribution' (Kay Lipson), 'The statistical re-education of Psychology' (Geoff Cumming), and 'Comparison of multimedia educational materials used in an introductory statistical methods course' (Richard Alldredge).

Session 6E, *Teachers' Training, Conceptions and Beliefs,* was organised by Lisbeth Cordani (Brazil). Papers on 'Probability and statistics in elementary school: a research of teachers' training', (Celi Espansadin), 'Teacher's training in a statistics teaching experiment' (Linda Gattuso) and 'Investigating the 'data sense' of preservice teachers' (Tim Burgess) were presented.

Session 6F, *Research into Teaching and Learning Statistics at Tertiary Levels* was organised by Flavia Jolliffe (UK) Presentations dealt with 'Modelling students' learning of introductory statistics' (Dirk Tempelaar), 'Choosing to study independently - when is it a bad idea?' (Glenda Francis), 'Evaluating the impact of multimedia lectures on student learning and attitudes' (Sterling C. Hilton), and 'Investigating patterns of interview conversations among lecturers in the Malaysian institutes of higher learning on the teaching of statistics at the introductory level' (Zamalia Mahmud).

Session 6G, *Research into Teaching and Learning Probability* was organised by Antonio Estepa (Spain). Presentations dealt with 'Teaching probability and statistics to 10 years old children' (Marie Berrondo), 'On the use of paradoxes in the teaching of probability' (Talma Leviatan) and 'Misconceptions in probability' (Lin Ju and Lionel Pereira-Mendoza).

Session 6H, *Round Table Discussion: Major Problems and Directions in Statistics Education* was organised Joan Garfield (USA) and Carmen Batanero (Spain) and chaired by Lisbeth Cordani (Brazil). Different graduate training programs for statistics education research were presented by Mike Shaughnessy and Carmen Batanero and debated by Gabriella Ottaviani.

5.7. TOPIC 7. TECHNOLOGY IN STATISTICS EDUCATION. Convenor: *Laurence Weldon*

This topic explored the way in which teaching and learning of statistics has changed as a result of modern technologies. In particular, animations based on Internet browsers which operate an many platforms have brought some unity to a field that has been quite disorganized. Moreover, the ease of distributing these animations to students for their interactive use has added interest and effectiveness to online education. Some products that were originally designed for distance education are now being used in the classroom - the effort needed to prepare these is great but widespread use makes them feasible. The sessions of this topic focused on these Internet technologies but will include other computer-based technologies as well.

Session 7A, *Java-Based Instructional Packages* was organised by Doug Stirling (New Zealand). Presentations dealt with 'Java applets and multimedia catalogues for statistics education', 'Applets for experimenting with statistical concepts' and ´Interactive content in web pages to teach statistics'.

Session 7B, *Computer-Based Demonstrations of Statistical Phenomena*, was organised by Andrej Blejec (Slovenia). Presentations were: 'Computer modules for teaching statistical concepts', 'Teaching statistical concepts with simulated data' and 'Hands-on survey research in a virtual environment'.

Sessions 7C, *Using Technology for Statistics Education in Engineering*, was organised by David Bacon (Canada) and there were discussions about 'Learning statistics in an engineering curriculum', 'Technology, statistical thinking and engineering students and 'Teaching experimental design to engineers: some experiences and advice'.

Session 7D, *Using Graphics Calculators in Statistics Education,* was organised by Kay Lipson (Australia) and included the following papers: 'Simulation as a tool to develop statistical understanding', 'Simulating experiments with the graphic calculator TI 83plus' and 'Teaching statistics with TI 83'.

Session 7E, *Statistics Education and the Internet*, was organised by Joe Wisenbaker (USA). Presentations dealt with 'St@tnet, an internet based software for teaching introductory statistics', 'News - groups and teaching statistics. Are they useful?' and 'A personal journey toward a virtual introductory statistics course: not (quite) ready for prime time'.

Sesion 7F, *Research-Based Design and Use of Software for Teaching Statistical Concepts* was organised by Cliff Konold (USA) and Bill Finzer (USA). Papers were 'Route-type and landscape-type software for learning statistical data analysis', 'Technology, statistics, and subtleties of measurement: bridging the gap between science and mathematics' and 'The Fathom experience—is research-based development of a commercial statistics learning environment possible?

Session 7G, *Software Tools Designed for Statistics Education* was organised by Rodney Carr (Australia). Presentations were: 'Using Excel to teach statistics in New Zealand secondary schools', 'A data analysis tool that organizes analysis by variable types', 'Using Fathom to promote interactive explorations of statistical concepts' and 'Live figures: interactive diagrams for statistical understanding'.

5.8. TOPIC 8. OTHER DETERMINANTS AND DEVELOPMENTS IN STATISTICS EDUCATION. Convenor: *Philip J. Boland*

Statistics is a very broad discipline encompassing so many areas of practical application. Don't we know that everyone needs to learn (and be taught) Statistics? The purpose of a topic on Other Determinants and Developments in Statistics Education at the ICOTS meetings is to include sessions which address special aspects of teaching statistics, but which do not necessarily fall under one of the main conference themes. Generally this topic is open to novel and sometimes-controversial sessions, and at ICOTS-6 there were five such sessions of a very high calibre.

Robert Delmas organised a session (8A) on *Learning Factors in Statistics Education*, although he was not able to actually attend the conference himself. In the session, Carl Lee made an interesting presentation on the issue of motivation and expectations in introductory statistics courses, while Verena Nolan discussed the influence of attitude, knowledge of English and mathematical ability in a course on quantitative techniques.

Kay McClain organised and spoke in a very interesting session (8B) on *Data Analysis and Statistical Learning*. Kay provided an analysis of a teacher development experiment, in which she concluded that the learning trajectory for the teacher's activity paralleled that of the students. Cliff Konold made an excellent presentation on how students use a "modal clump" in trying to express both the average and spread of a set of data. Katie Makar discussed the statistical thinking of teachers in analysing their own students' data.

Jerry Moreno chaired and organised a very interesting session (8E) on *Projects and Poster Competitions in Statistics Education.* A most interesting selection of posters from the American Statistical Association's annual poster competition was given by Linda Quinn. Loi Soh Loi discussed the impact of final year projects in a Singapore university business school. Saleha Naghmi Habibulluh spoke on her experience over many years in organising (national and international) statistical competitions and exhibitions in Pakistan. Susan Starkings gave a very nice presentation on the use of statistical projects as part of the secondary school curriculum in the UK.

Given the need to convey the importance of statistics to the general public, it was very appropriate to have Shen Shir Ming organise a session (8F) at this ICOTS-6 on *The Mass Media and Statistics*. Martin Podehl gave a fascinating presentation on the efforts of the National Statistical Office of Canada to co-ordinate with the news media on informing the public about social and economic issues. Yuen Ying Chan spoke on the role of statistics in journalism education.

Mbulaheni Nthangeni organised a session (8G) on *Teaching Statistics to Second-Language Students*, where Renette Blignaut and I. M. Vente gave a paper on Statistics Teaching Enhanced by Teamwork.

5.9. TOPIC 9. AN INTERNATIONAL PERSPECTIVE FOR STATISTICS EDUCATION. Convenor: *Vitalis Muba*

Session 9A, *Statistics Education in African Countries*, was organised by Fayez Mina (Egypt). Reda Mosad El-Said Asar presented 'An experimental approach for teaching statistics in the Egyptian schools', John W. Odhiambo spoke about the 'Teaching of statistics in Kenya'. Jules J.S. de Tibeiro presented the paper: 'Is it reasonable to teach statistics without probability or probability without statistics?' and Fayez M. Mina discussed 'Some features of future statistics education'.

In Session 9B, *Statistics Education in Spanish-speaking Countries*, organised by Teresita Teran (Argentina), Antonio Estepa summarised the state of stochastic education in the Ibero-American countries.

The following papers were presented in Session 9C, *Statistics Education in Asia*, organised by Ann-Lee Wang (Malaysia). Louisa Lam talked about the changes in the statistics syllabus and the way it is taught in schools in Hong Kong. She also discussed the implication of the way statistics is taught in schools. R. P. Suresh explained the academic background of students taking the Post-Graduate Diploma in Management in the Indian Institute of Management Kozhikode. He gave an illustration of how case studies may be used to teach probability concepts to these students. Y. Zhang gave a review of the development of statistical education at the tertiary level in China. The number of students taking the various types of statistics courses were touched on. He concluded by saying that statistical education is expanding in China.

5.10. CONTRIBUTED PAPERS AND POSTERS. Convenor: *Susan Starkings*

Over 50 interesting papers were submitted to the contributed paper section of the ICOTS 6 conference held in South Africa. The papers were grouped under the following headings:

- Teaching and Learning Statistics Using Electronic Media
- Concepts in Teaching Statistics
- Assessment in Statistics
- Statistics Education for Teachers
- Teaching Statistics at University
- Research into Teaching and Learning Statistics
- Teaching School Children Statistics

It is evident from the contributed sessions that authors have a great deal of knowledge and expertise in the area of statistical education. The diversification, of the papers presented, and the imaginative ways in which the authors have constructed these papers is commendable. Some interesting discussions should emerge as result of the papers presented.

The papers submitted advocated the teaching of statistics as a practical application that linked statistics to a student's everyday life. This entailed students collecting and generating data relevant to their daily interests and experiences and then using these data to construct and test hypothesis. Several authors provided research evidence to support new methods of teaching the subject. The use of technology now plays a prominent role, in the papers submitted, with interesting and novel ways of using this technology to enhance learning being demonstrated. Numerous practical examples were elucidated for every level of statistics being taught in educational establishments.

The common themes that emerged, from these papers, were the use of topical and relevant examples; that technology should be used as a tool for data analysis, and that the use of various instructional techniques is beneficial to both teachers and students alike. To sum up, contributors agreed that the learning of statistics should be achieved through doing real life practical problems that bring theory into practice.

POSTERS. Convenor: Andrew Dale

A total of 25 posters were on show during the conference – amongst them three posters by South African school children. The children presented their posters on the Tuesday and thoroughly enjoyed the excitement of discussing their posters with the many interested delegates. The posters presented by the children were the winning entries in local statistics poster competitions run at a few schools.

RECENT PUBLICATIONS

Recent publications from the Sixth International Conference on Teaching Statistics (ICOST6, Cape Town, 2002) and the Twenty Sixth Annual Conference of the International Group for the Psychology of Mathematics Education (PME26, Norwich, 2002) are not included in this summary because they are described elsewhere in this issue.

Batanero, C. (2002). Estadística y didáctica de la matemática: relaciones, problemas y aportaciones mutuas (Statistics and mathematics education: relationships, problems and mutual feedback). In C. Penalva, G. Torregrosa, and J. Valls (Eds.), *Aportaciones de la didáctica de la matemática a diferentes perfiles profesionales* (pp. 95-120). Universidad de Alicante.

In this work I remark on the interest of a didactical training for future statisticians. Then I analyse the components of this training and suggest a method and possible activities to carry out this training. These reflections are based on the University of Granada experience, where an optional course on Statistics Education was included in the recent Major in Statistical Sciences and Techniques. Our aim is to encourage other colleagues to organise similar courses within graduate official programmes or in postgraduate training.

Batanero, C., & Godino, J. D. (2002). *Estocástica y su didáctica para maestros* (Stochastics and stochastics education for primary teachers). Granada: Los autores.

This material was prepared to introduce basic ideas about statistics and probability, and to show didactical resources for teaching these themes at primary school level.

Bedwell, M. (2002). More happy returns to the birthday problem. *Teaching Statistics*, 24(2), 43-45.

This article shows how the birthday problem can be used to introduce the exponential distribution.

Bryce, G. R. (2002). Undergraduate statistics education: An introduction and review of selected literature. *Journal of Statistics Education*, 10(2). (http://www.amstat.org/publications/jse/v10n2/bryce.html)

A recent symposium on improving the Work Force of the Future: Opportunities in Undergraduate Statistics Education was held to focus attention on the importance of undergraduate statistics education. The symposium and the approval of curriculum guidelines for undergraduate degrees by the Board of Directors of the American Statistical Association have done much to define the current state of undergraduate education in statistics and suggest directions for improvement. This article summarizes the activities leading up to the symposium and provides a brief summary of six papers from the symposium that have been published. The article concludes with a discussion of some of the outstanding issues that remain to be addressed.

Chick, H. L. & Watson, J. M. (2001). Data representation and interpretation by primary school students working in groups. *Mathematics Education Research Journal*, 13, 91-111.

Twenty-seven grade 5/6 students, working in triads in a near-classroom environment, were video-taped as they considered a supplied data set over three 45-minute sessions. They were asked to hypothesize about associations in the data and represent these on a group poster. Each student was assigned to three categories: one for the observed level of interpreting the information provided in the data set, one for the observed level of representing the chosen data, and one for the degree of collaboration observed in the group. In addition, students were asked their views on the group work. Levels of interpretation and representation skills were related but there was no association of either with the degree of group collaboration. There was some association of degree of collaboration and students' views on group work. Descriptive aspects of the three characteristics—interpretation, representation and collaboration—are considered, as are implications of the results for future research and for the classroom.

Franklin, L. A. & Cooley, B. J. (2002). An experiential approach to integrating ANOVA concepts. *Journal of Statistics Education,* 10(1).

This paper presents a data set based on an industrial case study using design of experiments. The data set is pedagogically rich because it has a rather large total sample size from an industrial setting that naturally yields a large third order interaction term. The experiment is a 2^3 design and is initially presented with no replications. The sample size of the data is then doubled and the analysis repeated, comparing these results with previous results. The process is repeated until eight replications are available for each combination of factors and all parameters are estimated. With eight replications, the analysis shows all main effects and all interactions are statistically significant at the $\alpha = 0.05$ level. With smaller sample sizes, various main effects and interactions are not found to be statistically significant. Through this presentation the instructor can lead students in discussions about the effect of increased sample sizes, power, statistical significance (or insignificance), interaction terms, Type I and Type II errors as well as the importance and the role of the error term. In addition, students can manipulate the data set in a computer laboratory setting to illustrate many of the concepts inherent in the design of experiments and analysis of variance.

Gallese, E., Antonia, J., Molina, G., Alvarez, E., et al. (2000). Problematics of the teaching and learning of statistics in the Accountants career. In the *Minutes of the XXI University Seminary of Accounting,* Buenos Aires, Argentina.

In this study we will try to show the dynamics that the set up and the fulfillment experimented in the project of investigation of the teaching and learning of statistics in non statistics careers. We will give a vision that goes from the initial project, that dates back to January, 1999, passing through the modifications that it has undergone, the partial results obtained through the analysis of cases, the deepening of the current conditions of this problematic documented by specialized bibliography, till the transference to the medium in congresses, courses and specific seminaries.

Gallese, E., Antonia, J., Molina, G., Alvarez, E., et al. (2001). Improvement of the quality in the statistical education. In the *Minutes of the Sixth Seminary of Investigations in the Economic Sciences and Statistics College*, (November), Rosario, Argentina. ISSN: It 1666-3543.

In this presentation we expose some results of a deep bibliographical investigation such as: the recommendations of the joint Committee of the American Statistical Association (ASA) and of the Mathematical Association of America (MAA) on curriculum, as well as other recommendations of the International Association of the Teaching of the Statistics (IASE); the grounds for the change in the Accountants career that are debated worldwide. Also some conclusions obtained from field investigations are exposed. In order to find out what statistics the professionals apply in the current society, we present the results of a survey carried out by *probabilistic sampling* of the professionals of economical sciences. Also we went to companies and by means of the technique of cases we obtained information about what statistics the companies need and we give some recommendations of the contents that should be present in the statistic programs.

Gallese, E. & Ferreri, N. (2002). Meditations on the course of statistics for non statistic professionals. *Latin-American Magazine of Investigation in Educational Mathematics,* (accepted for publishing).

What class of statistics do we teach to our students of non-statistics careers? How do we do it? How do we evaluate their learning? These and other questions are subjects of debates, seminaries, publications and congresses in many countries at the present. It is already being spoken of " the arrival of the Age of the Statistical Education ". The opinions about the courses for non-statisticians are varied but many common elements appear: to abandon the traditional class and give more participation to the students through activities and projects, to emphasize on the fundamental concepts and not on the techniques or formulas, to reduce the space dedicated to probability, to work with real data...Probably, the course of statistics that these students will take will be the only one in their lives. Therefore, we have to teach them what Statistics is and convince them of its importance in the world. Are we conscious of our responsibility? In the present study, a summary of the opinion of different authors, on the courses for non-statisticians, is exposed; but also it is sought to establish a question: what do the statisticians of Argentina think about these topics?

Gallese, E., Lac Prugent, N. & Anido, M. (2001). Integral education facing the changes, in *Integral education in a scientific and technical world,* ICALA Editions, Rio Cuarto, pp. 321-327. ISBN: 950-99863-7-2.

This work treats on the Applied Econometrics Seminary that has the main purpose of emphasizing the interaction between the production of data, scientific investigation and teaching, orientated towards the labour market. This Seminary is given at the Mathematics Laboratory of the Economic Sciences and Statistics College, of the National University of Rosario, since 1997.

Gallese, E., Lac Prugent, N. M. & Antoni, E. J. (2001). The social learning and the contribution of the statistics by means of computational tools, in the *Minutes of the VIII meeting of Chairs of Social and Humanistic Sciences for the Economic Sciences. Mar del Plata*, (June), Argentina.

The Social Science has two levels. In the first one the comprehension and the significance are directed to the main lines of social thought. The second level is that of the social contemporary facts. Having in mind that the reading of the reality will have to be based on the scientific records of its information, it will be necessary that the learning of the Statistics leave in the students good clarified concepts of reliability and validity of the conclusions. Also, the principal recommendations of the Fourth and Fifth International Conference on Education of the Statistics (Morocco, 1994 and Singapore, 1998) aim at, based on the socioeconomic reality of the country, that the students may have the possibility of taking part in the selection of the topic to be developed and to use the databases elaborated by the official organizations. It is recommended that the universities, in their new plans, should have in consideration the existing urgent need of stimulating an interaction between the data producers (the blue collars) and the investigators (the white collars). On this basis, we will present the foundations for the elaboration of an extracurricular program for teaching of the Statistics in the Social Sciences by means of computational tools.

Hindls, R., Hronova, S. & Stepan, J. (2001). University statistical education from the unifying Europa`s point of view. *Statistics in Transition,* 5(2), 249-257, ISSN1234-7655.

Statistical education at Czech universities began in the period between the two world wars. After the Second World War, this education reached a high degree of independence and a good professional position at international level. Nowadays, introductory statistics and statistical methods are compulsory in almost all non-statistical specialities at the Faculties of Economics and there is a specialisation in statistics at the Charles University and the University of Economics. New requirement in the European Union lead to modification of course contents, namely in the area of national accounting, price and industrial statistics, sampling surveys, time series, conjuncture analysis and computational statistics.

Ho Yu, C. Andrews, C., Winograd, D., Jannasch-Pennell, A. & DiGangi, S. A. (2002). Teaching factor analysis in terms of variable space and subject space using multimedia visualization. *Journal of Statistics Education*, 10(1).

There are many common misconceptions regarding factor analysis. For example, students do not know that vectors representing latent factors rotate in subject space, rather than in variable space. Consequently, eigenvectors are misunderstood as regression lines, and data points representing variables are misperceived as data points depicting observations. The topic of subject space is omitted by many statistics textbooks, and indeed it is a very difficult concept to illustrate. An animated tutorial was developed in attempt to alleviate this problem. Since the target audience is intermediate statistics students who are familiar with regression, regression in variable space is used as an analogy to lead learners into factor analysis in subject space. At the end we apply the Gabriel biplot to combine the two spaces. Findings from textbook review, survey and the "think aloud" protocol were taken into account during the program development and are discussed here.

Kennedy, P. E. (2002). More on Venn diagrams for regression. *Journal of Statistics Education,* 10(1).

A Venn diagram capable of expositing results relating to bias and variance of coefficient estimates in multiple regression analysis is presented, along with suggestions for how it can be used in teaching. In contrast to similar Venn diagrams used for portraying results associated with the coefficient of determination, its pedagogical value is not compromised in the presence of suppressor variables.

Mills, J. D. (2002). Using computer simulation methods to teach statistics: A review of the literature. *Journal of Statistics Education,* 10(1).

The teaching and learning of statistics has impacted the curriculum in elementary, secondary, and post-secondary education. Because of this growing movement to expand and include statistics into all levels of education, there is also a considerable interest in employing effective instructional methods, especially for statistics concepts that tend to be very difficult or abstract. Researchers have recommended using computer simulation methods (CSMs) to teach these concepts; however, a review of the literature reveals very little empirical research to support the recommendations. The purpose of this paper is to summarize and critically evaluate the literature on how CSMs are used in the statistics classroom and its potential impact on student achievement. The recommendation is that more empirically and theoretically grounded research studies are needed to determine if these methods improve student learning.

Moreno, A. & Vallecillos, A. (2001). Exploratory study on inferentials concepts's learning in secondary level in Spain. In: M. Van der Heuvel-Panhuizen (Ed.): *Proceedings of the 25th Conference of the International Group for the Psychology of the Mathematics Education*, p. 343. The Netherlands: Freudenthal Institute, Faculty of Mathematics and Computer Science, Utrech University.

In this paper we in summary communicate the first results obtained in an exploratory study on the learning of statistical inference basic concepts for Spanish students in secondary level. This is a part of an ongoing research project, with the aim to determine problematic areas, the students' difficulties, favourable learning environments or methodological questions that can serve like base to improve the teaching of the statistical inference in the secondary level. The taking of data has been carried out in two courses of secondary level, in total 49 students of different age and previous statistical formation. The students have answered a written questionnaire whose content refers to inferential basic concepts as those of population and samples, the influence of the type and size of the sample in the realisation of inferences, the different sampling types role in the inferential process as well as the implications of the possible biases in the obtaining of data. The enunciates are posed in three different contexts, concrete, narrative and numeric. The analysis of the answers has been, fundamentally, of qualitative type. First, each one of the items separately; later, each researched conceptual nucleus, grouping for it all the results of the items that refer to the same one. As it is an exploratory study we intend especially to determine the open questions and to formulate queries to research later more than to reach accelerated conclusions. Nevertheless, we have obtained some first interesting results: a categorisation of answers about concepts, only seemingly easy, as those of population and samples; differences among the students' conceptions about the sampling process in connection with the age and the context in that the question it posed; appreciation lack about the importance of the randomness in the taking of data or the influence of the sampling type. All these questions have a certain social importance, given their incidence in the taking of decisions in situation of uncertainty for all the citizens, therefore, the improvement in the basic education in these aspects is so much an obligation of the educational system as of the teachers and the educational research.

Rangecroft, M. (2002). The language of statistics. *Teaching Statistics*, 24(2), 34-37.

The discussion of problems associated with the use of language, specifically vocabulary and symbolism, is extended from the teaching and learning of mathematics to particular areas of statistics.

Sommerfield, J. T. (2002). The Binomial and Hypergeometric probability distributions in jury selection. *Teaching Statistics*, 24(2), 38-42.

This article considers the composition of juries, asking whether this is representative of the general population from which juries were drawn. The binomial and hypergeometric distributions are used for probability calculations. Several example applications of both of these distributions are given, addressing racial, sex and age distributions in various cases.

Truran, J. & Arnold, A. (2002). Using consulting for teaching elementary statistics. *Teaching Statistics*, 24(2), 46-48.

Consulting in Statistics is usually deferred until at least near the end of a first degree, but this article shows how some aspects can be effectively taught to students in upper secondary or early tertiary courses in a way which reinforces their learning of standard basic concepts. We suggest that the existence of a real client adds a degree of realism not available in other ways, and emphasizes to students the importance of blending statistical calculations with meaningful communication.

O'Connell, A. A. (2002). Student perceptions of assessment strategies in a multivariate statistics course. *Journal of Statistics Education,* 10(1).

The purpose of this paper is to illustrate the use of several assessment strategies in an advanced course in statistics, and present the results of student ratings for each assessment strategy in terms of difficulty, appropriateness, level of learning achieved, and preference. The assessment strategies used include structured data analysis assignments, open-ended data analysis assignments, reviews of applied research articles, and annotating computer output of multivariate procedures. Findings indicate that students "prefer" instructor-directed or structured assignments overall, but feel they learn the most when the assessment is unstructured and requires greater self-direction. Suggestions for incorporating these assessment strategies into the multivariate classroom, as well as examples of each strategy, are included in this study.

Vallecillos, A. (2000). Understanding of the logic of hypothesis testing amongst university students. *Journal für Mathematik-Didaktik*, 2, 101-123.

Significance testing is one of the most controversial subjects in research work (Morrison & Henkel, 1970) and also one of the most misunderstood topics in statistics learning (Brewer, 1986). In this paper, we present the whole results of a theoretical and experimental study concerning University students' understanding about the logic of statistical testing. The theoretical study discusses epistemological issues concerning Fisher's and Neyman-Pearson's approaches to hypotheses testing and their relationship with the problem of induction in experimental sciences.

Watson, J. M. (2001). Profiling teachers' competence and confidence to teach particular mathematics topics: The case of chance and data. *Journal of Mathematics Teacher Education*, 4, 305-337.

This paper presents an instrument for assessing teacher achievement and teacher need in relation to the chance and data part of the mathematics curriculum. The development of a profiling instrument, to be used with teachers, had two main objectives. The first was to assist in assessing teacher achievement in the context of proposals for the adoption of professional standards for mathematics teachers. The second was to assess professional development needs for teachers in the light of changes to the mathematics curriculum. The background for the development of the instrument is presented, followed by a description of the instrument and the results of responses to it from 43 Australian teachers. Uses for the instrument and further development possibilities are also discussed.

Watson, J. M. (2001). Longitudinal development of inferential reasoning by school students. *Educational Studies in Mathematics,* 47(3), 337-372.

This study follows an earlier study of school students' abilities to draw inferences when comparing two data sets presented in graphical form (Watson and Moritz, 1999). Forty-two students who were originally interviewed in grades 3 to 9, were subsequently interviewed either three or four years later. The results for individual student development add to the credibility of the cross-age observations, as well as support the hierarchical framework suggested by the original study. Changes in levels of performance and strategies for drawing conclusions are documented. A further step from the original study is the consideration of how students used the variation displayed in the graphical presentation of the data sets as a basis for decision-making. Implications for teaching and for further research are discussed.

Watson, J. M. & Chick, H. L. (2001). Does help help?: Collaboration during mathematical problem solving. *Hiroshima Journal of Mathematics Education*, 9, 33-73.

This paper considers the circumstances surrounding instances where help is sought and/or provided in collaborative problem-solving situations. Video-taped observations of nine groups of three grade 5/6 students working over three 45-minute periods on an open-ended data handling task allowed for the documentation of questions asked, responses provided, and outcomes achieved. Help associated with questioning was provided in two contexts: by students in response to student questions and by teachers through questions. Hierarchical levels were defined for questions, answers, and outcomes. It was then possible to compare the levels of each component of this sequence for student-initiated and teacher-initiated questions. As well, unsolicited help provided by students was documented. Student soliciting of help through questioning was not as effective as unsolicited help offered through the questioning of teachers.

Watson, J. M. & Chick, H. L. (2001). Factors influencing the outcomes of collaborative mathematics problem solving—An introduction. *Mathematical Thinking and Learning*, 3(2&3), 125-173.

This study is an investigation of the factors that influence the effectiveness of collaboration on open-ended mathematical tasks. Students in grades 3, 6, and 9 worked in groups of three on two chance and data tasks: one related to fair dice and the other related to associations among variables presented on data cards. The groups' outcomes and the types of collaboration observed are investigated in relation to issues raised in the literature. Various phenomena are identified that influence cognitive "lifting," "hovering," and "falling," that is, improvement, no change, and reduction in levels of functioning. These phenomena include cognitive factors, social or interpersonal factors, and external factors.

Watson, J. M. & Chick, H. L. (2001). A matter of perspective: Views of collaborative work in data handling. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education*, (pp. 407-414). Utrecht: Freudenthal Institute.

This paper reports on selected aspects of a study of students' collaborative group work in a grade 5/6 classroom on an open-ended task from the chance and data part of the mathematics curriculum. Of interest were four questions. What were the students' perceptions of the task in relation to their beliefs about the content of the mathematics curriculum? During the collaborative sessions, were the observations of student knowledge, learning, and task outcomes consistent with student understanding displayed in individual interviews made a week after the group work was completed? Were students' later accounts of events that took place in their groups consistent with what was recorded on videotape? Were students' beliefs about collaborative group work expressed in the interviews consistent with the actions observed during the collaborative group work?

Watson, J.M., & Mortiz, J.B. (2001).The role of cognitive conflict in developing students' understanding of chance measurement. In J Bobis, B Perry, & M Mitchelmore (Eds.) *Numeracy and beyond (Proceedings of the 24th Annual Conference of the Mathematics Education Research Group of Australasia)*, (pp. 523-530). Sydney: MERGA.

In individual interviews, twenty students in each of grades 3, 6, and 9 responded to a task involving chance measurement, then viewed video recordings of other students' conflicting responses and decided which response they preferred. Seven students improved their levels of reasoning and seven agreed with higher-level prompts without expressing reasoning. Only two students agreed at some point with lower-level prompts, and both reverted to the level of their original response in conclusion. Educational implications are noted.

Watson, J. M. & Moritz, J. B. (2001). Development of reasoning associated with pictographs: Representing, interpreting, and predicting. *Educational Studies in Mathematics,* 48 (1), 47-81.

A developmental model involving four response levels is proposed concerning how students arrange pictures to represent data in a pictograph, how they interpret these pictographs, and how they make predictions based on these pictographs. The model is exemplified by responses from three related interview-based studies. In Study 1, examples of each response level are provided from 48 preparatory- to tenth-grade students. Students from higher grades were more likely to respond at higher levels. In Study 2, 22 students were interviewed longitudinally after a three-year interval; many improved in response level over time, although a few responded at lower levels. In Study 3, 20 third-grade students were interviewed and then prompted with conflicting responses of other students on video; many improved their initial responses to higher levels after exposure to the conflicting prompts. Associations among levels of representing, interpreting, and predicting were explored. Educational implications are discussed concerning reasonable expectations of students and suggestions to develop these skills in students at different grades.

RECENT DISSERTATIONS

Bialas, P. J. (2001). *Spreadsheet use in an elementary statistics course.* EdD Columbia University Teachers College. Supervisor: Phillip J. Smith.

The purpose of the study was to investigate the effects of the spreadsheet on achievement in selected statistical topics and the effects on beliefs about statistics of undergraduate students in an elementary statistics course. This study was instituted as part of the investigator's effort to enhance the statistical experience of undergraduate students. The study sought answers to the following questions: Does the use of the spreadsheet affect students' achievement on every topic selected for the study? Is the level of previous computer experience of students related to their achievement on the topics selected for the study? Does the use of the spreadsheet affect students' beliefs about statistics? Does students' achievement on the topics taught with the spreadsheet approach differ from achievement on the topics taught without the spreadsheet? The investigator conducted the experiment with students in one class at the beginning of the Fall 1999 academic semester in a community college setting. The investigator selected and taught eight Elementary Statistics course topics. The selected eight topics were grouped into two categories of four: topics taught with no spreadsheets and topics presented to students with the aid of spreadsheet files. During class sessions, students used computer labs and the spreadsheet program, Excel 5 and/or Excel 98. The instructor developed the curricular units and the test instruments. Test gains show that the spreadsheet approach to instruction was positively related to student achievement on every topic selected for the study. In addition, students' achievement on tests of topics taught with the spreadsheet was greater than their achievement on tests of topics taught with no spreadsheet. The use of the spreadsheet files seemed to affect students' beliefs about statistics. The analysis of students' responses to the statements on the questionnaire indicated that students were more in agreement with the questionnaire's statements after its second administration at the end of the study than they were after the first administration of the questionnaire.

Cazorla, I.M. (2002). *A relação entre a habilidade viso-pictórica e o domínio de conceitos estatísticos na leitura de gráficos* (The relationships between visual-pictoric ability and understanding of statistical concepts in graph reading). PhD University of Campinas, Brazil. Supervisor: Márcia Regina Ferreira de Brito.

This thesis aims to analyze factors that determine success when reading statistical graphics, based on Krutetskii's mathematical ability theory and Pinker's graphical comprehension theory. 814 undergraduate students attending Statistical courses were investigated. Six instruments were used: a questionnaire; two attitudes scale towards Statistics and Mathematics; and mathematical, statistics, and verbal aptitude tests. Findings show that success when reading statistical graphics lie on the understanding of the statistical concept, level of knowledge of graphics, visual-pictorial ability, and gender. Male students show more positive attitudes, and higher scores at cognitive tests, except at the verbal aptitude test. Instruction also shows to play a significant role on the development of statistical and graphic abilities.

Condor, J. A. (2001). *Effects of computer coaching in metacognitively cued elementary statistics instruction*. PhD University of South Florida. Supervisors: James, A. White and Michael F. Chappell.

Technology-based problem-solving models are being successfully implemented in the mathematics curriculum. This study focused on enhancing problem-solving ability by supplementing traditional instruction in statistics with metacognitively-cued, computer-coached activities. The purposes of this study were to investigate the: (1) differences in ability to solve basic, statistical word problems when comparing a metacognitively-cued, computer-tool (MCCT) group to a metacognitively-cued, computer-coached (MCCC) group; (2) differences in metacognitive ability while solving basic, statistical word problems when comparing a MCCT group to a MCCC group; (3) relationship between problem-solving ability and metacognitive ability while solving basic statistical word problems. A sample of 120 community college, elementary statistics students was divided into four sections with a MCCT and a MCCC group at one time period and a MCCT and a MCCC group at a different time period. Treatments lasted eight weeks of a summer semester. Dependent variables were ability to solve basic statistical word problems as measured by a teacher-made test and ability to think metacognitively while solving the word problems, as measured by the Assessment of Cognition Monitoring Effectiveness (ACME) procedure. The students were also measured on the quality of their responses to written metacognitive cues while solving a basic statistical word problem before each of the exams during the experiment. The dependent variables were measured at five different times throughout the semester. It was expected that the metacognitively-cued, computer-coached groups would show the most improvement and metacognitively-cued, computer-tool groups would show the least improvement on all measures. The data analysis revealed that the apparent difference in problem-solving ability between the MCCT groups and the MCCC groups grew as the study progressed, achieving statistical significance at the last testing, with the MCCC groups being significantly higher at the last testing. The MCCC groups also demonstrated significant higher metacognitive-ability. In addition, significant correlations were found between problem-solving ability and metacognitive ability, ranging from .28 to .66. The presence of some significant teacher effects suggests that the effectiveness of coaching software may be affected by instructional strategy.

Estrada, A. (2002). *Análisis de las actitudes y conocimientos estadísticos elementales en la formación del profesorado.* PhD Universidad Autónoma de Barcelona. Directores: Carmen Batanero y Jose María Fortuny.

En esta Tesis nos hemos interesado por el estudio de los conocimientos y las actitudes hacia la estadística de los profesores en formación, justificando el interés del estudio por la influencia de estas variables sobre la planificación de la acción educativa que cubra tanto los conocimientos matemáticos y didácticos como la educación de la afectividad del profesor.

El trabajo se ha llevado a cabo en dos fases. En la primera presentamos los resultados de pasar una escala de actitudes de construcción propia a una muestra de 66 profesores en ejercicio y 74 profesores en formación de educación primaria. No encontramos diferencias por género, aunque si una mejoría en aquellos alumnos que la estudiaron previamente. Contrariamente a nuestra hipótesis inicial, tampoco encontramos unas diferencias acusadas en las actitudes de profesores en formación y profesores en ejercicio, por lo que decidimos continuar el trabajo centrándonos exclusivamente en el primer grupo, que es sobre el que recae principalmente nuestra acción formativa.

Para la segunda fase del estudio, realizada con una muestra de 367 profesores en formación elegimos la escala SATS, después de realizar un comparativo de escalas de medición de actitudes. Asimismo utilizamos una parte del cuestionario Statistics Reasoning Assessment, construido por Konold y Garfield para obtener algunos datos sobre los conocimientos estadísticos de los profesores. Se confirma que la actitud de los profesores en formación es en general positiva independientemente del genero y de la especialidad. Aporta ademas una evaluación orientativa de los conocimientos estadísticos de los profesores en formación, encontrando errores conceptuales sobre los promedios, muestreo, valor atípico, así como en las interpretaciones de la probabilidad. En estos errores influye la especialidad cursada en los estudios de Magisterio. Se observa también una relación moderada entre actitudes y conocimientos estadísticos. Pensamos que esto sugiere la necesidad de apoyo desde los departamentos de didáctica y facultades de educación. Todos estos resultados se comparan con otros estudios previos de actitudes hacia la estadística.

La Memoria incluye también un estudio detallado de las investigaciones previas sobre actitudes hacia la estadística y errores en conceptos estadísticos elementales, así como un análisis del concepto actitud, sus componentes y medición y de las principales escalas de actitudes hacia la estadística.

Gunnarsson, C. L. (2001). *Student attitude and achievement in an online graduate statistics course.* EdD University of Cincinnati. Supervisor: Suzanne W. Soled.

The purpose of this study was twofold. First, the purpose was to design a web based graduate level statistics course for MBA students and to analyze the attitudes of the online students toward the course. The second purpose was to compare the students taking the course online versus the students taking the course in a traditional classroom setting. Achievement along with three mediating variables was investigated. The three mediating variables included: prior computer experience, prior math knowledge and experience and attitude toward the subject of statistics. The participants were forty-two graduate students in their first year of the MBA program, thirteen students took the class online, twenty-nine attended a traditional class. Students' attitudes toward learning in an online environment overall were favorable. Differences were found in the attitude toward the subject of statistics and prior computer experience; however, no casual relationship between class and achievement was detected. Students who learned in an online environment achieved comparably to students learning in a traditional classroom. The online course developed for this research can be used as an educationally equivalent managerial statistics course taught in a traditional classroom setting.

Scott, J. S. (2001). Modeling aspects of students' attitudes and performance in an undergraduate introductory statistics course. PhD University of Georgia. Supervisor: Joseph Wisenbaker.

This study examined the role of attitudes toward statistics, mathematics anxiety, mathematics attitude, mathematics background, demographic variables, and performance for students in an undergraduate introductory statistics course. The study participants were 155 students enrolled in five classes of introductory statistics at a four year college in metropolitan Atlanta. Using a self-selected ID to assure anonymity, the students completed the Survey of Attitudes Toward Statistics (SATS) at the beginning and end of the term. The SATS provides scale scores for Affect, Cognitive Competence, Value, and Difficulty. They also completed a mathematics attitude and anxiety measure, a demographic questionnaire, and a mathematics history. Students revealed their ID's after completion of the study. This allowed performance data from the course and prerequisite mathematics information to be linked with other student data. Students participating in this study had fairly positive attitudes concerning their Cognitive Competence and the Value of statistics at the beginning of the course. Their feeling of Affect was almost neutral and they expected the course to be somewhat difficult. Statistics attitudes were slightly less positive at the end of course. There were no statistically significant differences in attitudes between first time enrollees and those who were repeating the course or between students who did and did not complete the course. Pre-course SATS attitudes were generally not related to gender or age of the students nor to the years of high school mathematics or number of college mathematics courses. All of the SATS subscales were correlated with student grades in the prerequisite course. Pre-course Affect and Cognitive Competence scales were highly correlated to mathematics attitude, math self-concept and statistics self-confidence and moderately correlated with mathematics anxiety. Path analysis was used to develop a conceptual model for statistics attitude and performance in the course using mathematics attitude, mathematics anxiety, and prequisite grade as the exogeneous variables. In the path model, performance in the course was not influenced by either the pretest or posttest SATS. Performance during the statistics course did affect the posttest SATS scores.

Phyllis, H. (2001). The effects of using computer manipulatives in teaching probability concepts to elementary school students. EdD Columbia University Teachers College. Supervisor: Brice R. Vogelli.

The purpose of this study was to investigate the effectiveness of computer manipulatives compared to concrete manipulatives in teaching selected elementary probability topics. With the growing availability of computers in the classroom and the advancements in technological capabilities, computer manipulatives have the potential to have the same benefits of concrete manipulatives. It has been well documented that when used properly, concrete manipulatives benefit student's mathematical learning. Despite this fact, few teachers use concrete manipulatives because of classroom management issues. Several studies have shown that computer manipulatives are more manageable compared to their concrete counterparts and that computer manipulatives can facilitate students' mental operations better with the movements on screen. Thirteen fourth-grade students and two teachers were participants in the study. The students were separated into two groups of comparable ability. All students were to complete two activities which addressed nine probability-related target objectives. The first activity involved number cubes while the second involved spinners. When performing the number cube activity, half the students completed the activity using concrete cubes, the other half using computer cubes. To complete the second activity involving spinners, students who had used the computer number cubes for the first activity now used concrete spinners, and students who had used concrete number cubes for the first activity now used computer spinners. Students and teachers filled out a questionnaire and were interviewed at the completion of the study. Several comparisons showed that students using concrete manipulatives did just as well as those using computer manipulatives. Two out of four comparisons showed that students using concrete manipulatives scored better than those using computer manipulatives. Students and teachers reported that they enjoyed using computer manipulatives, and found them easy to use. Eight out of thirteen students saw no difference between the manipulatives with respect to their contribution to their learning, while about four out of thirteen students believed that concrete manipulatives were better for learning. Teachers did not change their belief that computer manipulatives are one of many tools that could be used to teach concepts however they reported that computer manipulatives will not replace concrete manipulatives.

INFORMATION ON PAST CONFERENCES

1. THE 26TH PME CONFERENCE, 21 - 26 JULY, 2002, NORWICH, UNITED KINGDOM

Reported by Chris Reading, creading@metz.une.edu.au

The following are abstracts of papers which were presented at PME26 and appear in:

Cockburn, A & Nardi, E, 2002, *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education*. University of East Anglia: United Kingdom.

1.1. REFEREED RESEARCH REPORTS, PRINTED IN FULL IN PROCEEDINGS

Afantiti Lamprianou, T. & Williams, J. A development scale for assessing probabilistic thinking and the tendency to use a representative heuristic. (Vol. 2, pp. 9-16).

We report a study of children's probability conceptions and misconceptions due to the representativeness heuristic. Rasch measurement methodology was used to develop a 13-item open response instrument with a sample (N=116) of 12-15 year olds. A hierarchy of responses at two levels is confirmed for this sample, and a third level is hypothesized. Each level is characterized by the ability to overcome typical 'representativeness' effects, namely 'recency', 'random-similarity' (at level 1), 'base-rate frequency' and 'sample size' (at level 2-3). Our interpretations were validated and anomalies identified through clinical interviews with children making the errors (n=8), suggesting another measure, which we named 'representativenss tendency' from 11 multiple choice errors.

Alatorre, S. A Framework for the study of intuitive answers to ratio-comparison (probability) tasks. (Vol. 2, pp. 33-40).

This paper describes a framework for the study of strategies used in ratio-comparison problems, which was constructed for the analysis of adults' responses to double urn probability tasks. This framework involves two systems, one for the interpretation and classification of answers (strategies), and one for the planning of numbers involved in ratio-comparison questions (situations). It was applied in an experiment with university students. Some results are reported, which refer to the relative occurrence of strategies, the difficulty levels of different situations and a classification of the subjects according to their performance.

Ginat, A. & Wolfson, M. On limited views of the mean as a point of balance. (Vol. 2, pp. 430-437).

It is clear to every high school student that the mean of two values is half their sum. Is it also clear that the mean is their point of balance? Not quite. In the course of studying pattern exploration disciplines of high school computer science majors, we noticed that a non-negligible number of students lack a clear view of the mean as a point of balance. The students were asked to design a computer program that inputs N, a positive integer, and outputs all the positive integer pairs <x,y> which average N/2. The majority of the students demonstrated limited orientation with patterns of the mean. In particular, a considerable number of them designed programs that "search for each x the y's which average N/2 with x". The student solutions, together with representative interviews, reflect diverse levels of mathematical insight and pattern recognition.

Kafoussi, S. Learning opportunities in a kindergarten about the concept of probability. (Vol. 3, pp. 161-168).

In this paper we describe the students' mathematical learning in a kindergarten during a classroom teaching experiment about the concept of probability. We present and analyze the learning opportunities that were created in the classroom as the children tried to resolve their problems, to reason mathematically and to communicate their thinking to others. The results of the research showed that kindergarten children made considerable progress in their probabilistic thinking, when they accepted the process of the experiment to check their different predictions as well as when they arrived at a consensus about the solution of a problem.

Lamprianou, I. & Afantiti Lamprianou, T. The nature of pupils' probabilistic thinking in primary schools in Cyprus (Vol. 3, pp. 273-280).

In this research work we explored the nature of 9-12 year old pupils' responses to probability problems. Analysis of pupils' arguments in 'Explain why' questions uncovered their thinking strategies, which we compared for pupils of different age and gender. The results revealed the existence of subjective elements and other errors in pupils' probabilistic thinking. The data were generated in year 2000 when the new mathematics books had just introduced probability extensively in the primary curriculum. Since the relevant literature in Cyprus is sparse, the results of the study form a general overview of the pupils' errors and build the basis for further in-depth and more focused research.

Monteiro, C. E. F. Investigating the interpretation of economics graphs. (Vol. 3, pp. 361-368).

Several studies in cognitive psychology have investigated the development of mathematical knowledge in cultural practices. For example, some authors have investigated situations where people deal with economic issues, which refer to quantitative relations and mathematical concepts. Print media use graphs to give information about economic topics e.g. variations of the rates in inflation, and wages. This study investigates the ways in which economists and business people, who subscribe to magazines and newspapers, interpret such graphs dealing with economics. The aim of the study was to identify how the background of the interviewees and the specific aspects of the graphs influenced the interpretative situation. Analysis of the data raised issues related to the teaching of graphing.

Nisbet, S. Representing numerical data: The influence of sample size. (Vol. 3, pp. 417-424).

Twenty secondary school students (in Grades 9 & 11) were given two datasets to represent graphically – one with 10 pieces of numerical data, and one with 30. Students were more likely to represent the large dataset in an organized form than the small dataset. The more mathematically able students found it easier to organise the data than their less able counterparts. Grade level had no effect. Possible explanations for the results are explored and the implications for teaching and the curriculum discussed.

Paparistodemou, E., Noss, R. & Pratt, D. Exploration of random mixtures in a 2D space continuum. (Vol. 4, pp. 33-40).

This paper focuses on how children express their ideas for randomness in two-dimensional continuous space, through tools for directing and redirecting the simulated movement of balls. It reports the findings of a study in which children aged between 6 and 8 years old engaged with a game-like environment to construct for themselves random behaviour by making special representations of sample space. In response to a range of tasks, the children manipulated the sample space in ways that generated corresponding outcomes in the game. We present some case studies of children's activities, which illustrate how the medium mediates the children's understanding of chance events.

Reading, C. Quantitative justification of SOLO cycles. (Vol. 4, pp. 97-104).

More and more research into student understanding is being undertaken, in which analysis of responses often involves the use of taxonomies to describe hierarchies. Recent use of the SOLO Taxonomy has identified that there can by cycles of levels within each mode in the hierarchy. Most recent research has involved qualitative description of these levels and identification of cycles based on subjective decisions. This report demonstrates how Threshold Values produced by Rasch Analysis can be used, as a quantitative measure, to justify the cycles of SOLO levels. The data used is from a study into statistical understanding, undertaken with Australian school students aged 12 to 18, in which the SOLO Taxonomy was used as the framework for developing a hierarchy of student responses to open-ended questions.

Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D. & Verschaffel, L. Secondary school students' illusion of linearity: Expanding the evidence towards probabilistic reasoning. (Vol. 4, pp. 305-311).

Many secondary school students have a strong tendency towards improper linear reasoning in the domain of geometry e.g. by believing that if the sides of a figure are doubled, the area is also doubled. In this paper, the evidence for this "illusion of linearity" is expanded to a new application domain: probabilistic reasoning. The paper reports an empirical investigation on the ability of 10th and 12th grade students to compare the probabilities of different situations. It is shown that most students have a good capability of comparing two events qualitatively, but at the same time incorrectly quantify this qualitative understanding into linear relationships between the varying quantities. It is shown how the research findings can shed a new light on some well-known probabilistic misconceptions.

Watson, J. M. & Kelly, B. A. Grade 5 students' appreciation of variation. (Vol.4, pp. 386-393).

This report focuses on one aspect of a larger study of school students' understanding of statistical variation. Although the study included students in grades 3, 5, 7, and 9, this paper will focus on grade 5 students only. Students experienced a unit of 10 lessons on the chance and data part of the mathematics curriculum conducted over an eight-week period. Lessons included a particular emphasis on variation and its role in statistical understanding. Pre- and post-tests were administered and improvements were found in overall performance and for variables reflecting appreciation of variation in chance, variation in data, and variation in sampling. Some comparisons are made with grade 3 students' performance.

Way, J. M. & Ayres, P. The instability of young students probability notions. (Vol.4, pp. 394-401).

This paper draws together the findings of a series of probability studies with students aged 11 to 14 years and incorporates new data not previously reported. The collective data emphasizes the fragility of probability notions and the strong effect of confirmation or refutation of outcome 'predictions' on probabilistic reasoning. Although some evidence of development of probabilistic reasoning with age was found, instruction appears to have a positive effect. The studies confirm the usefulness of videos of seemingly random experiments as research tools.

2. SHIRT ORAL PRESENTATIONS

Benson, C. T. Instructional strategies and tasks developed in a teaching experiment on probability modeling. (Vol.1, p. 293).

This study developed probability tasks and used them in a teaching experiment to investigate the impact of instruction on the modeling strategies of children in grades 3 and 4. Instructional strategies that were used are reported.

Cavallaro, M. I. & Anaya, M. Cognitive problems of university students with the concept of distribution. (Vol.1, p. 269).

The concept of distribution, and specifically the notion of Dirac's Delta Distribution, was found to be difficult for university students who deal with it in the advanced course of Mathematical Analysis. It is shown that these difficulties are related to mental models and schemes that the students have or develop at the moment when the concept of distributions is first presented to them.

Lann, A., & Falk, R. An average with unimaginative weights when the weights equal the values. (Vol.1, p. 288).

Our research focuses on learning and analysing people's intuitions: How do people intuitively fare in different 'self-weighted sampling' tasks (e.g. assessing the expected waiting time for a bus that arrives at varying intervals, or the mean class size obtained by questioning students)? Are there typical fallacies, and how could they be overcome?

Selva, A. C. V. & da Rocha Falcao, J. T. Understanding and using graphs: A didactic sequence for pre-school children. (Vol.1, p. 316).

We are particularly interested in investigating the bar graph comprehension among kindergarten in the exploration of additive-structure problems among six-year old children. We proposed a didactic sequence to children based on graphical representation of quantities derived from tridimensional histograms built with Lego-blocks.

Van Dooren, W., De Bock, D., Janssens, D. & Verschaffel, L. Broadening the scope of the illusion of linearity: A conceptual reanalysis of probabilistic misconceptions. (Vol.1, p. 273).

A summary communication of first results obtained in an exploratory study on the learning of statistical inference basic concepts for Spanish students in secondary level. This is part of an ongoing research project, with the aim to determine problematic areas, the students' difficulties, favourable learning environments or methodological questions that can serve like a base to the teaching of statistical inference in secondary schools.

Yañez, G., & Vallecillos, A. Some challenges of the use of computer simulations in probability teaching. (Vol.1, p. 330).

The first part of this work, besides looking for students' intuitive ideas about probability, looks for the spontaneous ideas they have about the frequentist approach and, in particular, the difficulties they face solving probability problems using computer simulation. Furthermore, we attempted to find the strategies they used for estimating the values of the requested probabilities.

FORTHCOMING CONFERENCES

IV Mathematics Education Conference in Venezuela, Trujillo - Venezuela, November, 12- 16, 2002

This conference is organised by the National Association in Mathematics Education (ASOVEMAT). The aims are to continue the previous Mathematics Education Conferences COVEM, and to foster mathematics education research and practice development. There is a thematic group in the Teaching of Probability, Statistics and Combinatorics. More information from ivcovem@yahoo.eshttp://es.geocities.com/ivcovem/

7th Asian Technology Conference in Mathematics, Multimedia University (Melaka Campus), Malaysia, 17-21 December, 2002

ATCM 2002 is to provide an interdisciplinary forum for teachers, researchers, educators and decision makers around the world in the fields of mathematics and mathematical sciences. It also provides a venue for researchers and developers of computer technology to present their results in using technology in both basic research and pedagogical research, and to exchange ideas and information in their latest developments. Further information: http://www.atcminc.com

Stochastic Thinking Group at the Third Conference of the European Society for Research in Mathematics Education (CERME 3), Bellaria , Italy, February 28th - March 3rd, 2003.

CERME is a conference organised by the European Society for Research in Education, and is designed to foster a communicative spirit. It deliberately and distinctively moves away from research presentations by individuals towards collaborative group work. Group 5 (Stochastic Thinking) will focus on stochastic thinking, including probability, statistics and the interface between these domains. Details of conference are given on the CERME3 web site http://fibonacci.dm.unipi.it/~didattica/CERME3/. More information available from Dave Pratt <dave.pratt@warwick.ac.uk>.

27 Congreso Nacional de Estadística e Investigación Operativa, Lérida, Spain, April, 8-11, 2003.

Statistics education will be included as a topic in the Annual Meeting of the Spanish Statistical Association and one plenary lecture will be devoted to statistics education. More information from the web page http://www.matematica.udl.es/seio2003/

V Simposio de Educacion Matematica, May, 5-9, 2003, Chivilcoy, Buenos Aires, Argentina.

This conference will be focused on "Research in Mathematics Education". It is organised by the Centro Regional Chivilcoy, Universidad Nacional de Luján, and sponsored by EDUMAT (Buenos Aires, República Argentina) and the Departament of Basic Sciences. More information is available from Oscar Luis Isnardi (institucionales@edumat.com.arthe). Web page: www.edumat.com.ar

Hawaii International Conference on Statistics and Related Fields. June 5-8, 2003

The main goal of the 2003 Hawaii International Conference on Statistics and Related Fields is to provide an opportunity for academicians and professionals from various Statistics and Related Fields from all over the world to come together to meet and learn from each other. The 2003 Hawaii International Conference on Statistics and Related Fields will provide a meeting place for academicians and professionals from related fields and with cross-disciplinary interests a chance to interact with each other from inside and outside their own particular disciplines, as well as for presenting research. Web page: http://www.hicstatistics.org/

The 2003 *Mathematics Education Research Group of Australasia* [MERGA] Conference, Deakin University, Geelong, Australia, July 6- 10

*Further information:*http://www.deakin.edu.au/fac_edu/numeracy_and_merino/merga/ or from the Chair of the organising committee: Judy Mousley (judym@deakin.edu.au)

IX Seminar on Applied Statistics: "Statistics in Education and Education in Statistics", Rio de Janeiro, Brazil, 7-10 July 2003. Interamerican Statistical Institute.

This seminar should provide an excellent opportunity for exchange of ideas, dissemination of recent work and developments that took place in Brazil and the Americas over the last few years, together with discussion of perspectives for advancement of both areas in the future. The goals are to attract wide participation from researchers, university teachers, and professionals, students and high-school teachers. More information is available from Pedro Luis do Nascimento Silva (pedrosilva@ibge.gov.br) or the web page. http://www.indec.mecon.gov.ar/newindec/proyectos/iasi_ingles/act_seminarios.htm

The PME 27th Conference will be held in Honolulu, Hawaii, at the Hawaii Convention Center, 13 - 18 July, 2003

You are cordially invited to participate in the Twenty-Seventh Annual Meeting of the International Group for the Psychology of Mathematics Education (PME) and the Twenty-Fifth Annual Meeting of PME-North American Chapter to be jointly held in Honolulu, Hawai'i on July 13-18, 2003.The conference will be held in the Hawai'i Convention Center. Website http://www.hawaii.edu/pme27. Coordinators of the Stochastical thinking learning and teaching Discussion Group are Christine Reading (creading@metz.une.edu.au) and Brian Greer.

Third International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-3), University of Nebraska, USA, July 23-28 of 2003

The third in a series of International Research Forums is to be held in the United States of America in July 2003. The Teachers College Institute and the Department of Educational Psychology at the University of Nebraska-Lincoln will host the Forum. This gathering offers an opportunity for a small, interdisciplinary group of researchers from around the world to meet for a few days to share their work, discuss important issues, and initiate collaborative projects. The topic of the third Forum will be *Reasoning about Variability*. This conference is supported by the International Association of Statistics Educators (IASE) and the American Statistical Association (ASA) Section on Statistical Education. More information from William T. Mickelson, <wmickelson2@unl.edu> or from the web site: http://tc.unl.edu/srtl

IASE Satellite Conference on Statistics Education and the Internet Max-Planck Institute for Human Development, Berlin, 11-12 August 2003

This conference will be organised by the IASE in cooperation with the Section on Stochastics of the German Mathematics Education Association, the Max-Planck-Institute for Human Development, and Stochastics Section of German Mathematical Association. The aim is to discuss the implications of the Internet for teaching and learning statistics: web based teaching, learning, materials and resources. An electronic proceedings and a limited number of printed copies of the set of papers presented will be produced after the conference. More information from: the Chair Larry Weldon (weldon@sfu.ca) or the Chair of the Local Organising Committee, Joachim Engel (JoaEngel@aol.com).

IASE IPM at ISI-54 Berlin, Germany, August 13-20, 2003

The list of IASE Invited Paper Meetings for Berlin is finalized. In addition to an exhibition "History of statistics and its teaching" organized by Maria-Gabriella Ottaviani mariagabriella.ottaviani@uniroma1.it, the following Invited Papers Meeting will be organized:

- IPM44 Teaching probability with a modelling approach. Organizer: Michel Henry (henry@math.univ-fcomte.fr)
- IPM45 Statistics training for consultants or collaborators. Organizer: Gabriella Belli (gbelli@vt.edu).
- IPM46 International co-operation in research on statistics education. Organizer: Lisbeth Cordani (lisbeth@maua.br).
- IPM47 Mathematics teachers teaching statistics. Organizer: Susan Starkings (starkisa@sbu.ac.uk).
- IPM48 Statistics education for media reports. Organizer: Maxine Pfannkuch (pfannkuc@math.auckland.ac.nz).
- IPM49 Teaching and learning approaches aimed at developing statistical reasoning, thinking or literacy Organizers: Joan Garfield (jbg@tc.umn.edu) and Dani Ben-Zvi (dani.ben-zvi@weizmann.ac.il) .
- IPM50 Statistics Teaching in the Internet Age. Organizer: Wolfgang Haerdle (haerdle@wiwi.hu-berlin.de).

- IPM68 Assessment of literacy, numeracy and other life skills. Organizer: Denise Lievesly (d.lievesly@unesco.org).
- IPM69 Impact of developments in information systems on statistics education (joint with IASC). Organizer: Annie Morin (Annie.Morin@irisa.fr).
- IPM70 Teaching biostatistics (joint with the International Biometrics Society). Organizer: Elisabeth Svensson (Sweden) (elisabeth.svensson@esa.oru.se) and Els Goetghebeur.
- IPM71 Educational implications of statistical method and modelling developments in psychometry. (Joint with the European Mathematical Psychology Group). Organizer: Helena Bacelar (Portugal) (hbacelar@fc.ul.pt) and Francesca Cristante.

Interested people are invited to contact the above mentioned organizers or the Chair of the IASE Programme Committee Gilberte Schuyten (Gilberte.Schuyten@rug.ac.be).

The 10th International Congress on Mathematics Education, July 4-11, 2004, Copenhagen

The congress will take place July 4-11 2004. The venue will be the Technical University of Denmark, located in a northern suburb of Copenhagen. The IASE will be collaborating in the organisation of specific statistics education activities in the conference. Chair International Programme Committee: Mogens Niss (ICME10-IPC@ruc.dk). Chair Local Organising Committee Morten Blomhøj (ICME10-LOC@ruc.dk). Conference web page: http://www.icme-10.dk/

IASE 2004 Research Round Table on Curricular Development in Statistics Education

The Round Table will be held in Lund, Sweden, and the dates will be coordinated with those of the Tenth International Congress on Mathematical Education, which will take place in Copenhagen, Denmark 4-11 July 2004.Lena Zetterqvist <lena@maths.lth.se> and Ulla Holt will be local organisers. **Those interested** can contact Gail Burrill, Division of Science and Mathematics Education, College of Natural Science, Michigan State University, 116 North Kedzie, East Lansing MI 48824, USA, E-mail: burrill@msu.edu.

IASE Activities at the 55th Session of the ISI, Sydney, Australia, 2005

Chris Wild is the IASE representative at the ISI Programme Co-ordinating Committee for ISI-55th Session, to be held in Sydney, Australia, April 5-12, 2005. As such he also is Chair of the IASE Programme Committee, which is in charge of preparing a list of Invited Paper Meetings to be organised by the IASE alone or in co-operation with other ISI Sections, Committees and sister societies. The committee will pay special attention to new topics that have been not discussed at the previous ISI Session. There is still time for you to propose a session theme for the IASE sessions for ISI55 in Sydney in 2005. Sessions that are of joint interest to IASE and another ISI section are also sought. Suggestions should normally include the name of the session organiser, a short description of the theme and an indicative list of possible speakers. Please email your proposals to Chris Wild at c.wild@auckland.ac.nz.

ICOTS-7, 2006 in Brazil

We are also glad to announce that the IASE Executive accepted the proposal made by the Brazilian Statistical Association to hold ICOTS-7 in 2006 in Brazil. This decision was announced at the ICOTS-6 farewell dinner. The proposal is also supported by the statistical associations in Argentina and Chile and we are already taking our first steps towards organising the conference. Pedro Morettin <pam@ime.usp.br> is the Chair of the Local Organising Committee and Lisbeth Cordani <lisbeth@maua.br> is acting as a link between the IASE Executive and the local organisers.

STATISTICS EDUCATION RESEARCH JOURNAL REFEREES JANUARY-NOVEMBER 2002

Acknowledgement: We are grateful to the following IASE members and researchers who acted as referees for Statistics Education Research Journal in the period January- November 2002:

Richard Allredge, USA
Andy Begg, New Zealand
George Bright, USA
George Cobb, USA
Lisbeth Cordani, Brazil
Jose Luis Cortina, USA
Joachim Engel, Germany
Candelaria Espinel, Spain
Sharleen Forbes, New Zealand
Michael Glencross, South Africa
Tjaart Imbos, the Netherlands
John Harraway, New Zealand
Michel Henry, France
Pilar Martín Guzmán, Spain
Robert delMas, USA
David Ospina, Columbia
Ernesto Sánchez, México
J. Michael Shaughnessy, USA
Luis Serrano, Spain
Gilberte Schuyten, Belgium
Elisabeth Svensson, Sweden
Peter Johnston-Wilder, UK
Joe Wisenbaker, USA