



Statistics Education Research Journal

Volume 3 Number 1 May 2004

Editors

Flavia Jolliffe
Iddo Gal

Assistant Editor

Christine Reading

Associate Editors

Carmen Batanero
Andrej Blejec
Joan B. Garfield
John Harraway
Annie Morin
M. Gabriella Ottaviani
Lionel Pereira-Mendoza
Maxine Pfannkuch
Mokaeane Polaki
Dave Pratt
Richard L. Scheaffer
Jane Watson
Chris Wild

International Association for Statistical Education
<http://www.stat.auckland.ac.nz/~iase>

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

Statistics Education Research Journal

The Statistics Education Research Journal is published 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;
- improve the teaching of statistics at all educational levels.

The Journal encourages the submission of quality papers, including research reports, theoretical or methodological analyses, and integrative literature surveys, that can advance scholarly knowledge, research methods, and educational practice in any of the broad areas related to statistical education or learning of statistics and probability at all educational levels and in all educational contexts. Contributions in English are preferred. Contributions in French and Spanish will also be considered. All papers are blind-refereed by at least two experts in the field.

Submissions

Manuscripts should be sent to co-editor Flavia Jolliffe (F.Jolliffe@kent.ac.uk), by email, as an attached document in Word format. Full details regarding submission are given in the Guidelines for Authors on the Journal's Web page: <http://www.stat.auckland.ac.nz/serj>.

© International Association for Statistical Education (IASE/ISI), May, 2004

Publication: IASE/ISI, Voorgurg, The Netherlands

Technical Production: University of New England, Armidale, NSW, Australia

ISSN: 1570-1824

International Association for Statistical Education

President: Chris Wild (New Zealand)

President-Elect: Gilberte Schuyten (Belgium)

Past-President: Carmen Batanero (Spain)

Vice-Presidents: Carol Joyce Blumberg (USA), Lisbeth Cordani (Brazil), Christine Reading (Australia), Susan Starkings (UK), Larry Weldon (Canada)

SERJ EDITORIAL BOARD

Editors

Flavia R. Jolliffe, Institute of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent, CT2 7NF, United Kingdom. Email: F.Jolliffe@kent.ac.uk

Iddo Gal, Department of Human Services, University of Haifa, Eshkol Tower, Room 718, Haifa 31905, Israel. Email: iddo@research.haifa.ac.il

Assistant Editor

Christine Reading, School of Education, Faculty of Education, Health and Professional Studies, University of New England, Armidale, NSW 2351, Australia. Email: creading@metz.une.edu.au

Associate Editors

Carmen Batanero, Departamento de Didáctica de las Matemáticas, Facultad de Ciencias de la Educación, Universidad de Granada, Granada 18071, Spain. Email: batanero@ugr.es

Andrej Blejec, National Institute of Biology, Vecna pot 111 POB 141, SI-1000 Ljubljana, Slovenia. Email: andrej.blejec@nib.si

Joan B. Garfield, Educational Psychology, 315 Burton Hall, 178 Pillsbury Drive, S.E., Minneapolis, MN 55455, USA. Email: jbg@umn.edu

John Harraway, Dept of Mathematics and Statistics, University of Otago, P.O.Box 56, Dunedin, New Zealand. Email: jharraway@maths.otago.ac.nz

Annie Morin, Institut de Recherche en Informatique et Systèmes Aléatoires, Université de Rennes 1, F35042 Rennes Cedex, France. Email: amorin@irisa.fr

M. Gabriella Ottaviani, Dipartimento di Statistica Probabilità e Statistiche Applicate, Università degli Studi di Roma "La Sapienza", P.le Aldo Moro, 5, 00185, Rome, Italy. Email: mariagabriella.ottaviani@uniroma1.it

Lionel Pereira-Mendoza, National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616. Email: lpereira@nie.edu.sg

Maxine Pfannkuch, Mathematics Education Unit, Department of Mathematics, The University of Auckland, Private Bag 92019, Auckland, New Zealand. Email: m.pfannkuch@auckland.ac.nz

Mokaeane Polaki, School of Education, National University of Lesotho, P.O. Box Roma 180, Lesotho. Email: mv.polaski@nul.ls

Dave Pratt, Centre for New Technologies Research in Education, Institute of Education, University of Warwick, Coventry CV4 7AL, United Kingdom. Email: dave.pratt@warwick.ac.uk

Richard L. Scheaffer, Department of Statistics, University of Florida, 907 NW 21 Terrace, Gainesville, FL 32603, USA. Email: scheaffe@stat.ufl.edu

Jane Watson, University of Tasmania, Private Bag 66, Hobart, Tasmania 7001, Australia. Email: Jane.Watson@utas.edu.au

Chris Wild, Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand. Email: wild@stat.auckland.ac.nz

TABLE OF CONTENTS

Editorial	2
New Associate Editors	3
Jose Carmona Márquez. <i>Una Revisión de las Evidencias de Fiabilidad y Validez de los Cuestionarios de Actitudes y Ansiedad Hacia la Estadística</i>	5
Sonia Kafoussi. <i>Can Kindergarten Children be Successfully Involved in Probabilistic Tasks?</i>	29
Sue Gordon. <i>Understanding Students' Experiences of Statistics in a Service Course</i>	40
María Virginia López, María del Carmen Fabrizio, Maria Cristina Plencovich and Hernan Giorgini. <i>Some Issues About the Status of Statistics Teaching in Agricultural Colleges in Argentina</i>	60
Paula R. Williamson and Gillian A. Lancaster. <i>Statistical Education for PhD Students in UK Medical Schools</i>	72
Forthcoming IASE Conferences	80
Other Forthcoming Conferences	86

EDITORIAL

Welcome to the May 2004 issue of SERJ, which is the first issue since Iddo Gal succeeded Carmen Batanero as co-editor with Flavia Jolliffe. This issue contains five papers. The paper by Jose Carmona Marquez, the second paper we have published in Spanish, reviews evidence for the reliability and validity of instruments assessing attitudes towards statistics and suggests priority research areas on this topic. Sue Gordon's paper explores the attitudes towards statistics of psychology students in Australia. Both these papers focus on dispositional aspects of learning statistics. Sonia Kafoussi's paper is one of a few papers published so far on probability work with very young children. Paula Williamson's paper is on the teaching of statistics to PhD students in medical schools in the UK, and the paper by Maria Virginia Lopez and colleagues is on statistics teaching in agricultural colleges in Argentina. These papers overall examine a variety of issues related to the learning, teaching, and content of statistics teaching in diverse cultural contexts.

Over the last few months, several important internal changes have occurred at SERJ. Since the journal continues to expand, and with an increasing number of submissions, we have felt it necessary to further expand the editorial board while retaining its international nature. Details of the seven new associate editors are given at the end of this editorial. We are very grateful that they have agreed to serve on the board. On the other hand, David Green (University of Loughborough, UK) has had to resign as an associate editor and we would like to thank him for his contribution to SERJ.

We have been changing the way we deal with submissions, with the intention of making working procedures more effective, and are in the process of updating the guidelines for authors and referees. The SERJ web page has moved and is now hosted by the University of Auckland as part of the IASE web page which carries archives of former newsletters as well as hundreds of papers presented at international conferences such as ICOTS-6 (Cape-Town, 2002), ISI-54 (Berlin, 2003), and other meetings. We thank IASE president and associate editor Chris Wild, and Rachel Cunliffe, the webmaster of the IASE website, for their efforts in this regard. We would also like to thank the University of New England, Armidale, Australia, for hosting the SERJ webpage and Chris Reading for developing the SERJ webpage and serving as its webmaster during the first two years of development.

As previously announced, our next issue in November 2004 will be devoted to the topic of reasoning about variability and variation. This Special Issue (Guest Editors: Joan Garfield and Dani Ben-Zvi) will include half a dozen papers first presented at the Third Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-3, Nebraska 2003), and then revised and submitted to SERJ for refereeing, as well as several introductory and reaction papers. In publishing this Special Issue, we will take advantage of the ability of an electronic journal such as SERJ to publish a substantial number of papers within a single issue. We will thus be able to offer our diverse audiences a broad coverage of a topic of central importance in statistics education, variability and variation, which has so far received little solid research attention.

FLAVIA JOLLIFFE AND IDDO GAL

NEW ASSOCIATE EDITORS

SERJ welcomes the following new Associate Editors, who have joined the Editorial Board for a 3-year appointment 2004-2007.

Andrej Blejec is Assistant Professor of Statistics and Computer Science in the Department of Biology at University of Ljubljana, Slovenia. He has more than 20 years of experience in teaching statistics for biologists and works as a statistical consultant at the National Institute for Biology. He is presently teaching undergraduate courses in Statistics and Computer Science and graduate courses in Biostatistics and Computational Statistics. He was a vice-president of the Slovenian Statistical Society (SSS) from 1997-2001, and is a president of the Section for Statistical Education of SSS. He has been a member of IASE since its foundation and serves as a National Correspondent for Slovenia. His main interest in statistical education is design and use of computer simulation systems for explanation of statistical concepts and use of computer technology for better teaching of statistics. At ICOTS6 he organized a session on the use of technology and computer simulations in statistics education. His current research is connected with application of stochastic processes in neuroscience and analysis of DNA micro-array experiments.

John Harraway is a Senior Lecturer in the Department of Mathematics and Statistics at the University of Otago, New Zealand. He has taught probability and statistics for over 25 years. This includes large classes in business, the biological sciences and biostatistics for students entering medicine, dentistry, pharmacy and physiotherapy. He was Chief Examiner in Statistics for the New Zealand Qualifications Authority from 1993 to 1997. He is author of two texts, *Regression Methods Applied* and *Introductory Statistical Methods for the Biological, Health and Social Sciences* (University of Otago Press, 1995 and 1997). In addition to consultancy which has led to recent publications on dolphin behaviour and habitat selection, his current research focuses on training researchers in the use of statistics. This has involved an analysis of all papers published in 1999 in 16 high profile journals, a workplace study of statistics use by recent PhD and Masters graduates and an investigation of statistical procedures used by PhD students. He organised the session on multivariate statistics at ICOTS6 and is Scientific Secretary for ICOTS7. He is a member of the Advisory Committee of the International Statistical Literacy Project.

Lionel Pereira-Mendoza is an Associate Professor in Mathematics Education at the National Institute of Education, Nanyang Technological University, Singapore. He is also currently Associate Dean of Educational Research for the Institute. During his career he had taught in England, Canada and Singapore. He has been involved in IASE and ICOTS activities for many years and was the local organiser of ICOTS in Singapore. His particular research interest is in the area of graphing and how primary children interpret graphical information from their environment and currently has some graduate students working in this area. In addition, he is leading an international team involving the USA, China, South Korea, Thailand and Singapore looking at Mathematics and Science Education in selected areas of these countries.

Maxine Pfannkuch is a senior lecturer in the Mathematics Education Unit in the Departments of Mathematics and Statistics at The University of Auckland, New Zealand. She graduated with an MSc in mathematics, then taught for many years in a secondary school, after which she trained secondary mathematics teachers. Five years ago she completed a doctorate in statistics education. She is presently teaching introductory undergraduate mathematics and statistics courses as well as graduate courses in statistics education and assessment. Her research is focused upon the development of secondary students' statistical thinking and the development of undergraduate students' statistical literacy.

Mokaeane Victor Polaki is a Senior Lecturer in Mathematics Education in the School of Education at the National University of Lesotho (NUL) in Southern Africa, and teaches mathematics and mathematics education courses to prospective teachers in the same university. Additionally, he is the Head of the newly formed School of Education (formerly Faculty of Education) in the same University. He has published a number of research articles on the development of children's probabilistic thinking and is currently working on students' statistical literacy. In addition to having been the National Correspondent for the International Association for Statistical Education (IASE), he has served as the Regional Representative for the Southern African Association For Research in Mathematics, Science and Technology Education (SARMSTE). In Lesotho he is also a member of the National Mathematics Panel, National Curriculum Committee, and is currently helping the Examinations Council in conducting the assessment of children's literacy and numeracy in Lesotho's elementary schools.

Dave Pratt is a Senior Lecturer in the Institute of Education at the University of Warwick, England. He is Director of the Centre for New Technologies Research in Education (CeNTRE). His teaching encompasses teacher training, mostly at secondary level, and research into the relationship between technological tools and mathematical thinking. His doctoral research focused this broad interest onto technology and students' probabilistic thinking, which has been the stimulus for probability as a continuing strand of his overall research. His work on probability has triggered progress on formulating a theoretical framework on the micro-evolution of mathematical knowledge. He is also on the editorial board of the International Journal of Computers for Mathematical Learning.

Jane Watson is a Reader in Mathematics Education at the University of Tasmania, Hobart, Australia, where she oversees the training of pre-service teachers in mathematics at the primary, middle, and secondary school levels, and directs research projects related to the Chance and Data part of the school curriculum. Over the last decade she has played a major role in four nationally funded professional development projects for teachers in mathematics and statistics education, as well as 13 nationally funded research grants ranging from one to three years. Eight of these grants have focused on students' and teachers' understandings of Chance and Data and have included longitudinal understanding, the effect of collaboration on higher order thinking, and the effect of cognitive conflict on understanding. Current projects are studying understanding of variation and a model for an underlying construct of statistical literacy. Over the years she has published over 200 papers and other multi media work, such as a CD ROM, web site on quantitative literacy in the newspaper, and a radio essay for national broadcast. Reports of her research have appeared in major research journals (e.g., Journal for Research in Mathematics Education, Educational Studies in Mathematics, For the Learning of Mathematics, Mathematical Thinking and Learning, Focus on Learning Problems in Mathematics). In Australia, she was the winner of a prestigious Ian Clunies Ross National Science and Technology Medal in 1999 for her work in statistical literacy, using technology to reach teachers and students with the message that statistics is crucial for survival in today's society.

UNA REVISIÓN DE LAS EVIDENCIAS DE FIABILIDAD Y VALIDEZ DE LOS CUESTIONARIOS DE ACTITUDES Y ANSIEDAD HACIA LA ESTADÍSTICA

JOSE CARMONA MÁRQUEZ
 Universidad de Huelva, España
 carmona@uhu.es

RESUMEN

La importancia del dominio afectivo en el ámbito de la educación estadística es en la actualidad ampliamente reconocida. Con objeto de evaluar la influencia de ese componente afectivo en la formación estadística de los estudiantes, se han elaborado en las dos últimas décadas un número importante de instrumentos para la medición de las actitudes y de la ansiedad hacia la estadística. Sin embargo, muchos de estos instrumentos han sido muy poco investigados y, en general, sus propiedades psicométricas no parecen estar bien establecidas. El objetivo de este estudio es realizar una revisión sistemática de los trabajos de investigación que aportan evidencias empíricas acerca de la fiabilidad y validez de dichos instrumentos. Los resultados de esta revisión muestran ciertas carencias que hacen recomendables nuevos estudios acerca de estos instrumentos de medición de las actitudes y la ansiedad hacia la estadística. Finalizamos proponiendo algunas líneas de investigación que creemos pueden tener cierto impacto en la investigación futura sobre el tema.

Palabras clave: Medición de la actitud; Ansiedad; Estadística; Revisión de la literatura

ABSTRACT

Nowadays, the importance of affectivity in statistics education is widely acknowledged. In the past two decades a large number of instruments to measure attitudes and anxiety toward statistics have been developed in order to assess the influence of emotional factor on students' training. However, many of these instruments have not been analysed in depth and their psychometric properties have not properly been assessed. The aim of this paper is revising the research works that have contributed empirical evidence for the reliability and validity of these instruments. Our results suggest the need for more research about instruments to measure attitudes and anxiety about statistics. We conclude by suggesting some priority research areas that might be relevant in future studies on this topic.

Keywords: Attitude Measurement, Anxiety, Statistics, Literature Review

1. INTRODUCCIÓN

Muchos alumnos suelen llegar a las asignaturas de estadística con preconcepciones y actitudes negativas hacia la materia, asociadas habitualmente con niveles altos de ansiedad cuando se enfrentan a las clases, ejercicios o exámenes de estas asignaturas. Por ejemplo, en un estudio sobre la prevalencia de la ansiedad hacia la estadística, Onwuegbuzie (en prensa) estima que aproximadamente un 75% de los alumnos experimentan niveles elevados de ansiedad. Otros autores han señalado que los alumnos suelen percibir estas asignaturas como obstáculos en el camino para la obtención del título (Perney y Ravid, 1990). Un indicador claro de esta situación es el apodo con el

que muchos alumnos en las universidades anglosajonas suelen referirse a la asignatura de estadística: ‘sadística’ -*sadistics*- (Rosenthal, 1992). Dillon (1982) ha llegado incluso a etiquetar como ‘estadísticofobia’ -*statisticophobia*- esos sentimientos de los alumnos.

A pesar de la importancia del dominio afectivo para un aprendizaje adecuado y un uso efectivo de la estadística, existen pocas investigaciones que aborden el estudio de los constructos de las actitudes y la ansiedad hacia la estadística. En este sentido, sigue siendo válida la afirmación de Gal y Ginsburg (1994) cuando señalaban que “el conjunto de investigaciones sobre las actitudes, las creencias y los afectos relacionados directamente con la educación estadística es reducido y problemático” (párrafo 17). Uno de los problemas de este campo de estudios es la falta de consenso acerca de cuáles son los instrumentos de medida adecuados para medir estos constructos. Como en otras áreas de investigación que están en sus etapas iniciales de desarrollo, existe cierta proliferación de instrumentos de medida que suelen ser usados en pocas ocasiones y cuyo fundamento teórico es poco claro (Schuessler, 1993).

El avance de la investigación del dominio afectivo en la educación estadística requiere, sin embargo, de instrumentos que proporcionen datos fiables y válidos acerca de las actitudes y la ansiedad hacia la estadística. En la literatura sobre educación estadística ya existen algunas revisiones de los instrumentos elaborados para medir estos constructos. Así, por ejemplo, Cashin y Elmore (1997) presentan una recopilación de la mayoría de investigaciones realizadas hasta esa fecha, tanto de actitudes como de ansiedad hacia la estadística, con especial énfasis en las evidencias de fiabilidad y validez presentadas en los mismos. El trabajo de Estrada (2001), en cambio, se centra en el análisis de los instrumentos para medir actitudes, más desde un punto de vista teórico que empírico. De la misma forma, Onwuegbuzie y Wilson (2003) presentan una revisión muy interesante del constructo ansiedad hacia la estadística, en la que, sin embargo, no se profundiza en las características psicométricas de los cuestionarios elaborados para medir dicha ansiedad.

El objetivo de este trabajo es poner a disposición de los investigadores una revisión de la información disponible acerca de los diferentes instrumentos propuestos para medir estos constructos. Se pretende actualizar el trabajo de Cashin y Elmore (1997) y complementar los trabajos de Estrada (2001) y Onwuegbuzie y Wilson (2003) con una revisión que incluya artículos de revista, comunicaciones a congresos y tesis doctorales obtenidos a través de búsquedas en las bases de datos PSYCINFO (Base de datos de información psicológica de la American Psychological Association) y ERIC (Educational Resources Information Center). Fruto de esta revisión se encontraron 112 trabajos de investigación en los que se usaba alguna medida de las actitudes o ansiedad hacia la estadística.

Esta revisión pretende ser un intento sistemático de presentar las evidencias de fiabilidad y validez de estos instrumentos que permita obtener conclusiones acerca de los defectos o virtudes de los mismos. Antes de abordar el análisis de los resultados, hemos planteado dos apartados iniciales cuyo objetivo es introducir conceptualmente las actitudes y la ansiedad hacia la estadística y describir las características principales de los instrumentos de medida analizados. A continuación se incluyen los dos apartados en los que se da cuenta de los resultados de la revisión: las evidencias de fiabilidad y de validez de las medidas de actitudes o ansiedad hacia la estadística. Por último se extraen algunas conclusiones que creemos representan un punto de partida conveniente para las próximas investigaciones en este campo.

2. LAS ACTITUDES Y LA ANSIEDAD HACIA LA ESTADÍSTICA

Tal como señalan la mayoría de los teóricos de la psicometría, la medición de un constructo requiere de una delimitación conceptual precisa del mismo. Para que se puedan elaborar adecuadamente los ítems de un instrumento de medida, es necesario especificar cuáles son las características comunes y diferenciales de esos ítems. Un paso previo a la elaboración de instrumentos de medida debe ser, por tanto, definir el constructo y especificar las dimensiones o facetas que lo conforman. Aunque el objetivo de este trabajo no es elaborar un instrumento de medida, sí lo es evaluar la calidad de instrumentos elaborados. Por ello, en lo que sigue describiremos brevemente qué vamos a entender por actitudes y ansiedad hacia la estadística.

Según Eagly y Chaiken (1998) una actitud es una “tendencia psicológica que es expresada a través de la evaluación de una entidad particular favorable o desfavorablemente en cierto grado”

(p.269). En el caso de las actitudes hacia la estadística esa tendencia se forma a lo largo del tiempo como consecuencia de las emociones y sentimientos experimentados en el contexto del aprendizaje de las matemáticas y la estadística (Gal, Ginsburg y Schau, 1997).

Esta concepción general de las actitudes hacia la estadística es compartida por la mayoría de autores, aunque no existe acuerdo acerca de cuáles son las dimensiones que estructuran este dominio psicológico. Las dos visiones más aceptadas son las propuestas por Wise (1985) y por Schau, Stevens, Dauphinee y del Vecchio (1995). Para Wise (1985) existen dos dimensiones relevantes en el dominio de las actitudes de los alumnos hacia la estadística: las actitudes hacia la asignatura de estadística en que están matriculados y las actitudes hacia el uso de la estadística en su campo de estudios. Schau et al. (1995), por otra parte, distinguen cuatro dimensiones en las actitudes hacia la estadística: (a) Afectos -sentimientos positivos o negativos en relación a la estadística-, (b) Competencia cognitiva -actitudes acerca del conocimiento y las habilidades intelectuales aplicadas a la estadística-, (c) Valor -actitudes acerca de la utilidad, relevancia y valía de la estadística en la vida personal y profesional-, y (d) Dificultad -actitudes acerca de la dificultad de la estadística como materia-. Como veremos más adelante, estas propuestas han servido como base para la elaboración de dos de los cuestionarios de actitudes hacia la estadística más usados en la actualidad.

La ansiedad hacia la estadística puede ser definida como una fobia específica, es decir, como un miedo desmesurado hacia un objeto, en este caso la estadística. Tal como señala Zeidner (1991) la ansiedad hacia la estadística puede manifestarse a través de la aparición de una preocupación excesiva, pensamientos perturbadores, tensión y cierta excitación fisiológica, en determinadas situaciones académicas. Este carácter situacional es una de las características definitorias de la ansiedad y representa un elemento diferenciador con respecto a las actitudes hacia la estadística. Además, constituye el fundamento de la dimensionalidad de este constructo, de manera que es habitual distinguir entre ansiedad hacia el examen, ansiedad en clase o ansiedad al resolver problemas, entre otros tipos de ansiedad hacia la estadística que pueden surgir en situaciones académicas específicas.

3. INSTRUMENTOS DE MEDIDA DE LAS ACTITUDES Y LA ANSIEDAD HACIA LA ESTADÍSTICA

El análisis de la literatura existente arroja un resultado de 17 instrumentos de medida de las actitudes y la ansiedad hacia la estadística. Exceptuando dos instrumentos elaborados a partir de escalas bipolares, a la manera del diferencial semántico de Osgood (Birenbaum y Eylath, 1994; Green, 1993), todos los instrumentos revisados son escalas tipo Likert. En lo que sigue vamos a describir brevemente estos cuestionarios, poniendo un mayor énfasis en aquellos que han sido usados más frecuentemente.

3.1 CUESTIONARIOS DE ACTITUDES HACIA LA ESTADÍSTICA

Los estudios acerca de las actitudes hacia la estadística se remontan al menos a mediados de los años 50 con un trabajo de Bending y Hughes (1954). Esta temática, sin embargo, recibió muy poca atención hasta los años 80 cuando varios autores diseñaron algunos de los instrumentos específicos para la medida de estas actitudes (Roberts y Bilderback, 1980; Wise, 1985). Los cuestionarios más usuales para medir las actitudes hacia la estadística son el SAS -*Statistics Attitude Survey*- de Roberts y Bilderback (1980), el ATS -*Attitudes Toward Statistics scale*- de Wise (1985) y el SATS -*Survey of Attitudes Toward Statistics*- de Schau et al. (1995).

Statistics Attitude Survey (SAS)

Roberts y Bilderback (1980) pretendían elaborar una escala afectiva que por estar “formulada en lenguaje estadístico” (p. 236) fuera más relevante que otros instrumentos de medida de tipo afectivo –escalas de actitudes hacia las matemáticas- para la predicción de las notas en las asignaturas de estadística. A partir de un conjunto inicial de 50 ítems, desarrollados tomando como referencia el dominio de contenido presentado por Dutton (1951) para su escala de actitudes hacia las matemáticas,

seleccionaron los 34 ítems que componen el SAS, de manera que tuvieran una correlación alta con el total de la escala. Cada uno de estos ítems tiene cinco opciones de respuesta, desde ‘Totalmente de acuerdo’ hasta ‘Totalmente en desacuerdo’.

A pesar de que el SAS “fue construido teniendo en cuenta varios conceptos tales como la competencia percibida en estadística y la utilidad de la estadística, no se supuso que era factorialmente complejo” (Roberts y Reese, 1987, p. 759). En general parece deducirse de sus escritos que Roberts considera que este cuestionario como esencialmente unidimensional, aunque en ocasiones use argumentos no demasiado convincentes. Así, por ejemplo, en Roberts y Bilderback (1980) se señala que las altas estimaciones de la consistencia interna indican que “los ítems miden el mismo constructo” (p. 237), a pesar de las advertencias que se pueden encontrar en la literatura psicométrica acerca de que las estimaciones de consistencia interna no son buenos índices de la dimensionalidad de una prueba (Crocker y Algina, 1986).

Según Wise (1985), la inspección de los ítems del SAS revela que al menos un tercio de ellos abordan cuestiones como la capacidad de los estudiantes en la solución de problemas de estadística o la comprensión de conceptos estadísticos. Para Wise estos ítems son inapropiados por dos razones: en primer lugar parecen estar midiendo más el logro que las actitudes de los estudiantes y, en segundo lugar, son inapropiados para alumnos que acaban de comenzar la asignatura de estadística.

Attitudes Toward Statistics scale (ATS)

Los ítems del ATS difieren de los del SAS en tres aspectos (Wise, 1985): (a) se elaboraron para que sean por naturaleza actitudinales, (b) pretenden ser contestables por los estudiantes tanto en el primero como en el último día de clase, y (c) se construyeron para medir dos dominios separados, a saber, las actitudes de los estudiantes hacia la asignatura en la que están matriculados (ATS-Asignatura) y sus actitudes hacia el uso de la estadística en su campo de estudios (ATS-Campo). A partir de un conjunto inicial de 40 ítems con cinco opciones de respuesta cada uno, desde ‘Totalmente de acuerdo’ hasta ‘Totalmente en desacuerdo’, se seleccionaron los 29 ítems que componen el cuestionario definitivo, a través de análisis realizados por jueces expertos acerca la adecuación del contenido y de su aplicabilidad el primer día de clase, y de análisis empíricos teniendo en cuenta la correlación corregida ítem-total.

Survey of Attitudes Toward Statistics (SATS)

Las características que Schau et al. (1995) sugieren para un uso óptimo de un cuestionario de actitudes hacia la estadística, tanto en situaciones de investigación como de instrucción son las siguientes:

- (a) que cubran las dimensiones más importantes de las actitudes hacia la estadística; (b) que sean aplicables en la mayoría de los departamentos que ofrecen cursos de introducción a la estadística y sirvan como medidas relevantes a lo largo del curso con sólo cambios menores en el tiempo verbal; y (c) sean cortas, de modo que su administración implique un mínimo tiempo de clase e incluyan ítems que midan tanto actitudes positivas como negativas (p. 869).

Estos autores señalan, además, que tanto el desarrollo de la encuesta como la validación de contenido deben tener en cuenta a los estudiantes, puesto que son ellos los que rellenan el cuestionario. Por último, señalan que la estructura del instrumento resultante debe ser evaluada a través de técnicas de análisis confirmatorio, como el análisis factorial confirmatorio. Según Schau et al. (1995), ninguno de los instrumentos existentes hasta ese momento había sido elaborado teniendo en cuenta todas estas características. Para subsanar esta situación estos autores elaboraron el SATS atendiendo a todos esos factores.

A partir de una técnica de grupo nominal con alumnos y profesores de asignaturas de estadística, generaron inicialmente 80 ítems cuya estructura, alcanzada por consenso entre los participantes en la técnica grupal se compone de las cuatro dimensiones ya señaladas en el apartado anterior: Afecto, Competencia Cognitiva, Valor y Dificultad. De la muestra inicial de ítems, se seleccionan 28 ítems basándose en varios análisis: grado de acuerdo de los expertos acerca de la dimensión de la cual es indicador cada uno de los ítems; análisis empíricos clásicos basados en la correlación corregida ítem-total de la dimensión, correlación múltiple al cuadrado; valor del coeficiente alfa de Cronbach cuando

el ítem es eliminado de la escala; y análisis empírico del ajuste de los ítems a la estructura propuesta con técnicas de análisis factorial confirmatorio.

Otros cuestionarios de actitudes hacia la estadística

Otros cuestionarios elaborados para medir actitudes hacia la estadística y que han sido usados en un menor número de trabajos son los siguientes:

- a) *Students' Attitudes Toward Statistics* de Sutarso (1992), cuestionario compuesto de 24 ítems, que no parece diferir mucho de otros cuestionarios más usuales como el SAS y el ATS (Gal y Ginsburg (1994);
- b) *Statistics Attitude Scale* de McCall, Belli y Madjidi (1991) cuestionario compuesto por 20 ítems, que ha sido posteriormente usado en Glencross y Cherian (1992) y Cherian y Glencross (1997) con poblaciones sudafricanas;
- c) *Attitude Toward Statistics* de Miller, Behrens, Green y Newman (1993), instrumento compuesto de 25 ítems que pretenden medir los constructos valor de la estadística, orientación a objetivos y habilidad percibida en relación a la estadística;
- d) *Quantitative Attitudes Questionnaire* de Chang (1996), cuestionario diseñado para medir actitudes hacia la metodología cuantitativa, incluyendo la estadística pero no limitándose a ella. En su versión corta está compuesto por 20 ítems elaborados para medir cuatro dimensiones: utilidad de la metodología cuantitativa, valor de esa metodología para la investigación en ciencias sociales, eficacia o percepción de las propias habilidades en relación con lo cuantitativo, y conocimiento acerca de estos temas;
- e) *Escala de actitudes em relação à Estatística* de Cazorla, Silva, Vendramini y Brito (1999), escala unidimensional compuesta por 20 ítems y adaptada de la escala de actitudes hacia las matemáticas de Aiken y Dreger (1961).

De todos los instrumentos analizados, los únicos que, según nuestra revisión, han sido traducidos y usados en español son el SAS de Roberts y Bilderback, usado por Auzmendi (1991), Cuesta, Rifá y Herrero (2001) y Mejía (1995), el ATS de Wise, usado por Gil (1999) y el SATS de Schau et al., usado por Carmona y Moreno (1999), Estrada (2002) y Huedo, López, Martínez y Nortes (2003). Por otra parte, existen dos instrumentos elaborados directamente en español: la *EAE -Escala de Actitudes hacia la Estadística-* de Auzmendi (1991) y otro cuestionario desarrollado más recientemente con el mismo nombre por Velandrino y Parodi (1999).

La *EAE* de Auzmendi (1991) es una escala tipo Likert compuesta por 25 ítems, que ha sido elaborada para medir indistintamente actitudes hacia la estadística y hacia las matemáticas, con pequeñas modificaciones de cada ítem. Para la elaboración del cuestionario se partió de la selección de las cinco dimensiones más comunes en las escalas de actitudes hacia las matemáticas. Se seleccionaron a continuación ocho ítems por cada una de estas dimensiones y se probaron los 40 ítems resultantes en una muestra de alumnos universitarios. Se eligieron los 25 ítems definitivos, cinco por cada factor, usando como criterio de selección de ítems la saturación de éstos en el factor correspondiente en un análisis factorial. Según Auzmendi (1992) las dimensiones o factores de los que consta la *EAE* son: (a) Utilidad subjetiva que tiene para el estudiante el conocimiento de estadística; (b) Ansiedad o temor que se manifiesta ante la materia; (c) Confianza o seguridad que se tiene al enfrentarse a la estadística; (d) Agrado o disfrute que provoca el trabajo estadístico; y (e) Motivación que siente el estudiante hacia el estudio y uso de la estadística.

La *Escala de Actitudes hacia la Estadística* de Velandrino y Parodi (1999) es una escala tipo Likert desarrollada a partir de un conjunto inicial de 60 ítems. A partir de una evaluación de la adecuación del contenido por parte de jueces expertos -profesores de metodología- y de análisis clásicos de ítems se seleccionaron los ítems que componen el cuestionario definitivo. Según los autores, los resultados de un análisis factorial de la matriz de correlaciones entre ítems sirvió de apoyo a la estructura de tres dimensiones propuesta: *utilidad* percibida, *conceptualización* general de la estadística, y *disposición* y capacidad para la *formación* y preparación en estadística.

3.2 CUESTIONARIOS DE ANSIEDAD HACIA LA ESTADÍSTICA

Además de estos instrumentos para la medida de las actitudes, en la literatura se han planteado diversos instrumentos para medir la ansiedad hacia la estadística. Debemos señalar, sin embargo, que en muchos de los trabajos analizados se plantean indistintamente ambos tipos de instrumentos. A nuestro entender existen razones, tanto prácticas como teóricas, para este tratamiento indiferenciado. A un nivel práctico, la mayoría de los cuestionarios propuestos no difieren en gran medida, estando compuestos de ítems muy parecidos. Teóricamente, las actitudes suelen ser entendidas como reacciones evaluativas ante un objeto cuyos antecedentes o manifestaciones pueden ser emociones, cogniciones o conductas manifiestas. La ansiedad hacia un objeto es una reacción fundamentalmente emocional de miedo o aprensión hacia ese objeto. En este sentido, la ansiedad hacia la estadística puede ser entendida como la manifestación de una actitud negativa hacia la estadística o como un antecedente de dicha actitud. Los indicadores de ambos constructos, por tanto, pueden ser muy similares.

Una característica que introduce cierta diferenciación entre los instrumentos de actitudes y los instrumentos de ansiedad hacia la estadística es el hecho de que la ansiedad suele ser concebida como una reacción específica a la situación (Onwuegbuzie, 2000a; Zeidner, 1991), con lo que en los instrumentos para medir ansiedad suelen estar más centrados que los instrumentos de actitudes en situaciones concretas, como el examen y las clases de estadística. Los cuestionarios de ansiedad hacia la estadística más usuales son el STARS -*Statistical Anxiety Rating Scale*- de Cruise, Cash y Bolton (1985) y las diversas adaptaciones del MARS -*Mathematics Anxiety Rating Scale*- de Suinn, Edie, Nicoletti y Spinelli (1972).

Statistical Anxiety Rating Scale (STARS)

El STARS es un cuestionario compuesto de 51 ítems que evalúan la ansiedad que surge en los estudiantes en distintas situaciones académicas relacionadas con la estadística. Cada uno de los ítems consta de cinco opciones de respuesta, desde 1 ‘poca ansiedad’ hasta 5 ‘muchísima ansiedad’ provocada por la situación descrita. Cruise et al. (1985) identifican seis factores, usando un análisis factorial: (a) Valor, o percepción de los estudiantes acerca de la relevancia de la estadística; (b) Ansiedad surgida en situaciones en las que hay que interpretar datos estadísticos; (c) Ansiedad en clase; (d) Autoconcepto de cálculo, que surge cuando se tiene que resolver un problema estadístico que implica cálculos matemáticos; (e) Miedo a pedir ayuda a profesores o compañeros en diversas situaciones; y (f) Miedo a los profesores de estadística, el cual está relacionado con la percepción que se tiene del profesor.

Adaptaciones del Mathematics Anxiety Rating Scale (MARS)

En los trabajos en los que se estudia la ansiedad hacia la estadística también se han usado con bastante frecuencia algunas de las adaptaciones del MARS de Suinn et al. (1972), como el *Statistics Anxiety Inventory* de Zeidner (1991), la *Statistics Anxiety Scale* propuesta por Bessant (1997), o la versión reducida del MARS diseñada por Plake y Parker (1982) para medir ansiedad hacia las matemáticas y la estadística. Estas distintas versiones varían en función del número de ítems: así, mientras la propuesta por Plake y Parker tiene 24 ítems, el cuestionario de Zeidner tiene 40 y el de Bessant 42 ítems. En todos los casos, los ítems de los cuestionarios contienen pequeñas descripciones de situaciones, fundamentalmente académicas, relacionadas con la estadística o las matemáticas ante las cuales los sujetos deben responder en una escala de cinco puntos desde 1 ‘No me provoca ansiedad’ hasta 5 ‘Me provoca mucha ansiedad’. Los ítems de estos cuestionarios suelen ser idénticos a los del cuestionario original, con la modificación de incluir la palabra ‘estadística’ en lugar o junto a la palabra ‘matemáticas’. Todos estos cuestionarios incluyen dos dimensiones o factores denominados respectivamente ‘Ansiedad ante el aprendizaje o contenido de la estadística’ y ‘Ansiedad ante los exámenes de estadística’.

Otros cuestionarios de ansiedad hacia la estadística

Un cuestionario de ansiedad hacia la estadística con un uso menor que los anteriores es el *Statistics Anxiety Scale* de Pretorius y Norman (1992), elaborado a partir de la *Mathematics Anxiety Scale* de Betz (1978), y compuesto por 10 ítems idénticos a los de Betz, salvo en que se han cambiado las palabras referidas a matemáticas por palabras referidas a estadística.

Además en algunos de los trabajos sobre ansiedad hacia la estadística se han usado cuestionarios específicamente concebidos para medir la ansiedad hacia los exámenes como el TAI *-Test Anxiety Inventory-* de Spielberger (1980), usado por Bandalos, Yates y Thorndike-Christ (1995), Benson (1989), Musch y Bröder (1999) y Finney y Schraw (2003), y el STAS *-Statistics Test Anxiety Scale-* de Fitzgerald, Jurs y Hudson (1996) usado sólo por sus autores. El TAI es una escala tipo Likert compuesta por 20 ítems desarrollados para medir ansiedad hacia cualquier tipo de examen y que en las aplicaciones analizadas se ha usado sin modificar, con la única instrucción de que los alumnos respondan pensando en el examen de estadística. Según sus autores, el TAI mide dos dimensiones: a) Preocupación o componente cognitivo de la ansiedad hacia los exámenes que refleja las preocupaciones con la ejecución en el examen, y b) Emotividad o medida de las reacciones afectivas y fisiológicas como el nerviosismo o la tensión ante las evaluaciones. El STAS es una versión modificada del *Test Anxiety Scale* de Sarason (1972) en la que se ha reemplazado la palabra ‘examen’ por la palabra ‘estadística’; consta de 34 ítems con formato de respuesta verdadero-falso en los que se pide a los sujetos que indiquen su reacción ante diversas cuestiones relacionadas con los exámenes.

4. FIABILIDAD DE LOS INSTRUMENTOS

A continuación vamos a revisar los resultados de estos estudios desde la perspectiva de sus aportaciones al análisis de la calidad psicométrica de los cuestionarios de actitudes y ansiedad hacia la estadística. Los datos que se comentan corresponden fundamentalmente a los cinco cuestionarios más usuales -SAS, ATS, SATS, R-MARS y STARS- y a la EAE de Auzmendi, debido a que es el cuestionario en español más usado.

En casi todos los estudios revisados se informa acerca de la consistencia interna de las escalas a través del coeficiente Alfa de Cronbach. En general, estos indicadores de la consistencia interna suelen ser altos, con valores en la mayoría de los casos superiores a 0.80 -mínimo aceptable recomendado por algunos autores (Henson, 2001)- y en muchos casos superiores incluso a 0.90. Los valores mínimos y máximos del coeficiente alfa para las escalas y subescalas revisadas pueden consultarse en la Tabla 1.

A diferencia de las medidas de consistencia interna, en muy pocas ocasiones se informa acerca de la estabilidad de las puntuaciones, cuya estimación requiere administrar el instrumento en dos ocasiones, obteniéndose un coeficiente test-retest o coeficiente de correlación de Pearson entre las puntuaciones obtenidas en ambas ocasiones. Según nuestro conocimiento, sólo los trabajos de Cruise et al. (1985) y de Wise (1985) aportan medidas de la estabilidad de las puntuaciones que pueden ser interpretadas como estimaciones de la fiabilidad. Wise (1985) obtiene un coeficiente test-retest de 0.82 para la subescala Campo y de .91 para la subescala Asignatura de su cuestionario, con un tiempo entre aplicaciones de la prueba de dos semanas. Cruise et al. (1985), en cambio, dejaron transcurrir cinco semanas entre las aplicaciones de la prueba y obtuvieron unos coeficientes test-retest que varían entre 0.67 y 0.83, dependiendo de la subescala. Aunque existen otros trabajos en los que se aporta información acerca de la estabilidad de las puntuaciones (Pretorius y Norman, 1992; Shultz y Koshino, 1998), esta información no puede ser asumida a nuestro entender como un indicador de la fiabilidad debido a que el tiempo transcurrido entre aplicaciones es demasiado grande, aproximadamente tres meses en ambos casos, y a que en ese tiempo los alumnos han estado cursando una asignatura de estadística, lo cual puede haber modificado de forma natural sus actitudes.

Tabla 1. Valores mínimos y máximos del coeficiente alfa de Cronbach obtenidos en las escalas SAS, ATS, STARS, RMARS, SATS o EAE por distintos autores.

Escala	Componente	Alfa Mínimo	Alfa Máximo
SAS		0.89 (Mejía, 1995)	0.95 (Roberts y Bilderback, 1980)
ATS	Total	0.89 (Gil, 1999)	0.91 (Roberts y Reese, 1987)
	Campo	0.83 (Waters et al., 1988a)	0.94 (Perney y Ravid, 1990)
	Asignatura	0.77 (Rhoads y Hubele, 2000)	0.93 (Perney y Ravid, 1990)
STARS	Valor	0.78 (Onwuegbuzie y Daley, 1999)	0.96 (Onwuegbuzie, 2000b)
	Interpretación	0.78 (Onwuegbuzie y Daley, 1999)	0.91 (Baloglu, 2003)
	Examen	0.68 (Cruise et al., 1985)	0.91 (Baloglu, 2002)
	Cálculo	0.77 (Onwuegbuzie y Daley, 1999)	0.93 (Onwuegbuzie, 1993)
	Ayuda	0.62 (Baloglu, 2002)	0.89 (Cruise et al., 1985)
	Profesores	0.64 (Baloglu, 2003)	0.88 (Elmore et al., 1993)
Versiones de MARS	Total	0.86 (Sigurdsson, 1991)	0.98 (Plake y Parker, 1982)
	Contenido	0.94 (Zeidner, 1991)	0.94 (Zeidner, 1991)
	Examen	0.92 (Zeidner, 1991)	0.92 (Zeidner, 1991)
EAE	Total	0.89 (Sánchez-López, 1996)	0.90 (Quilter y Chester, 2001)
	Utilidad	0.64 (Auzmendi, 1991)	0.80 (Auzmendi, 1991)
	Ansiedad	0.81 (Auzmendi, 1991)	0.84 (Auzmendi, 1991)
	Seguridad	0.74 (Auzmendi, 1991)	0.84 (Auzmendi, 1991)
	Agrado	0.79 (Auzmendi, 1991)	0.83 (Auzmendi, 1991)
	Motivación	0.61 (Auzmendi, 1991)	0.71 (Auzmendi, 1991)
SATS	Afecto	0.81 (Schau et al., 1995)	0.89 (Finney y Schraw, 2003)
	Competencia	0.77 (Schau et al., 1995)	0.90 (Finney y Schraw, 2003)
	Valor	0.80 (Schau et al., 1995)	0.91 (Finney y Schraw, 2003)
	Dificultad	0.64 (Schau et al., 1995)	0.86 (Finney y Schraw, 2003)

5. VALIDEZ DE LOS INSTRUMENTOS

La validación es un proceso por el cual se aportan evidencias que apoyen la interpretación propuesta de los datos recogidos mediante la prueba, en nuestro caso del cuestionario de actitudes o ansiedad hacia la estadística. Según nuestro análisis, las evidencias aportadas en las investigaciones no están basadas en los procesos de respuesta o en las consecuencias de la evaluación, a pesar de la recomendación en este sentido de los *Standards* (AERA/APA/NCME, 1999). Una posibilidad para obtener información acerca de los procesos de respuesta es la propuesta por Gal et al. (1997), quienes consideran conveniente complementar las preguntas habituales -formato de respuesta tipo Likert- con preguntas abiertas de cuyas respuestas, para inferir las motivaciones y causas de las actitudes de los alumnos. Sin embargo, esta recomendación no ha sido seguida en la mayoría de los trabajos analizados, con la excepción de algunos trabajos recientes (Suanpang, Petocz y Kalceff, 2003; Watson, Kromrey, Ferron, Lang y Hogarty, 2003).

A continuación se presentan los resultados de las investigaciones revisadas organizados en función del tipo de evidencia que aportan en el proceso de validación de los distintos cuestionarios, que puede estar basada en (a) el contenido del cuestionario, (b) su estructura interna, o (c) su relación con otras variables.

5.1 VALIDEZ DE CONTENIDO

El análisis de aquellas investigaciones en las que se usa alguno de los seis cuestionarios revisados en mayor profundidad en este trabajo, mostró que sólo para tres de ellos se aportan evidencias de validez relativas al contenido. En estos casos, esas evidencias se obtienen en la fase de elaboración del instrumento y se presentan en los artículos originales, que describen esa fase de desarrollo de los cuestionarios. Los cuestionarios para los que se aporta este tipo de evidencias son el STARS de Cruise et al., ATS de Wise y el SATS de Schau et al.

Cruise et al. (1985) realizaron una evaluación de la adecuación del contenido de los ítems a la estructura de seis factores propuesta para el STARS presentando a cinco profesores de estadística y cinco estudiantes de doctorado descripciones de las seis subescalas e ítems potenciales para cada una de ellas. La tarea de los evaluadores era determinar a cuál de las subescalas pertenecía cada uno de los ítems. Los resultados obtenidos por estos autores mostraron unos índices de acuerdo que iban de 0.60 a 1, con una media de 0.91, lo cual fue interpretado como una evidencia de la validez aparente del contenido de los ítems.

Wise (1985) describe el proceso de evaluación del contenido de los ítems inicialmente elaborados de la siguiente manera: “*La evaluación de los 40 ítems en términos de su validez de contenido por dos profesores de introducción a la estadística en educación dio lugar a la eliminación de cinco ítems*” (p. 403). Puesto que Wise no es más explícito no podemos estar seguros acerca de qué tipo de evaluación llevan a cabo esos expertos en el contenido. Tras la eliminación de otros cinco ítems debido a sus bajos índices de discriminación, dos profesores de estadística distintos a los anteriores realizan una evaluación adicional -en este caso explícita- de la adecuación de los 30 ítems restantes para ser respondidos el primer día de clase, y cuyo resultado es la eliminación de otro ítem.

La evaluación del contenido de los ítems del SATS difiere de la realizada por Wise (1985) en dos aspectos. Por una parte, los jueces evaluadores no son sólo profesores: el grupo de jueces constaba de dos profesores de introducción a la estadística, dos alumnos graduados y dos alumnos no graduados matriculados en cursos de iniciación a la estadística. Y por otra, en Schau et al. (1995), tanto la estructura conceptual, como la relevancia de cada ítem fue determinada inductivamente a partir del consenso entre los jueces. En primer lugar, los jueces clasificaron un conjunto inicial de frases y palabras referidas a las actitudes hacia la estadística en categorías que etiquetaron y definieron con posterioridad. En esta fase “el grupo alcanzó una estructura consensuada de los ítems consistente en cuatro dimensiones” (Schau et al., 1995, p. 869), etiquetadas como Afecto, Competencia Cognitiva, Valor y Dificultad. Posteriormente esas palabras y frases fueron reescritas en 80 ítems potenciales que de nuevo fueron clasificados por cada uno de los jueces en una de las dimensiones propuestas. “Los 60 ítems (15 por categoría) que dieron lugar al mayor consenso (habitualmente 80% de acuerdo o mayor) fueron reescritos para ser usados de forma adecuada al inicio de un curso de introducción a la estadística” (Schau et al., 1995, p. 870).

Del resto de los cuestionarios descritos, sólo el *Quantitative Attitudes Questionnaire* de Chang (1996) aporta un análisis de la validez de contenido. En este caso, al igual que en el SATS, las dimensiones que conforman el constructo se derivaron a partir de las opiniones de un grupo de jueces, aunque Chang usó sólo profesores de metodología. El conjunto inicial de 65 ítems fueron revisados por esos mismos jueces para determinar su validez de contenido, reteniéndose aquellos 45 ítems que suscitaron un mayor consenso.

5.2 EVIDENCIAS BASADAS EN LA ESTRUCTURA INTERNA

El análisis de las evidencias de validez basadas en la estructura interna de una prueba consiste en el estudio del ajuste entre la estructura conceptual del constructo y las relaciones empíricas entre ítems o partes de la prueba. Aunque los análisis de datos implicados debieran depender fundamentalmente de la concepción del constructo, en las investigaciones revisadas el estudio de la estructura interna de los cuestionarios se ha realizado en todos los casos por medio de análisis factoriales de las respuestas a los ítems de esos cuestionarios, tanto exploratorios como confirmatorios.

El SAS de Roberts y Bilderback (1980) es uno de los pocos cuestionarios de actitudes o ansiedad hacia la estadística que se postulan unidimensionales, junto al *Statistics Attitude Scale* (McCall et al., 1991; Glencross y Cherian, 1992), el *Statistics Anxiety Scale* (Pretorius y Norman, 1992), y la *Escala de atitudes em relação à Estatística* (Cazorla et al., 1999). Sin embargo, cuando se ha puesto a prueba esa supuesta unidimensionalidad del SAS a través de análisis factoriales de las respuestas a sus ítems (Cuesta et al., 2001; Waters, Martelli, Zakrajsek y Popovich, 1988b) el resultado muestra como solución más satisfactoria una estructura bifactorial.

La estructura bidimensional propuesta por Wise (1985) para su ATS, sin embargo, ha sido probada en varias ocasiones usando análisis factoriales exploratorios y confirmatorios. En el trabajo original de Wise, por ejemplo, se usó un análisis factorial exploratorio con rotación varimax. La solución bifactorial obtenida dio cuenta del 49% de la varianza y los factores se interpretaron, como ya se indicó anteriormente, como actitudes hacia la asignatura y actitudes hacia el campo de la estadística. Esa estructura fue obtenida de nuevo en análisis factoriales exploratorios del ATS en los trabajos de Waters et al. (1988b) y en Woehlke (1991). En el estudio realizado por Schau, Dauphinee y Del Vecchio (1993) se aplicó un análisis factorial confirmatorio en el que se obtuvo que todos los ítems saturaban de forma significativa en su correspondiente factor, aunque el ajuste global del modelo no fue satisfactorio. Gil (1999), por otra parte, encontró resultados que no apoyaban la estructura bifactorial al realizar un análisis factorial exploratorio con rotación varimax de una versión en español del ATS, puesto que obtuvo una solución de cinco factores que explicaba aproximadamente el 51% de la varianza.

Cruise et al. (1985) realizaron un análisis factorial exploratorio con rotación varimax a partir de las respuestas de 1150 estudiantes al STARS, a partir del cual se obtuvieron precisamente los seis factores señalados en el anterior apartado de este trabajo. En un estudio más reciente, sin embargo, Baloglu (2002), usando técnicas de análisis factorial confirmatorio, concluye que el modelo de seis factores no muestra un ajuste adecuado a los datos.

Con respecto a las versiones del MARS de Suinn et al. (1972), los análisis factoriales realizados tanto en Plake y Parker (1982) como en Zeidner (1991) muestran que una estructura bifactorial da cuenta adecuadamente de las respuestas a los ítems de ambas versiones del cuestionario: las soluciones obtenidas en estos AF exploratorios dan cuenta respectivamente del 60% y del 45% de la varianza total.

La dimensionalidad del SATS de Schau et al. (1995), a diferencia del resto de cuestionarios, ha sido estudiada en distintas ocasiones a través de análisis factoriales confirmatorios. En Schau et al. (1993) se intentó confirmar la estructura de cuatro factores propuesta por los autores usando un AF confirmatorio al nivel de los ítems. Los resultados no fueron satisfactorios, debido a un pobre ajuste del modelo. Sin embargo, tanto en Schau et al. (1993) como en Schau et al. (1995) se realizó un AF confirmatorio de las puntuaciones en subgrupos de ítems *-parcels-*, formados a partir de la suma de las puntuaciones a varios ítems del mismo factor, obteniéndose en ambos casos un ajuste adecuado del modelo de cuatro factores. En Dauphinee, Schau y Stevens (1997) se encontró además que esa estructura factorial es invariante para alumnos y alumnas.

La Escala de Actitudes hacia la Estadística de Auzmendi ha sido sometida en dos ocasiones a un AF de componentes principales con rotación varimax. En la tesis doctoral de Auzmendi (1991), el resultado de dicho análisis muestra una estructura de cinco factores que da cuenta del 60.7% de la varianza total. En Sánchez-López (1996), en cambio, se obtuvo una estructura factorial distinta, constituida por cuatro factores que daban cuenta del 53.5% de la varianza.

En algunos de los estudios en los que se usan los otros cuestionarios revisados también se realizan análisis factoriales. Resumiendo, podemos señalar que mientras los resultados de los trabajos de Cazorla et al. (1999), Glencross y Cherian (1992), y Pretorius y Norman (1992) parecen apoyar la hipótesis de una estructura unifactorial de los cuestionarios respectivos; Velandrino y Parodi (1999) obtienen una estructura de tres factores para su Escala de Actitudes hacia la Estadística; Chang (1996), usando AF confirmatorio, una estructura de cuatro factores en su *Quantitative Attitudes Questionnaire*; y Miller et al. (1993) obtienen mediante análisis de componentes principales una estructura de cinco factores que muestra un ajuste aceptable a las subescalas teóricamente determinadas de su cuestionario *Attitude Toward Statistics*.

5.3 EVIDENCIAS BASADAS EN LAS RELACIONES CON OTRAS VARIABLES

Las evidencias basadas en la relación entre las puntuaciones en el test y variables externas proporcionan una de las fuentes más habituales de validación de las medidas psicológicas. La determinación de cuáles son esas variables externas depende de la teoría relativa al constructo que se

pretende medir. Aunque en nuestro caso dichas relaciones son, en principio, muy numerosas, hemos clasificado los resultados de las investigaciones revisadas en tres tipos:

1. En primer lugar, puesto que uno de los objetivos básicos de los cuestionarios de actitudes o ansiedad hacia la estadística es la predicción del rendimiento de los alumnos en las asignaturas de estadística, los intentos de validación de estos cuestionarios han incluido habitualmente el estudio de la relación entre las puntuaciones obtenidas en ellos y alguna medida del rendimiento de los alumnos, como son las notas en las asignaturas.
2. También ha sido habitual estudiar la relación que existe entre los diferentes cuestionarios de actitudes hacia la estadística. En principio es esperable que estas medidas estén altamente relacionadas, puesto que son medidas del mismo constructo.
3. Por último, existe un conjunto de investigaciones en las que se estudia la utilidad de algunas variables para predecir las actitudes hacia la estadística. En la literatura revisada hemos encontrado tres grupos de variables usadas de esta manera: características personales de los alumnos, como el género y determinadas características de personalidad; experiencia formativa en matemáticas y estadística principalmente; y pensamientos auto referidos sobre las capacidades personales.

En las investigaciones revisadas se han usado diversos métodos de análisis de datos para el estudio de esas relaciones, que van desde simples correlaciones bivariadas, hasta modelos de ecuaciones estructurales. Un modelo global de esas relaciones incluye variables predictoras de las actitudes hacia la estadística, varias medidas de las actitudes, y el rendimiento en las asignaturas de estadística, tal como se refleja en la Figura 1.

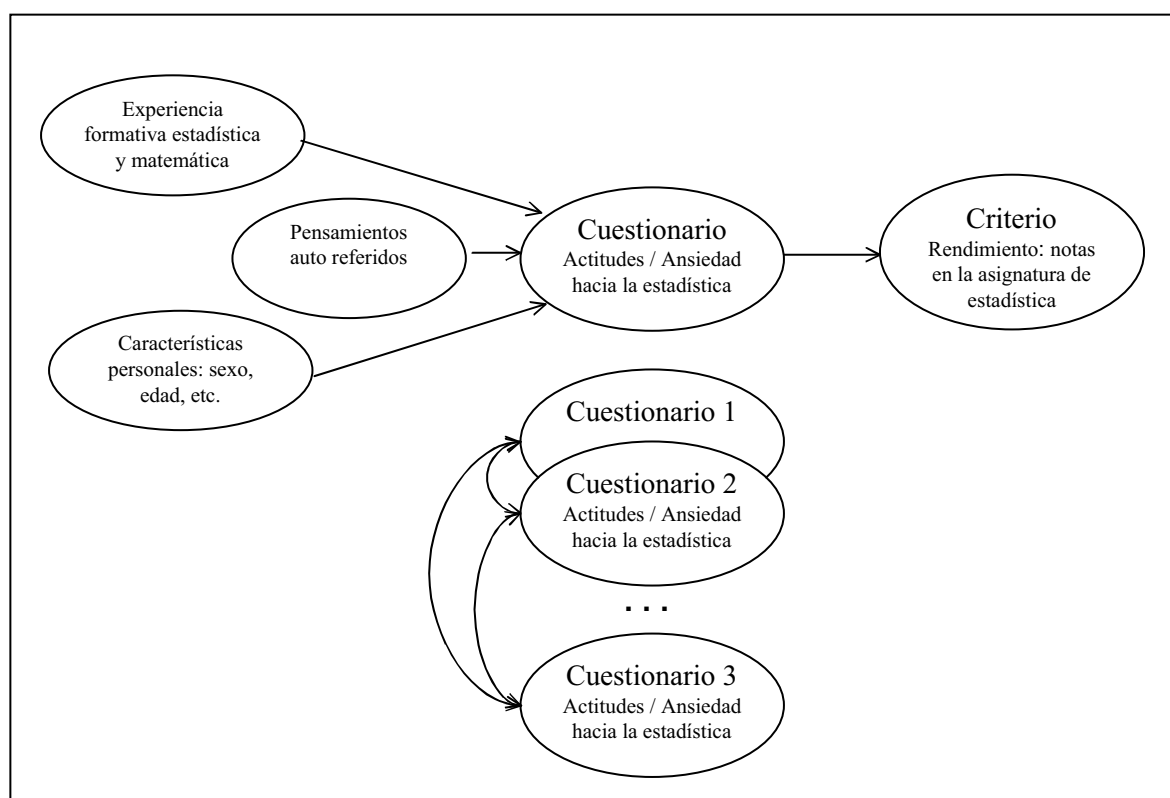


Figura 1. Modelo global de relaciones de las actitudes / ansiedad hacia la estadística con otras variables

Puesto que el objetivo de este apartado es revisar las evidencias que contribuyen a la validación de los cuestionarios de actitudes y ansiedad hacia la estadística, a continuación se presentan esas evidencias agrupadas en las tres categorías señaladas. Si atendemos únicamente al orden explicativo

planteado en el modelo global, esos tres tipos de evidencias pueden etiquetarse como evidencias de validez de pronóstico, concurrente y retrospectiva.

Evidencias de validez de pronóstico

En general, en los cuestionarios analizados se ha mostrado una relación positiva o negativa entre el rendimiento en las asignaturas de estadística y las actitudes o ansiedad hacia la estadística, que puede ser de moderada o de baja intensidad. Esto se refleja, por ejemplo, en el metaanálisis realizado por Fitzgerald (1997) en el que se obtiene una estimación global de esta relación de $r = -0.15$ a partir de 27 estudios, en los que se informa de la relación entre ansiedad hacia la estadística y rendimiento. De la misma forma, en nuestra revisión, encontramos 13 estudios en los que se informa de la relación entre el ATS de Wise y el rendimiento de los alumnos en asignaturas estadísticas, siendo la correlación promedio de 0.21. Resultados similares se obtienen con otros cuestionarios. Así, por ejemplo, Roberts y Reese (1987) encuentran una correlación entre el SAS de Roberts y Bilderback y las notas en una asignatura de estadística de 0.14; Schutz, Drogosz, White y Distefano (1998) informan correlaciones de 0.21 y 0.16, entre las notas y SATS-Afecto y SATS-Valor, respectivamente.

Aunque en la literatura se han propuesto diferentes variables moderadoras de esta relación (véase Fitzgerald, 1997), la variable que parece tener un efecto moderador más patente es el tiempo transcurrido entre la administración del cuestionario y la realización del examen de la asignatura, de manera que cuanto mayor sea este tiempo, menor es la correlación encontrada. Este resultado se obtiene en varios trabajos de investigación, independientemente del cuestionario usado: RMARS (Harvey, Plake y Wise, 1985), SAS (Mejía, 1995; Roberts y Saxe, 1982), ATS (Green, 1994; Shultz y Koshino, 1998; Waters et al., 1988a) o SATS (Finney y Schraw, 2003).

Por otra parte, existen varios trabajos en los que las actitudes o la ansiedad hacia la estadística forman parte de un modelo multivariado para predecir el rendimiento de los alumnos, aunque sus resultados no permiten extraer conclusiones definitivas acerca del papel que juegan las actitudes o la ansiedad. La variedad de los modelos probados hace difícil la comparación entre ellos, aunque una característica común en muchos de estos estudios es que la importancia de la formación matemático-estadística en la predicción del rendimiento parece ser mayor que la ansiedad o actitud de los alumnos (Musch y Bröder, 1999; Schutz et al., 1998; Sorge y Schau, 2002). Además, aunque existen estudios en los que las actitudes o la ansiedad tienen efectos significativos importantes sobre el rendimiento (Fitzgerald et al., 1996; Nasser, 1999; Onwuegbuzie, Slate, Paterson, Watson y Schwartz, 2000), en otros trabajos esos efectos son poco importantes e incluso no significativos (Cashin, 2001; Scott, 2002; Wisenbaker, Nasser y Scott, 1999).

Evidencias de validez concurrente

En varios de los trabajos revisados se ha usado más de uno de los cuestionarios de actitudes o ansiedad hacia la estadística para intentar determinar la convergencia entre medidas. En general, se espera una correlación positiva entre cuestionarios de actitudes o entre cuestionarios de ansiedad y negativa entre cuestionarios de actitud y ansiedad. Las correlaciones encontradas en la mayoría de estos trabajos son bastante elevadas y del signo esperado, indicando una alta convergencia entre las distintas medidas.

Roberts y Reese (1987) fueron los primeros en usar conjuntamente el SAS y el ATS. En una muestra compuesta por 280 alumnos universitarios obtuvieron una correlación entre ambos cuestionarios de 0.88, lo cual interpretaron como una evidencia de que ambos cuestionarios pueden ser entendidos como formas alternativas de la misma prueba. Waters et al. (1988a) por su parte informan de una correlación de 0.73 entre el SAS y la subescala ATS-Asignatura, y de una correlación de 0.83 entre el SAS y la subescala ATS-Campo.

Schau et al. (1995) calcularon las correlaciones entre las diversas escalas del SATS y el ATS usando una amplia muestra de alumnos universitarios, en las que se puede destacar la alta relación entre la subescala ATS-Asignatura y las subescalas SATS-Afecto y SATS-Competencia Cognitiva ($r = 0.79$ y $r = 0.76$, respectivamente) y entre ATS-Campo y SATS-Valor ($r = 0.76$). Cashin y Elmore

(2000), por su parte, han utilizado conjuntamente el SAS, ATS y SATS obteniendo correlaciones significativas entre los tres instrumentos.

En el trabajo de Harvey et al. (1985) se usó un cuestionario de actitudes hacia la estadística (ATS) junto a un cuestionario de ansiedad hacia la estadística (RMARS) en dos muestras distintas. En la primera muestra obtuvieron una correlación entre ambos instrumentos de 0.06. En la segunda muestra, la correlación entre el RMARS aplicado al principio del curso y la subescala ATS-Campo fue de -0.12 y con la subescala ATS-Asignatura fue de -0.75; mientras que esas correlaciones fueron respectivamente de -0.03 y -0.60 cuando la medida del RMARS se obtuvo en un momento posterior.

Recientemente, Watson et al. (2003) aplicaron conjuntamente el SATS y el STARS a una muestra de 200 graduados universitarios matriculados en Facultades de Educación. La correlación entre las puntuaciones totales del SATS y del STARS fue de -0.89. Además se obtuvieron correlaciones entre la puntuación total en el SATS y cada una de las seis subescalas del STARS que iban desde -0.38 a -0.87.

Por último, Auzmendi (1991) para validar su Escala de Actitudes hacia la Estadística aplicó ese cuestionario junto al SAS en una muestra de 101 estudiantes de Psicología y Pedagogía, obteniendo una correlación entre ambos instrumentos de 0.86.

Evidencias de validez retrospectiva

En este apartado vamos a abordar aquellos estudios en que las actitudes hacia la estadística funcionan como variable dependiente. Dependiendo del tipo de variables predictoras hemos clasificado los resultados de estos estudios en aquellos en los que se analiza la capacidad predictiva de determinadas variables personales como el género y características de personalidad, variables relacionadas con la experiencia formativa, y variables que describen los pensamientos auto-referidos relativos a las propias capacidades.

Variables personales

A pesar de que existe cierta tradición de considerar que las mujeres tienen una mayor ansiedad o actitudes más negativas hacia la estadística que los hombres, heredada de la literatura sobre actitudes hacia las matemáticas, en la revisión realizada hemos encontrado resultados contradictorios. Algunos de los trabajos en los que se encuentran diferencias estadísticamente significativas entre hombres y mujeres son: Onwuegbuzie (1995), usando el STARS; Roberts y Saxe (1982) y Cuesta et al. (2001), usando el SAS; y Zeidner (1991) y Sigurdsson (1991), usando versiones del MARS. Mientras Onwuegbuzie (1995), Roberts y Saxe (1982) y Sigurdsson (1991) encuentran que las diferencias son las esperadas, Zeidner (1991) encuentra un patrón más complejo, según el cual las mujeres tenían una mayor ansiedad hacia el examen de estadística que los hombres y éstos tenían niveles ligeramente superiores de ansiedad hacia el contenido de la estadística. En otros trabajos el género de los alumnos no produce diferencias significativas en las puntuaciones de los cuestionarios. Así, por ejemplo, ocurre en algunos trabajos en los que se usa el ATS (Araki y Shultz, 1995; Cashin, 2001; Gil, 1999; Harvey et al., 1985), el RMARS (Harvey et al., 1985), el STARS (Baloglu, 2003; Tomazic y Katz, 1988), o el SATS (Sorge, Schau, Hubele y Kennedy, 2000).

En cuanto a la edad de los estudiantes, de nuevo se han encontrado resultados contradictorios. Con respecto al SAS de Roberts y Bilderback, se han encontrado correlaciones consistentemente bajas con la edad, aunque en algún caso esa correlación tenía un valor positivo de 0.15 (Roberts y Saxe, 1982), mientras en otro su valor era negativo de -0.14 (Roberts y Reese, 1987). Usando un cuestionario de ansiedad como el STARS también se obtienen resultados contrapuestos: mientras Tomazic y Katz (1988), usando un MANOVA con las subescalas del STARS como variables dependientes, no encuentran diferencias significativas entre estas subescalas en función de la edad; Baloglu (2003), usando un MANCOVA para determinar la influencia de la edad y el género sobre las puntuaciones en las subescalas del STARS, y controlando el efecto de la experiencia matemática previa, encuentra que los estudiantes mayores tenían niveles más altos de ansiedad hacia el examen y las clases de estadística que los estudiantes más jóvenes.

Onwuegbuzie y sus colaboradores han realizado en la última década varios estudios en los que analizan la relación entre varias características personales y la ansiedad hacia la estadística, medida a

través del STARS. En un estudio con 90 alumnos de ciencias sociales y conductuales matriculados en asignaturas de carácter metodológico, Onwuegbuzie y Daley (1997) encontraron mayores niveles de ansiedad en aquellos alumnos menos orientados hacia una inteligencia lingüística y lógico-matemática y más orientados hacia una inteligencia espacial e interpersonal. Por otra parte, Onwuegbuzie (1999) encontró que los alumnos de raza negra tienen niveles de ansiedad estadísticamente superiores a los mostrados por alumnos de raza blanca. En un estudio realizado por Onwuegbuzie y Daley (1999) para estudiar la relación entre el perfeccionismo y la ansiedad hacia la estadística, encuentran que aquellos alumnos que mantienen expectativas no realistas acerca de otras personas (perfeccionistas orientados hacia otros) y aquellos que tienen una necesidad excesiva de alcanzar objetivos prescritos por otras personas (perfeccionistas socialmente determinados) suelen tener niveles superiores de ansiedad hacia la estadística. Tenemos que señalar, sin embargo, que este último resultado ha sido puesto en cuestión recientemente por un estudio de Walsh y Ugumba-Agwunobi (2002) en el que se encuentra que la ansiedad hacia la estadística está más asociada al perfeccionismo orientado hacia uno mismo que al perfeccionismo orientado a otros.

Experiencia formativa

En la literatura especializada se ha supuesto que la experiencia formativa con asignaturas de matemáticas y de estadística es el principal determinante de las actitudes de los alumnos. En las investigaciones realizadas al respecto la experiencia previa ha sido operativizada por medio de variables como el número de asignaturas de matemáticas y de estadística cursadas, las notas obtenidas en estas asignaturas, y los conocimientos en matemáticas o estadística medidos a través de algún examen inicial.

Teóricamente es esperable que cuanto mayor sea la experiencia formativa previa en matemáticas o estadística –número de asignaturas- o mayor haya sido el aprovechamiento de la misma –notas o conocimientos-, más positiva sea la actitud y menor la ansiedad de los alumnos. Los resultados de la mayoría de las investigaciones confirman esta hipótesis, aunque los resultados son más claros cuando la relación se establece con las notas o los conocimientos que con el número de asignaturas.

Se encuentran también correlaciones significativas del número de asignaturas de matemáticas con las puntuaciones del SAS (Roberts y Saxe, 1982) y del RMARS (Royse y Rompf, 1992). Perney y Ravid (1990) obtuvieron puntuaciones significativamente más altas en ambas subescalas del ATS en aquellos alumnos que tenían una formación matemática más extensa que en otros con una menor formación. Onwuegbuzie (1998) por otra parte encontró correlaciones significativas entre el número de asignaturas de estadística cursadas con anterioridad y las puntuaciones en ATS-Campo y ATS-Asignatura. En el trabajo de Roberts y Reese (1987), sin embargo, se obtuvieron correlaciones muy bajas y no significativas entre el número de asignaturas de estadística y el número de asignaturas de matemáticas, por una parte, y las puntuaciones en el ATS y en SAS por otra.

En la mayoría de las investigaciones analizadas se han encontrado correlaciones significativas entre las puntuaciones en los cuestionarios de actitudes o ansiedad hacia la estadística y alguna medida de los conocimientos de matemáticas al empezar el curso. Se encuentran esas correlaciones significativas cuando se usan las puntuaciones en el SAS (Roberts y Saxe, 1982), el ATS (Harvey et al., 1985; Kottke, 2000; Woehlke, 1991), el RMARS (Harvey et al., 1985; Plake y Parker, 1982) y el SATS (Schutz et al., 1998). El promedio de estas correlaciones, tomando los valores absolutos de las mismas, es igual a 0.36, bastante superior al promedio de 0.19 obtenido cuando se consideran las correlaciones entre puntuaciones en los cuestionarios y número de asignaturas de matemáticas o estadística. Tomazic y Katz (1988), por otra parte, encontraron diferencias significativas en las puntuaciones de las seis subescalas del STARS entre los cuatro grupos de alumnos formados al dividir la muestra en función de que hubieran sacado buenas notas o no en matemáticas y de que hubieran cursado recientemente o no las asignaturas de matemáticas. Royse y Rompf (1992), por último, obtuvieron menos ansiedad, medida a través del RMARS, en los alumnos de su muestra que no habían suspendido nunca una asignatura de matemáticas que en aquellos que sí habían suspendido.

Cuando se ha estudiado si el conocimiento previo en estadística está relacionado con las actitudes, se han obtenido correlaciones significativas de promedio 0.30 con el SAS (Roberts y Saxe, 1982) y el ATS (Kottke, 2000). En Schutz et al. (1998), sin embargo, se encuentran correlaciones prácticamente

nulas con el SATS-Afecto y el SATS-Valor, lo cual puede ser debido a que se usó un tipo de prueba atípica para medir los conocimientos de estadística consistente en rellenar los vacíos de un árbol conceptual.

En el ámbito español, tanto Gil (1999) usando el ATS como Cuesta et al. (2001) con el SAS han obtenido relación entre las actitudes hacia la estadística y el tipo de bachillerato cursado. Concretamente, en Gil, los alumnos que habían cursado bachillerato de ciencias, con una mayor carga en asignaturas de matemáticas, tenían actitudes más positivas que los alumnos que habían cursado bachillerato de letras.

Una cuestión de investigación diferente se refiere no a la formación previa sino al efecto de determinadas características de la asignatura de estadística sobre las actitudes o la ansiedad de los alumnos. Para estudiar esta cuestión ha sido habitual aplicar el cuestionario elegido tanto al inicio como al final del curso y tomar la diferencia entre ambas medidas como un indicador del efecto de la intervención. Los resultados de estas investigaciones apuntan bien a la inexistencia de cambios en las actitudes o la ansiedad de los alumnos (Elmore, Lewis & Bay, 1993; Sgoutas-Emch y Johnson, 1998; Shultz y Koshino, 1998), la existencia de cambios sólo en algunos aspectos de la ansiedad o de las actitudes (Elmore y Lewis, 1991; Sorge et al., 2000; Waters et al., 1988a) o a la existencia de mejoras en la actitud o reducción de la ansiedad (Forte, Healey y Campbell, 1994; Harvey et al., 1985; Tomazic y Katz, 1988; Perney y Ravid, 1990; Quilter y Chester, 2001; Roberts y Saxe, 1982).

En la mayoría de estas investigaciones la existencia de cambios en las actitudes se atribuye al hecho de haber cursado la asignatura de estadística (Harvey et al., 1985; Perney y Ravid, 1990; Roberts y Saxe, 1982; Sorge et al., 2000; Waters et al., 1988a) o a alguna característica de la impartición de las mismas, como el uso de un determinado tipo de software (Forte et al., 1994), el uso de un aula de informática (Elmore y Lewis, 1991), el uso de conferencias a través de la web (Quilter y Chester, 2001) o el uso de una determinada aproximación conceptual –aproximación no cuantitativa a la enseñanza de la estadística- (Tomazic y Katz, 1988).

Aunque en todos estos trabajos, los diseños de investigación incluían la aplicación a todos los alumnos de una medida al inicio y otra al final del curso; existen tres trabajos en los que además se usó un grupo control (Elmore y Lewis, 1991; Elmore et al., 1993; Sgoutas-Emch y Johnson, 1998).

Pensamientos auto-referidos

Por último, algunas de las investigaciones revisadas sugieren que los pensamientos referidos a las propias capacidades, sobre todo las capacidades relacionadas con las matemáticas y la estadística, condicionan las actitudes y la ansiedad de los alumnos. Los resultados de los trabajos revisados muestran correlaciones moderadas e incluso altas entre ambos tipos de variables. Así, por ejemplo, Zeidner (1991) encuentra una correlación inversa significativa entre la ansiedad hacia la estadística medida con el SAI y las auto-percepciones acerca de las habilidades matemáticas ($r = -0.38$). Perney y Ravid (1990), por su parte, hallaron correlaciones desde 0.24 a 0.74 entre una escala de autoconcepto en matemáticas y las subescalas del ATS medidas al inicio y al final del curso.

La relación entre el autoconcepto en matemáticas y la ansiedad hacia la estadística ha sido estudiada también en Benson (1989) y en Bandalos et al. (1995), usando una escala elaborada por Benson para medir el autoconcepto y el TAI de Spielberger para medir la ansiedad hacia los exámenes de estadística, y modelos de ecuaciones estructurales en los que se integran ambas variables. En estas investigaciones se encuentra que cuanto más positivo sea el autoconcepto en matemáticas menor será la ansiedad hacia el examen de estadística, efecto que se observa en las puntuaciones en el TAI (Benson, 1989) y las puntuaciones en el TAI-Emotividad (Bandalos et al., 1995). En otro trabajo, Onwuegbuzie (2000a) usó un análisis de correlación canónica para estudiar qué combinación de las cinco dimensiones de la autopercepción relativas a las competencias académicas, medidas a través del SPPCS de Neemann y Harter (1986), podría estar correlacionada con alguna combinación de las dimensiones de la ansiedad hacia la estadística medida con el STARS. Los resultados de este análisis muestran que las dimensiones de la auto-percepción que permiten predecir mejor las seis dimensiones del STARS son, por este orden, la Habilidad intelectual percibida, la Competencia escolar percibida y la Creatividad percibida.

En cuanto a la autopercepción de las capacidades estadísticas, Schutz et al. (1998) preguntaron a sus alumnos acerca de la confianza que tenían en dominar el material de introducción a la estadística y encontraron correlaciones significativas entre las respuestas a esta pregunta y las subescalas SATS-Afecto ($r = 0.51$) y SATS-Valor ($r = 0.44$). Silva, Cazorla y Brito (1999), por su parte, solicitaron a los alumnos que clasificaran su conocimiento de estadística, optando entre ‘pésimo’, ‘pasable’ y ‘bueno’. Posteriormente usaron un análisis de varianza con los resultados de esa pregunta como medida de la autopercepción acerca de la capacidad estadística y las puntuaciones en la *Escala de actitudes em relação à Estatística* de Cazorla como variable dependiente. El resultado de este análisis mostró diferencias significativas entre los tres grupos, con actitudes más positivas para aquellos que consideraban tener más conocimientos de estadística.

Finney y Schraw (2003) han desarrollado medidas específicas de la autoeficacia de los alumnos en estadística y las han aplicado a una muestra de alumnos junto a otros cuestionarios, incluido el SATS de Schau et al. Los resultados de este estudio muestran correlaciones de las subescalas del SATS con una medida de la confianza en las propias capacidades para aprender las destrezas necesarias para resolver tareas específicas relacionadas con la estadística (SELS) que van desde 0.37 a 0.65.

6. CONCLUSIONES DE LA REVISIÓN REALIZADA

A pesar de su importancia para el proceso de enseñanza de la estadística, las investigaciones sobre las actitudes y la ansiedad hacia la estadística han sido y siguen siendo más escasas que las que se dan en campos afines, como el estudio de las actitudes hacia la ciencia o de las actitudes hacia las matemáticas (Gal et al., 1997). Esto no ha sido óbice, sin embargo, para que haya existido una cierta proliferación de instrumentos de medida que, en la mayoría de los casos, son usados en muy pocas investigaciones. Una de las consecuencias de este hecho es que no existen suficientes evidencias de la validez de muchos de los instrumentos de medida de las actitudes o la ansiedad hacia la estadística existentes.

En la revisión realizada hemos encontrado 17 instrumentos para medir actitudes o ansiedad hacia la estadística. Todos ellos son cuestionarios compuestos por ítems con opciones de respuesta propias de las escalas de valoración, habitualmente formatos de respuesta tipo Likert. Los cuestionarios de actitudes más usados son, por este orden, el ATS de Wise, el SATS de Schau y el SAS de Roberts. En cuanto a los cuestionarios de ansiedad hacia la estadística los más usados son el STARS de Cruise, seguido de la versión del MARS de Suinn elaborada por Plake y Parker. Los resultados más destacados de las investigaciones en las que se estudian las propiedades psicométricas de estos cuestionarios son las siguientes:

- a) La mayoría de las escalas que componen los cuestionarios analizados han mostrado una alta fiabilidad de consistencia interna. Por otra parte, sólo existen dos estudios de la fiabilidad entendida como estabilidad temporal de las puntuaciones, obteniéndose en ambos casos valores aceptables del coeficiente de estabilidad, para el ATS en un caso y el STARS en el otro.
- b) En cuanto a las evidencias de validez basadas en la estructura interna, la estructura del dominio de las actitudes hacia la estadística ha sido concebida teóricamente de muy diversas formas, desde una estructura unidimensional en el SAS de Roberts hasta una compuesta por cinco dimensiones en la EAE de Auzmendi. Lo mismo se puede decir del dominio de la ansiedad hacia la estadística, con propuestas que van desde la única dimensión del SAS de Pretorius y Norman hasta las seis dimensiones del STARS de Cruise et al. En todos los casos en los que se ha investigado si las respuestas de los sujetos a los cuestionarios respectivos se ajustaban a estas diferentes estructuras se han usado técnicas de análisis factorial. Los resultados de muchas de estas investigaciones apoyan las estructuras preconcebidas, siendo destacables las evidencias a favor de la estructura bidimensional del ATS obtenidas con técnicas de análisis factorial exploratorio y las evidencias a favor de la estructura de cuatro dimensiones del SATS obtenidas con técnicas de análisis factorial confirmatorio.

- c) En cuanto a las evidencias de validez del contenido, es destacable que para la mayoría de los cuestionarios analizados no se aporten este tipo de evidencias. Sólo con cuatro cuestionarios (STARS, ATS, SATS y QAQ) se realizan estudios para investigar el contenido de sus respectivos ítems a través de las evaluaciones de jueces expertos. Estas evaluaciones, básicamente juicios acerca de la congruencia ítem-dimensión, se realizan en las fases iniciales de desarrollo de los cuestionarios y sirven fundamentalmente como un método de selección de ítems.
- d) En cuanto a las evidencias de validez basadas en las relaciones con otras variables, cuando las puntuaciones en los cuestionarios se han usado para predecir el rendimiento de los alumnos en las asignaturas de estadística se ha encontrado que ambas variables están relacionadas débilmente. En cambio, en aquellos estudios en que se han usado más de uno de los cuestionarios revisados se ha hallado por lo general una alta convergencia entre ellos. Con respecto a las variables que permiten predecir las actitudes o la ansiedad de los alumnos, tanto la formación previa en matemáticas y estadística como el autoconcepto acerca de las capacidades relacionadas con estas materias han mostrado relaciones consistentes con las puntuaciones en los cuestionarios de actitudes o ansiedad. Los resultados son menos claros cuando se usa el género de los alumnos, su edad u otras características personales para intentar predecir las actitudes.

A partir de esta revisión de las evidencias de validez es posible concluir que algunos de estos cuestionarios, sobre todo el ATS y el SATS, pueden ser usados con ciertas garantías para evaluar las actitudes hacia la estadística. De igual forma, consideramos que, en líneas generales, el STARS es un cuestionario adecuado para medir ansiedad hacia la estadística. Sin embargo, también se han detectado algunas carencias en el proceso de elaboración y validación de estos instrumentos que hacen recomendables nuevas investigaciones para el desarrollo de instrumentos de medida de las actitudes y la ansiedad hacia la estadística o para completar el análisis de los existentes con suficientes garantías de validez.

En este sentido, estamos de acuerdo con Gal et al. (1997) cuando afirmaban que uno de los principales problemas de la investigación en este campo es la falta de modelos teóricos que guíen los trabajos. Esta falta de fundamentación teórica se refleja, por ejemplo, en uno de los aspectos más importantes para la elaboración y validación de un instrumento de medida, a saber, la determinación de su estructura interna. En la práctica, los autores de cada uno de los cuestionarios revisados han propuesto estructuras diferentes del dominio de las actitudes hacia la estadística. Sin embargo, casi ninguno de estos autores ha justificado la elección de esas estructuras. En aquellos casos en los que se ha propuesto alguna justificación, por ejemplo en el SATS, éstas suelen estar basadas en el conocimiento, no necesariamente bien informado, de profesores y alumnos de estadística acerca de ese dominio conductual más que en una teoría acerca del mismo. Otra consecuencia de la falta de teorías se refleja en la dificultad para determinar claramente una red nomológica de relaciones entre las actitudes hacia la estadística y otras variables. Así, por ejemplo, en cuanto a la relación entre género y actitudes de los alumnos, no está claro si es esperable que los alumnos tengan actitudes más, menos o igual de favorables hacia la estadística que las alumnas.

Además de la falta de modelos teóricos, existen otras carencias más específicas de los estudios de elaboración y validación de estos instrumentos, algunas de las cuales comentamos a continuación. En primer lugar, puesto que muchos de estos cuestionarios se van a usar en distintos momentos del curso para estudiar la evolución de las actitudes o la ansiedad de los alumnos, es bastante conveniente estudiar la estabilidad de las puntuaciones, cosa que sólo se ha hecho en el caso del ATS y el STARS. Por otra parte, faltan estudios que aporten evidencias basadas en el contenido, y cuando éstos se realizan se usan tareas clásicas de evaluación del contenido que han sido cuestionadas debido a que a la información previa que se proporciona a los jueces condiciona en gran medida sus juicios (Sireci, 1998). Un último aspecto de carácter técnico que, desde nuestro punto de vista, constituye una limitación de estos estudios de validación es el uso casi exclusivo de técnicas de análisis factorial exploratorio en el estudio de la estructura interna de los cuestionarios, con la excepción del SATS.

A pesar de estas limitaciones, el análisis de los trabajos revisados permite extraer algunas recomendaciones para futuras investigaciones en este campo:

En primer lugar, consideramos que la estructura de cuatro dimensiones propuesta para el SATS es una candidata plausible para describir el dominio de las actitudes hacia la estadística. Por otra parte, no existen evidencias suficientes que permitan determinar si una estructura de dos dimensiones (Plake y Parker, 1982; Zeidner, 1991) para el dominio de la ansiedad hacia la estadística es más adecuada que una estructura de seis dimensiones (Cruise et al., 1985). Por ello deberían diseñarse investigaciones para poner a prueba diferentes estructuras internas para este dominio desde una óptica confirmatoria, en la línea de lo realizado por los autores del SATS (Schau et al., 1995; Dauphinee et al., 1997). Además, en los últimos años se ha propuesto una interesante alternativa para el estudio de la estructura de dominios conductuales (Deville, 1996), que creemos tiene gran potencial en el estudio de las actitudes y la ansiedad hacia la estadística. Se trata de una propuesta que pone el foco en una especificación clara del contenido, una redacción de ítems que se ajuste a esas especificaciones y el uso de técnicas de análisis de datos que permitan el análisis conjunto de las evidencias relativas a la estructura interna y al contenido. Tal como recomiendan los *Standards for Educational and Psychological Testing* (AERA/APA/NCME, 1999) en esta propuesta se integran diferentes tipos de evidencias de validez, lo cual viene además a subsanar una de las carencias detectadas en la revisión realizada, a saber, la escasez de estudios de validación de contenido.

En cuanto a la relación de las actitudes y la ansiedad hacia la estadística con el rendimiento de los alumnos en las asignaturas de estadística, estos constructos suelen mostrar una capacidad muy limitada para predecir el rendimiento. En los últimos años se han propuesto, sin embargo, interesantes modelos multivariados para la predicción del rendimiento en asignaturas de estadística (Lalonde y Gardner, 1993; Onwuegbuzie, 2003; Tremblay, Gardner y Heipel, 2000) que deberían ser puestos a prueba en nuevos estudios. Otra prometedora línea de trabajo en este ámbito, que sin embargo no ha sido investigada hasta el momento, es el estudio de la relación de las actitudes y la ansiedad hacia la estadística con algunas variables relacionadas con el rendimiento en estadística, como el abandono de las asignaturas optativas o la falta de asistencia a las convocatorias de examen. Los resultados del trabajo de Walsh y Ugumba-Agwunobi (2002) que muestran una relación significativa entre la ansiedad hacia la estadística y una cierta tendencia a retrasar la realización de tareas, parecen avalar la conveniencia de realizar dichos estudios acerca de algunas posibles manifestaciones conductuales de esta tendencia.

Por último, es necesario destacar que en la mayoría de trabajos analizados se ha encontrado que la formación previa en estadística y matemáticas parece ser un antecedente bastante consistente de las actitudes y la ansiedad hacia la estadística. Esta relación, sin embargo, tiende a ser más débil que la encontrada entre las actitudes o la ansiedad con indicadores de la percepción que tienen los alumnos acerca de sus capacidades en estas materias (Schutz et al., 1998). Por ello consideramos conveniente que en los modelos de trabajo de futuras investigaciones se incorporen medidas específicas de este aspecto del autoconcepto, como las medidas de autoeficacia en estadística desarrolladas por Finney y Schraw (2003).

REFERENCIAS

- Aiken, L. R., y Dreger, R. M. (1961). The effect of attitudes in performance in mathematics. *Journal of Educational Psychology*, 52(1), 19-24.
- American Educational Research Association, the American Psychological Association y the National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, DC, EE. UU.: AERA.
- Araki, L. T., y Shultz, K. S. (1995, Abril). *Student attitudes toward statistics and their retention of statistical concepts*. Comunicación presentada al Annual Meeting of the Western Psychological Association, Los Angeles.
- Auzmendi, E. (1991). *Evaluación de las actitudes hacia la estadística en estudiantes universitarios y factores que las determinan*. Tesis doctoral no publicada, Universidad de Deusto, Bilbao, España.

- Auzmendi, E. (1992). *Las actitudes hacia la matemática-estadística en las enseñanzas medias y universitaria*. Bilbao, España: Ediciones Mensajero.
- Baloglu, M. (2002). Psychometric properties of the Statistics Anxiety Rating Scale. *Psychological Reports, 90*, 315-325.
- Baloglu, M. (2003). Individual differences in statistics anxiety among collage students. *Personality and Individual Differences, 34*, 855-865.
- Bandalos, D. L., Yates, K., y Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal of Educational Psychology, 87*(4), 611-623.
- Bending, A. W., y Hughes, J. B. (1954). Student attitude and achievement in a course in introductory statistics. *Journal of Educational Psychology, 45*, 268-276.
- Benson, J. (1989). Structural components of statistical test anxiety in adults: An exploratory model. *Journal of Experimental Education, 57*, 247-261.
- Bessant, K. C. (1997). The development and validation of scores on the Mathematics Information Processing Scale (MIPS). *Educational and Psychological Measurement, 57*(5), 841-857.
- Betz, N. E. (1978). Prevalence, distribution and correlates of math anxiety in college students. *Journal of Counseling Psychology, 25*, 441-448.
- Birenbaum, M., y Eylath, S. (1994). Who is afraid of statistics? Correlates of statistics anxiety among students of educational sciences. *Educational Research, 36*(1), 93-98.
- Carmona, J., y Moreno, R. (1999). Structuring the domain of attitudes toward statistics. En R. Meyer Schweizer (Ed.), *Facet Theory Conference: Design and Analysis* (pp. 323-334). Berna, Suiza: FTA-c/o Institut für Soziologie.
- Cashin, S. E. (2001). *Effects of mathematics self-concept, perceived self-efficacy, and attitudes toward statistics on statistics achievement*. Tesis doctoral no publicada, Southern Illinois University, Carbondale, IL, EE. UU.
- Cashin, S. E., y Elmore, P. B. (1997, Marzo). *Instruments used to assess attitudes toward statistics: A psychometric evaluation*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Chicago.
- Cashin, S. E., y Elmore, P. B. (2000, Octubre). *The Survey of Attitudes Toward Statistics scale: A construct validity study*. Comunicación presentada al Annual Meeting of the Mid-Western Educational Research Association, Chicago, IL, EE. UU.
- Cazorla, I. M., Silva, C. B., Vendramini, C., y Brito, M. R. F. (1999). *Adaptação e validação de uma Escala de Atitudes em Relação à Estatística*. Comunicación presentada a la Conferência Internacional Experiências e Expectativas do Ensino da Estatística: Desafios para o Século XXI, Florianópolis, Brasil. [Online: <http://www.inf.ufsc.br/cee/pasta1/art2.html>].
- Chang, L. (1996). Quantitative Attitudes Questionnaire: Instrument development and validation. *Educational and Psychological Measurement, 56*(6), 1037-1042.
- Cherian, V. I., y Glencross, M. J. (1997). Sex, socio-economic status, and attitude toward applied statistics among postgraduate education students. *Psychological Reports, 80*, 1385-1386.
- Crocker, L., y Algina, J. (1986). *Introduction to classical and modern test theory*. Nueva York: Holt, Rinehart & Winston.
- Cruise, R. J., Cash, R. W., y Bolton, D. L. (1985). Development and validation of an instrument to measure statistical anxiety. *ASA Proceedings of the Section on Statistical Education* (pp. 92-97). Alexandria, VA, EE. UU.: American Statistical Association.
- Cuesta, M., Rifá, H., y Herrero, F. J. (2001, Septiembre). *Un estudio exploratorio, en estudiantes de psicología, de una escala de actitudes hacia la estadística*. Póster presentado en el VII Congreso de Metodología de las Ciencias Sociales y de la Salud, Madrid. Resumen descargado el 6 de Marzo de 2002 de http://www.uned.es/VIIICongresoMetodologia/poster/sesionI_titulos.htm

- Dauphinee, T. L., Schau, C., y Stevens, J. J. (1997). Survey of Attitudes Toward Statistics: Factor structure and factorial invariance for women and men. *Structural Equation Modelling*, 4(2), 129-141.
- Deville, C. W. (1996). An empirical link of content and construct validity evidence. *Applied Psychological Measurement*, 20(2), 127-139.
- Dillon, K. M. (1982). Statisticophobia. *Teaching of Psychology*, 9(2), 117.
- Dutton, W. A. (1951). Attitudes of prospective teachers toward mathematics. *Elementary School Journal*, 52, 84-90.
- Eagly, A. H., y Chaiken, S. (1998). Attitude structure and function. En D. T. Gilbert, S. T. Fiske y G. Lindzey (Eds.), *The handbook of social psychology* (4ª Ed, Vol. 1, pp. 269-322). Nueva York, NY, EE. UU.: McGraw-Hill.
- Elmore, P. B., y Lewis, E. L. (1991, Abril). *Statistics and computer attitudes and achievement of students enrolled in Applied Statistics: Effect of a computer laboratory*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Chicago.
- Elmore, P. B., Lewis, E. L., y Bay, M. L. G. (1993). *Statistics achievement: A function of attitudes and related experiences*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Atlanta, GA, EE. UU. (ERIC Document Reproduction Service No. ED360324)
- Estrada, A. (2001). *Actitudes hacia la estadística e instrumentos de evaluación*. Comunicación presentada a las Jornadas Europeas de Estadística: La Enseñanza y la Difusión de la Estadística. Palma de Mallorca, España.
[Online: http://www.caib.es/ibae/esdeveniment/jornades_10_01/doc/estadistica.doc]
- Estrada, A. (2002). *Análisis de las actitudes y conocimientos estadísticos elementales en la formación del profesorado*. Tesis doctoral no publicada, Universidad Autónoma de Barcelona, España.
- Finney, S. J., y Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, 28, 161-186.
- Fitzgerald, S. M. (1997). *The relationship between anxiety and statistics achievement: A meta-analysis*. Tesis doctoral no publicada, University of Toledo, Toledo, OH, EE. UU.
- Fitzgerald, S. M., Jurs, S. J., y Hudson, L. M. (1996). A model predicting statistics achievement among graduate students. *College Student Journal*, 30(3), 361-366.
- Forte, J. A., Healey, J., y Campbell, M. H. (1994). Does MicroCase statistical software package increase the competence and comfort of undergraduate social work and social science majors? *Journal of Teaching in Social Work*, 10(1/2), 99-115.
- Gal, I. y Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Toward an assessment framework. *Journal of Statistics Education*, 2(2). Descargado el 6 de Marzo de 2003 de: <http://www.amstat.org/publications/jse/v2n2/gal.html>
- Gal, I., Ginsburg, L., y Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. En I. Gal y J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 37-51). Netherlands: IOS Press.
- Gil, J. (1999). Actitudes hacia la estadística. Incidencia de las variables sexo y formación previa. *Revista Española de Pedagogía*, 214, 567-590.
- Glencross, M. J., y Cherian, V. J. (1992). Attitudes toward applied statistics of postgraduate education students in Transkei. *Psychological Reports*, 70, 67-75.
- Green, K. E. (1993, Abril). *Affective, evaluative and behavioral components of attitudes toward statistics*. Comunicación presentada en Annual Meeting of the American Educational Research Association, Atlanta, GA, EE. UU.
- Green, K. E. (1994, Abril). *The affective component of attitude in statistics instruction*. Comunicación presentada en Annual Meeting of the American Educational Research Association, Nueva Orleans, LA, EE. UU.

- Harvey, A. L., Plake, B. S., y Wise, S. L. (1985). *The validity of six beliefs about factors related to statistics achievement*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Chicago, IL, EE. UU. (ERIC Document Reproduction Service No. ED262965)
- Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development*, 34(3), 177-189.
- Huedo, T., López, J. A., Martínez, R., y Nortes, A. (2003). *Contenidos y actitudes en estadística: Un estudio en maestros en formación*. Comunicación presentada al 27 Congreso Nacional de Estadística e Investigación Operativa, Lleida, España.
[Online: http://www.udl.es/usuarios/seio2003/treballs/05_1_5.pdf]
- Kottke, J. L. (2000). Mathematical proficiency, statistics knowledge, attitudes toward statistics, and measurement course performance. *College Student Journal*, 34(3), 334-347.
- Lalonde, R. N., y Gardner, R. C. (1993). Statistics as a second language: A model for predicting performance in psychology students. *Canadian Journal of Behavioural Science*, 25(1), 108-125.
- McCall, C. H., Belli, G., y Madjidi, F. (1991). The complexities of teaching graduate students in educational administration introductory statistical concepts. *PICTeachSt3*, 2, 495-497.
- Mejía, J. C. (1995). *Reliability and validity study of a Spanish translation of the Statistical Attitude Survey*. Tesis doctoral no publicada, Pennsylvania State University, PA, EE. UU.
- Miller, R. B., Behrens, J. T., Green, B. A., y Newman, D. (1993). Goals and perceived ability: Impact on student valuing, self-regulation and persistence. *Contemporary Educational Psychology*, 18, 2-14.
- Musch, J., y Bröder, A. (1999). Test anxiety versus academic skills: A comparison of two alternative models for predicting performance in a statistics exam. *British Journal of Educational Psychology*, 69, 105-116.
- Nasser, F. (1999). *Prediction of college students achievement in Introductory Statistics Course*. Comunicación presentada a la 52nd ISI –International Statistical Institute- Session, Helsinki, Finlandia. [Online: <http://www.stat.fi/isi99/proceedings/arkisto/varasto/nass0341.pdf>]
- Neemann, J., y Harter, S. (1986). *Manual for the self-perception profile for college students*. Manuscrito no publicado, University of Denver, CO, EE. UU.
- Onwuegbuzie, A. J. (1993). The interaction of statistics test anxiety and examination condition in statistics achievement of post baccalaureate non statistics majors. *Dissertation Abstracts International*, 54(12), 4371A.
- Onwuegbuzie, A. J. (1995). Statistics test anxiety and female students. *Psychology of Women Quarterly*, 19, 413-418.
- Onwuegbuzie, A. J. (1998). Teachers' attitudes toward statistics. *Psychological Reports*, 83, 1008-1010.
- Onwuegbuzie, A. J. (1999). Statistics anxiety among African American graduate students: An effective filter? *Journal of Black Psychology*, 25(2), 189-209.
- Onwuegbuzie, A. J. (2000a). Statistics anxiety and the role of self-perceptions. *Journal of Educational Research*, 93(5), 323-335.
- Onwuegbuzie, A. J. (2000b). Attitudes toward statistics assessments. *Assessment and Evaluation in Higher Education*, 25(4), 321-339.
- Onwuegbuzie, A. J. (2003). Modeling statistics achievement among graduate students. *Educational and Psychological Measurement*, 63(6), 1020-1038.
- Onwuegbuzie, A. J. (en prensa). Prevalence of statistics anxiety among graduate students. *Journal of Research in Education*.
- Onwuegbuzie, A. J., y Daley, C. E. (1997). The role of multiple intelligences in statistics anxiety. *Research and Reflections on Leadership and Organizational Change*, 3(2). Descargado el 28 de Octubre de 2000 de: <http://www.gonzaga.edu/rr/v3n2/onwuegbuzie.htm>

- Onwuegbuzie, A. J., y Daley, C. E. (1999). Perfectionism and statistics anxiety. *Personality and Individual Differences*, 26, 1089-1102.
- Onwuegbuzie, A. J., Slate, J., Paterson, F., Watson, M., y Schwartz, R. (2000). Factors associated with underachievement in educational research courses. *Research in the Schools*, 7, 53-65.
- Onwuegbuzie, A. J., y Wilson, V. A. (2003). Statistics anxiety: nature, etiology, antecedents, effects and treatments-a comprehensive review of the literature. *Teaching in Higher Education*, 8(2), 195-209.
- Perney, J., y Ravid, R. (1990, Abril). *The relationship between attitudes towards statistics, math self-concept, test anxiety and graduate students' achievement in an introductory statistics course*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Boston, MA, EE. UU.
- Plake, B. S., y Parker, C. S. (1982). The development and validation of a revised version of the Mathematics Anxiety Rating Scale. *Educational and Psychological Measurement*, 42, 551-557.
- Pretorius, T. B., y Norman, A. M. (1992). Psychometric data on the Statistics Anxiety Scale for a sample of south African students. *Educational and Psychological Measurement*, 52(4), 933-937.
- Quilter, S. M., y Chester, C. (2001). The relationship between web-based conferencing and instructional outcomes. *International Journal of Instructional Media*, 28(1), 13-22.
- Rhoads, T. R., y Hubele, N. F. (2000). Student attitudes toward statistics before and after a computer-integrated introductory statistics course. *IEEE Transactions on Education*, 43(2), 182-187.
- Roberts, D. M., y Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement*, 40, 235-238.
- Roberts, D. M., y Reese, C. M. (1987). A comparison of two scales measuring attitudes toward statistics. *Educational and Psychological Measurement*, 47, 759-764.
- Roberts, D. M., y Saxe, J. E. (1982). Validity of a Statistics Attitude Survey: A follow-up study. *Educational and Psychological Measurement*, 42, 907-912.
- Rosenthal, B. (1992). No more sadistics, no more sadists, no more victims. *UMAP Journal*, 13, 281-290.
- Royse, D., y Rompf, E. L. (1992). Math anxiety: A comparison of social work and non-social work students. *Journal of Social Work Education*, 28(3), 270-277.
- Sánchez-López, C. R. (1996). Validación y análisis ipsativo de la Escala de Actitudes hacia la Estadística (EAE). *Análisis y Modificación de Conducta*, 22, 799-819.
- Sarason, I. G. (1972). Experimental approaches to test anxiety: Attention and the assessment of information. En C. D. Spielberger (Ed.), *Anxiety: Current trends in theory and research* (vol. 2, pp. 283-403). Nueva York, NY, EE. UU.: Academic Press.
- Schau, C., Dauphinee, T. L., y Del Vecchio, A. (1993, Abril). *Evaluation of two surveys measuring students' attitudes toward statistics*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Atlanta, GA, EE. UU.
- Schau, C., Stevens, J., Dauphinee, T. L., y Del Vecchio, A. (1995). The development and validation of the Survey of Attitudes Toward Statistics. *Educational and Psychological Measurement*, 55, 868-875.
- Schuessler, K. F. (1993). Scaling social life feelings: Persistent issues. En D. Krebs y P. Schmidt (Eds.), *New directions in attitude measurement* (pp. 21-40). Berlin, Alemania: Walter de Gruyter.
- Schutz, P. A., Drogosz, L. M., White, V. E., y Distefano, C. (1998). Prior knowledge, attitude, and strategy use in an introduction to statistics course. *Learning and Individual Differences*, 10(4), 291-308.
- Scott, J. S. (2002). *Modeling aspects of students' attitudes and performance in an undergraduate introductory statistics course*. Tesis doctoral no publicada, University of Georgia, Athens, GA, EE. UU.
- Sgoutas-Emch, S. A., y Johnson, C. J. (1998). Is journal writing an effective method of reducing anxiety toward statistics?. *Journal of Instructional Psychology*, 25(1), 49-57.

- Shultz, K. S., y Koshino, H. (1998). Evidence of reliability and validity for Wise's Attitude Toward Statistics scale. *Psychological Reports*, 82, 27-31.
- Sigurdsson, J. F. (1991). Computer experience, attitudes toward computers and personality characteristics in psychology undergraduates. *Personality and Individual Differences*, 12(6), 617-624.
- Silva, C. B., Cazorla, I. M., y Brito, M. R. F. (1999). *Concepções e atitudes em relação à estatística*. Comunicación presentada a la Conferência Internacional Experiências e Expectativas do Ensino da Estatística: Desafios para o Século XXI, Florianópolis, Brasil.
[Online: <http://www.inf.ufsc.br/cee/pasta1/art2.html>]
- Sireci, S. G. (1998). Gathering and analysing content validity data. *Educational Assessment*, 5(4), 299-321.
- Sorge, C., y Schau, C. (2002). *Impact of engineering students' attitudes on achievement in statistics: A structural model*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Nueva Orleans, LA, EE. UU.
[Online: <http://www.unm.edu/~cschau/AERA2002.pdf>]
- Sorge, C., Schau, C., Hubele, N., y Kennedy, D. (2000, Agosto). *Measuring engineering students' attitudes toward statistics and probability using Internet data collection techniques*. Comunicación presentada al Joint Statistical Meetings, Indianapolis, IN, EE. UU.
- Spielberger, C. (1980). *Preliminary Professional Manual for the Test Anxiety Inventory*. Palo Alto, CA, EE. UU.: Consulting Psychologists Press.
- Suanpang, P., Petocz, P., y Kalceff, W. (2003). *Student attitudes to learning business statistics online vs traditional methods*. Comunicación presentada a HERDSA Conference 2003, Christchurch, Nueva Zelanda.
[Online: <http://surveys.canterbury.ac.nz/herdsa03/pdfsnon/N1177.pdf>]
- Suinn, R. M., Eddie, C. A., Nicoletti, J., y Spinelli, P. R. (1972). The MARS, a measure of mathematics anxiety: Psychometric data. *Journal of Clinical Psychology*, 28, 373-375.
- Sutarso, T. (1992). *Some variables in relation to students' anxiety in learning statistics*. Comunicación presentada a Annual Meeting of the Mid-South Educational Research Association, Knoxville, TN, EE. UU. (ERIC Document Reproduction Service No. ED353334).
- Tomazic, T. J., y Katz, B. M. (1988). Statistical anxiety in Introductory Applied Statistics. *ASA Proceedings of the Section on Statistical Education* (pp. 239-244). Alexandria, VA, EE. UU.: American Statistical Association.
- Tremblay, P. F., Gardner, R. C., y Heipel, G. (2000). A model of the relationship among measures of affect, aptitude, and performance in introductory statistics. *Canadian Journal of Behavioural Science*, 32(1), 40-48.
- Velandrino, A. P., y Parodi, L. H. (1999). *La Escala de Actitudes Hacia la Estadística (EAE): Desarrollo y Propiedades Psicométricas*. Comunicación presentada a la Conferência Internacional Experiências e Expectativas do Ensino da Estatística: Desafios para o Século XXI, Florianópolis, Brasil. [Online: <http://www.inf.ufsc.br/cee/comunica/EAE.html>]
- Walsh, J. J., y Ugumba-Agwunobi, G. (2002). Individual differences in statistics anxiety: the roles of perfectionism, procrastination and trait anxiety. *Personality and Individual Differences*, 33, 239-251.
- Waters. L. K., Martelli, T. A., Zakrajsek, T., y Popovich, P. M. (1988a). Attitudes toward statistics: An evaluation of multiple measures. *Educational and Psychological Measurement*, 48, 513-516.
- Waters. L. K., Martelli, T. A., Zakrajsek, T., y Popovich, P. M. (1988b). Factor analyses of two measures of attitudes toward statistics. *Educational and Psychological Measurement*, 48, 1037-1041.
- Watson, F., Kromrey, J., Ferron, J., Lang, T., y Hogarty, K. (2003). *An assessment blueprint for Encstat: A statistics anxiety intervention program*. Comunicación presentada al AERA Annual Meeting, San Diego, CA, EE. UU.

- Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement*, 45, 401-405.
- Wisembaker, J., Nasser, F., y Scott, J. S. (1999). *A cross-cultural comparison of path models relating attitudes about and achievement in Introductory Statistics Courses*. Comunicación presentada a la 52nd ISI -Internacional Statistical Institute- Session, Helsinki, Finlandia. [Online: <http://www.stat.fi/isi99/proceedings/arkisto/varasto/wise0263.pdf>]
- Woehlke, P. L. (1991). *An examination of the factor structure of Wise's Attitude Toward Statistics scale*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Chicago, IL, EE. UU. (ERIC Document Reproduction Service No. ED337500)
- Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British Journal of Educational Psychology*, 61(3), 319-328.

JOSÉ CARMONA MÁRQUEZ
Departamento de Psicología
Universidad de Huelva
Avda. Fuerzas Armadas s/n.
21007 Huelva, España

CAN KINDERGARTEN CHILDREN BE SUCCESSFULLY INVOLVED IN PROBABILISTIC TASKS?

SONIA KAFOUSSI
University of Aegean, Greece
kafoussi@rhodes.aegean.gr

SUMMARY

This paper describes a classroom teaching experiment, concerning the concept of probability, with children aged 5 in a kindergarten school. The teaching experiment was based on constructivist and interactionist theories about the learning of school mathematics and lasted one month. The collection of the information was based on the tape-recorded interviews with the children (each child was interviewed prior to the research program, at the end of the program and one month later) and the videotaped teaching sessions. During the program, we identified three critical steps in the development of the children's probabilistic thinking: a) the interpretation of the "different" outcomes in a two stage experiment, b) the acceptance of the realization of the experiment for resolving their conflicting viewpoints, and c) estimating the outcomes in a problem. At the end of the program the majority of the children managed to overcome their subjective interpretations and seemed to develop a primitive quantitative reasoning in probabilistic tasks.

Keywords: Statistics education research; Probabilistic thinking; Kindergarten; Instructional activities

1. INTRODUCTION

Recently, the interest of mathematics educators in stochastic concepts has increased, as a consequence of the recognition of the role of these concepts in our daily life. The use of data and graphs to communicate information, as well as the ability to take decisions under uncertainty, is increasing in our society. Moreover, this interest becomes even greater due to the fact that students have much difficulty with these concepts, as many researchers have shown (cf. Fischbein & Schnarch, 1997; Kapadia & Borovcnik, 1991, pp. 73-105; Shaughnessy, 1992).

In order to cope with students' problems concerning the concept of probability, it has recently been proposed that students could be engaged in probabilistic tasks at an early age (cf. National Council of Teachers of Mathematics, 1989). Although a lot of research has been done on young children's probabilistic thinking (cf. Acredolo et al., 1989; English, 1993; Fischbein & Gazit, 1984; Piaget & Inhelder, 197; Schroeder, 1988), the results sometimes seem to be contradictory, as the researchers start from different theoretical viewpoints. These differences are mainly related to the analysis of this concept and to the role of the instruction towards its acquisition. As is well known, the nature of this concept has been connected with different interpretations (classical, frequentist, subjective) and the comparison of them is not always an easy task. Moreover, the work of Piaget and his colleagues and the work of Fischbein seem to have crucial differences with respect to their educational implications (cf. Greer, 2001).

However, substantial research in the classroom is still necessary on the development of children's thinking on probabilistic tasks. Towards this effort, this paper provides focus on the development of children's probabilistic thinking in a kindergarten classroom. More specifically, we will present the learning opportunities that occurred in the social setting of the classroom, as the children tried to

overcome their subjective interpretations and develop a primitive quantitative reasoning when they were engaged in probabilistic tasks.

2. THEORETICAL BACKGROUND

The research program described in this paper is based on the cognitive framework that Jones and his colleagues (1997, 1999) have constructed in connection with the development of children's thinking in probability. According to this model, four key constructs have been used for the understanding of the probability concept: sample space, probability of an event, probability comparisons and conditional probability. Furthermore, young children's probabilistic thinking has been described across four levels for each of the four constructs: the subjective level, the transitional level, the informal quantitative level and the numerical level. For the purpose of this paper only the first and the second levels of their theoretical framework are presented in full here.

At the first level, the children:

- *Can list an incomplete set of outcomes for a one-stage experiment.*
- *Predict most/least likely events based on subjective judgments.*
- *Recognize certain and impossible events.*
- *Compare the probability of an event in two different sample spaces, usually based on various subjective or numeric judgments.*
- *Cannot distinguish "fair" probability situations from "unfair" ones.*
- *Following one trial of a one-stage experiment do not give a complete list of outcomes even though the complete list was given prior to the first trial.*
- *Recognize when certain and impossible events arise in a non-replacement situation.*

(p. 111, Jones et al., 1997).

At the second level, the children:

- *List a complete set of outcomes for a one-stage experiment and sometimes list a complete set of outcomes for a two-stage experiment using limited and unsystematic strategies.*
- *Predict most/least likely events based on quantitative judgments, but may revert to subjective judgments.*
- *Make probability comparisons based on quantitative judgments (may not quantify correctly and may have limitations when non-contiguous events are involved).*
- *Begin to distinguish "fair" probability situations from "unfair" ones.*
- *Recognize that the probability of some events changes in a non-replacement situation, however recognition is incomplete and is usually restricted only to events that have previously occurred.*

(p. 111, Jones et al., 1997).

The above findings were used as broad guidelines for the organization and the development of the instructional activities given to the children as well as the analysis of the children's activity and the interpretation of the results. The development of the instructional activities given to the children was related to their abilities according to these levels. Therefore, the instructional activities were concerned with the listing of all the possible outcomes for a one and a two-stage experiment, the prediction of the most/least likely event in a random experiment, the probability comparisons, the listing of the possible events after the realization of one trial in a random experiment and the distinction between fair and unfair games.

Moreover, the classroom teaching experiment was based on the constructivist and the interactionist views about mathematical learning. According to these approaches, mathematical learning is characterized as both an individual and collective process. Learning of school mathematics is a process in which students reorganize their mathematical activity to resolve situations that they find personally problematic as well as a process of enculturation into the mathematical practices of wider society (Cobb & Bauersfeld, 1995). Learning opportunities can arise for students

from their personal engagement with the mathematical activities as well as from their interaction with the other members of the classroom. Teaching of school mathematics can be characterized as a process in which the students and the teacher negotiate their mathematical meanings and interactively constitute the truth about a “taken-as-shared mathematical reality” (Cobb, Wood, Yackel & McNeal, 1992).

3. METHODOLOGY

All fifteen children that attended the kindergarten classroom in a typical state school of Athens participated in the classroom teaching experiment. The research project commenced in November 2001 and it lasted one month. The children were engaged in mathematical activities related to probability, three times per week. Every lesson lasted 30 minutes. The program was implemented in collaboration with an experienced and well-informed kindergarten teacher.

In the course of the program, we developed a set of instructional activities on the concept of probability concerned with the following themes: sample space, probability of an event, probability comparisons and conditional probability. The instructional activities used in the classroom were developed before and during the progress of the program, according to the evolution of the children’s ideas, so that the activities could enhance problematic situations for them.

A typical lesson can be described as follows. The teacher introduced the instructional activity to the whole class, the children made their predictions about the concrete problem by explaining their thinking and the teacher recorded the different ideas of the children. The children checked their predictions through the realization of the experiment and they recorded the results. Then, they compared their initial predictions to the experimental results and discussed the solution of the problem.

The children had not received any previous instruction in probability theory. Moreover, they had not yet received any instruction in whole numbers. Each child was interviewed for one hour prior to the instructional program, at the end of the program and one month later. The interview was realized by the researcher and included 14 tasks. The children were asked:

- a) in 5 tasks to report all the possible outcomes for a probability situation (for example, “what may happen if a die with different colors is thrown?” or “In a transparent bag we put three balls (1 red, 1 green and 1 blue). Imagine that you close your eyes and you draw two balls from the bag. Which colored balls could come out?”),
- b) in 4 tasks to predict the most likely outcome in a random experiment (for example, to predict which color is “the easiest” to be drawn from a box with colored balls or to come out if we turn a colored spinner),
- c) in 2 tasks to choose the probability situation for the most likely realization of an event, that is, to predict among different sample spaces, the one for the most likely realization of an event (for example, we put in a transparent box 2 red balls and 1 white ball and in another transparent box 1 red ball and 1 white ball and asked the children which box they would choose, if they wanted to draw a white ball), and
- d) in 3 tasks to report the outcomes of an experiment, after they had seen the outcome of one trial (for example, in a box we put 6 balloons (2 yellow, 2 blue and 2 green). The children closed their eyes and they drew one balloon. They got the balloon out of the box. Then they were asked to predict what could happen in the next trial).

All the items were presented with the use of manipulative materials.

All the teaching sessions were videotaped and the interviews were tape-recorded. The transcripts of the interviews and the teaching sessions provided the data for the analysis of the students’ learning. Analytical descriptive narrative was the method used for the analysis of the results (Erickson, 1986). The choice of this method was considered the most appropriate as:

Analytic narrative is the foundation of an effective report of fieldwork research. The narrative vignette is a vivid portrayal of the conduct of an event of everyday life, in which the sights and sounds of what was being said and done are described in the natural sequence of their occurrence in real time (Erickson, 1986, pp. 149-150).

4. THE SEQUENCE OF THE INSTRUCTIONAL ACTIVITIES

All the instructional activities were related to the interests and the experiences that children have at this age (5 years old). They were presented through small stories in puppet shows (where the basic persons were a grandfather, a grandmother and a squirrel), dramatic metaphors (teacher in role and children in role play) and games. All fifteen children participated together in the realization of every activity. Four categories of activities were discussed with the children:

In the first category of the activities the children were asked to recognize and distinguish events that always, sometimes or never happen in their daily life as well as in random experiments. The children were asked to make certain movements, according to the instructions of the teacher, when they heard these words and they were also asked to talk about events that always, never or sometimes happen at school. Then, they were asked to comment on different pictures from everyday life, by discussing which of these events take place always, sometimes or never and to construct a poster with the title “never”, “sometimes” and “always”. These activities gave the opportunity to the children to begin to estimate that an event could be certain, possible or impossible.

Afterwards, they had the opportunity to talk about the same words in random experiments. The random experiments were concerned with two or three possible outcomes. For example, the two heroes from the puppet show, the grandmother and the grandfather, were in complete disagreement about the color that could occur if they tossed up a “piece” with two colors (a “piece” is used in the Greek game of backgammon and in our activity it was presented like a coin with two different colored sides). The grandfather said that the red color would always occur and the grandmother said that sometimes the red color would occur and sometimes the yellow. The children had to express their opinion by discussing which of the heroes was right and why, and then to realize the experiment.

When the children had been engaged in this category of activities, they had the opportunity to concentrate on the colors that existed in front of them in the experiments. This was a significant step in their thinking. As we had observed from the first interview, the children sometimes mentioned colors that they desired but which did not exist in the task. Furthermore, we had the opportunity to initiate them into the process of the experiment. However, we should note that, at this phase, the realization of the experiment did not always influence the children’s answers. This means that some children insisted that a specific color could always come out, even though the results of the experiment did not support this opinion.

The second category of the instructional activities was concerned more systematically with the sample space. The children had to discuss all the possible events in a one or a two stage random experiment. Especially for the two stage experiments, the instructional activities were concerned with ordered and unordered pairs. For example, there were activities like the following: a) They had to design Christmas cards with two flowers, in order to help Father Christmas give the cards together with his presents to the children. The children had to pick one color from a box with three colors to paint the first flower and then they had to put it back. Then they had to pick one color for the second flower. After painting the flowers, they would discuss how many different cards had been made by the class. b) The squirrel wanted to make a curtain for the window in his house. He went in a shop that sells curtain material and the shop assistant showed him three plain materials with different colors. He would like the curtain to be made with two different materials. What colors could the curtain have? (The children were shown the materials).

The third category of the activities was concerned with the probability of an event and the probability comparisons. The tasks were related to situations that all the possible outcomes were equally likely. We used boxes with balls, dice and spinners. The children were asked to predict the color that was “the easiest” to come out or the color that could come out “more often”. For example, we discussed with the children activities like the following: a) The children were asked to help the

squirrel to solve a problem he had: “Every morning my mother puts fruits in the basket for my lunch at school. Today, she gave me one apple and three oranges. I will close my eyes and take a fruit at random, because I like both fruits. Which fruit is the easiest to come out?” b) We showed the children one spinner with two colors, red and green, the proportions of the colors were $\frac{3}{4}$ and $\frac{1}{4}$ respectively. The children were separated into two groups: the “red” and the “green”. They played the following game: “One child from every group turns the spinner. According to the color that comes out, the respective group gains one marble. The winner is the team that has gained the most marbles after 20 turns.” After the end of the game the children had to discuss the fairness of the game.

The last category of the activities was related to conditional probability. The children were asked to find all the possible outcomes in a random experiment after having completed one trial in a one-stage experiment. Moreover, they had the opportunity to discuss if an event becomes more or less likely to appear according to the results of previous trials, in non-replacement experiments. For example, the children participated in activities like the following. The squirrel would like to decorate the Christmas tree with colored balloons. He went in a shop and he found a box with red and yellow balloons (5 red and 3 yellow). He decided to pick a balloon at random. The squirrel picked one balloon out of the box and showed it to the children. The children were asked to predict the color that was “the easiest” to come out as the second balloon. Then, the squirrel picked a second balloon out of the box and showed it to the children. The children were then asked to predict the color that was “the easiest” to come out as the third balloon.

5. RESULTS

In sections 5.1 to 5.3 we present the children’s answers at the interviews prior to the teaching experiment, the critical moments on the development of the children’s probabilistic thinking during the teaching experiment and their solutions in the interview tasks after the teaching experiment.

5.1. BEFORE THE TEACHING EXPERIMENT

Prior to the teaching experiment, all the participating children seemed to interpret the tasks in probability in a subjective manner. However, we could identify some qualitative differences in their thinking. More specifically, the children’s responses could be classified in two general categories.

In the first category (4 children), the children gave answers that were strongly influenced by their favorite color for all the tasks. Ada was a representative for this category. She answered the red color as the only one that could appear in all the tasks. So, she reported only one outcome in a one or two stage experiment (e.g. she gave only one pair in a two-stage experiment including the red color) and she insisted that the red color was the most likely (“the easiest”) to come out in all the experiments. Having already seen the results of one trial in an experiment without replacement, when she was asked to describe the possible outcomes on the next trial, she repeated the color that she had got in the first trial.

In the second category, the children could give all the outcomes for a one-stage experiment. However, they had great difficulty in a two-stage experiment. For example, in the following task: “In a transparent bag we have put three balls (1 red, 1 green and 1 blue). Imagine that you close your eyes and you draw two balls from the bag. Which colored balls could come out?” the children gave only one combination. These children considered that they could not give another combination, because they had only three colors in the bag and so “it is only one that is left over”. The children answered in a similar way in tasks designed in terms of ordered pairs (for example in the same context of the above task, we asked them to imagine that they wanted to draw two balls, one ball for themselves and one for a friend). That means that it was very difficult for them to understand the context of these tasks, as they could not imagine repetitions of the two-stage experiment so as to think that for every trial all the balls could be inside the bag again. In the above tasks they considered that one trial of the experiment was identical with the solution of the problem. We could say that the children could propose a number of pairs according to the number of the balls inside the bag. Then they stopped when they could not make other pairs with the rest of the colors. This tendency was also recorded in a

similar task, where we had put two boxes with two balls in each of them. The children were asked to predict the colors that could come out if they drew a ball from the first box and a ball from the second box. They proposed two different pairs of balls (one ball from each box), thinking that they had exhausted all the colors.

Furthermore, these children did not have a consistent way of thinking about their answer to the question that this task brought out, regarding the probability of an event and the comparison of probabilities. Sometimes they gave a right answer and sometimes a wrong answer. However, in all cases it was very difficult for them to provide an explanation for their answer. In the case when they presented their arguments, they seemed to be influenced by the position of the materials (especially in the experiments with the balls) or their favorite color (especially in the experiments with the spinners), that is they based their judgments on subjective beliefs. In the comparison probability tasks, they usually gave the right answer for the probability situation related to the spinners and the wrong answer for the one related to the boxes with the balls. This difference in their answers shows that the comparison tasks were easier for them when they had to compare sizes rather than numbers of objects.

After the completion of a trial in an experiment, they did not usually mention the color that they had picked, but they gave all the other outcomes at the next trial.

We could remark, at this point, that according to the theoretical framework of Jones et al. (1997), all the children attained the first level for the probability of an event, the probability comparisons and the conditional probability. However, in relation to the sample space, except for the children in the first category, their problems were with the two-stage experiments, as they could find the possible outcomes in the one -stage experiments.

5.2. DURING THE TEACHING EXPERIMENT

We could identify three critical steps in the development of the children's probabilistic thinking during the classroom teaching experiment. These steps will be presented through representative episodes that took place in the classroom. The common characteristic in all these incidents was that the children found a way to resolve their problems that gave them the chance to *overcome their subjective interpretations on the probabilistic tasks*. These solutions influenced the development of the children's probabilistic thinking during the program, with the help of the teacher who always referred to them when the same problems appeared again.

These steps are described in the following way: a) the interpretation of the "different" outcomes in a two stage experiment; b) the acceptance of the realization of the experiment for resolving conflicting viewpoints; and c) estimating the outcomes in a problem.

The interpretation of the different outcomes in a two-stage experiment

In order to introduce the children to the idea of searching for all the possible outcomes in a two-stage experiment, we asked them to participate in the *Christmas cards* activity that was described earlier (section 4). This activity seemed to be very fruitful, as it gave them the opportunity to discuss what it means to have different outcomes in a two-stage experiment with ordered pairs. We could identify two important moments in the discussion during the engagement of the children with this task.

The first instance took place when it was Thomas's turn to pick the two colors. After his picking the first color (it was the red one) and replacing it, he picked the second one and it was also red. Thomas told the teacher with surprise that this could not happen, because "we have to choose two different colors". As the teacher tried to pose Thomas' s thoughts as a topic of discussion with the rest of the class, many children agreed with Thomas's thought. However, some children argued that they had to accept this outcome making comments like "But this color came out". This argument showed that some of the children had already begun to accept the randomness of the outcomes in an experiment. Although, this acceptance was also a consequence of the previous activities concerning the one-stage experiments, the interesting point at this moment was that this acceptance of the

randomness of the colors that could come out during the realization of an experiment allowed them to accept as a “normal pair” the pair with the same color twice. This way, they managed to overcome this difficulty.

When the teacher asked them to find the cards that were same, the children began to put together the cards having the same colors. As they began to put together the cards with the red flower on the left and the green flower on the right, on the next card, the flower on the left was green and the other one was red. The following discussion between the teacher and the children took place in the classroom.

Teacher: Where will you put this card?

Thomas: Here. (*Together with the other cards*)

Anna: No, this is different. The green goes first and then the red.

Socrates: This is wrong. They are not different. They are the same.

Thomas: It's a *little* different!

Socrates: Yes, it's a *little* different.

Teacher: OK. They are not the same, however they are a little different, so can we put this card alone, not with the others?

Thomas: Yes, here. (*He is showing a place near the other cards*).

Teacher: Can we make other different cards?

Anna: Yellow, red.

Teacher: OK. We have the red-yellow card and we can make the yellow-red one.

Marina: Two green.

Teacher: Very nice.

Thomas: Two blue.

Teacher: Bravo, Thomas.

On the above episode, as the children tried to negotiate their different interpretations about the word “different” in the concrete context, they managed to resolve their conflict and construct an acceptable characterization for the difference of the cards. The acceptable interpretation of the word “different” gave them the possibility to find all the different ordered pairs in subsequent similar activities. For example, in another activity, the children had to construct different Christmas trees with two balls by the same manner. When the children checked if they had made all the different trees, they used the same words (“little different”) to describe their results.

Two days later, the children had to find all the different combinations with two colors for a curtain, when they had to choose among four colors. The children found some of the combinations with the process of the experiment. When the teacher asked them if they had found all the different ways for the colors of the curtain, Marina said that there was one more as “we have red and green, red and blue, we don't have red and yellow”. The way that she managed to justify her answer showed that she had made a first step towards the generation of a strategy for the listing of the possible outcomes in a two-stage experiment. This does not mean that the children managed to construct a systematic strategy to find all the possible outcomes in a two-stage experiment. However, it was a major advance for them to develop such arguments concerning these probabilistic situations.

The acceptance of the realization of the experiment for resolving their conflicting viewpoints

A second critical step was the acceptance of the realization of the experiment for resolving conflicting viewpoints. At the beginning, there were children that insisted on their solutions even though they knew the results of the experiment. When they legitimated the way to resolve their disagreement by respecting the results of the experiment and began to change their opinion according to them, they were based more often on quantitative judgments, by using the words “more” or “less” and sometimes numbers to justify their answers. The children seemed to recognize that there was a solution to the problem independent of their desire and they were trying to find a logical explanation for their answer.

Moreover, one type of argument that some children used to justify their opinion in the probabilistic situations was related to the process of executing the experiment. This was expressed as follows: “we close our eyes and we pick a ball”, or “we mixed the balls” and it seemed to constitute a fruitful contribution to the development of the children’s thinking. We could maintain that this type of argument was a different expression of the acceptance of the notion of randomness during the realization of the experiment. More specifically, these statements served as a counter argument to the solutions that were based on subjective judgments like: “this one, because it is higher up in the bag”, or “it is the blue one, because the two green are below”, which referred to the position of the objects.

The representative episode described below took place when children were engaged in the following activity concerning the probability of an event.

The grandmother and the grandfather of the squirrel would like to make a scarf for him. The grandmother found a bag with balls of wool with different colors (1 blue and 3 green). As the grandmother and the grandfather were in complete disagreement about the color of the scarf, they decided to pick a ball at random. The children were asked to predict the color that was “the easiest” to come out. (The bag was painted on cardboard as in diagram 1. The black ball represents the blue color and the grey balls represent the green color.)

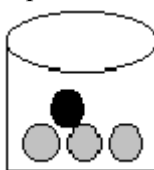


Diagram 1

The following dialogue took place among the members of the classroom.

Anna: The green color, because there are 3 green balls.

Teacher: Is there another opinion?

Socrates: The blue color, because this is higher up in the bag.

Teacher: What do you want to say Paul?

Paul: I say the green ...we close our eyes and we pick a ball.

Teacher: Paul said that we close our eyes when we pick a ball. So, we cannot see where it is. As we are picking a ball, we can mix the balls. (*He is doing the movement.*)

Socrates: It is the blue one, because the three green are below.

Anna: It is the green, because there are more. Let’s do it.

Socrates: Yes!

Teacher: OK! Let’s do it.

Anna, Paul and Socrates were three of the children whose responses in the interview could be classified in the second category. However, in the progress of the program, Anna’s responses were based on quantitative judgments, as she often used the words “more” or “less” to justify her thinking. It was the first time that she compared numbers to justify her answer. On the other hand, Socrates was still influenced by subjective judgments, when he tried to justify his answer. The level of Paul’s thinking about the probability of an event could be characterized as transitional, between subjective and naive quantitative thinking. In this episode, as he tried to justify his answer, he was thinking about the process that the children used to make the experiment and he used an argument which could upset Socrates’ argument. This argument seemed to be a fruitful contribution to the development of Socrates’ thinking. Although Socrates insisted on his answer, when Anna told him that they could find the solution to their problem by executing the experiment, he willingly accepted her idea. This was the first time that the children legitimated a way to resolve their disagreement by respecting the results of the experiment. After the realization of the experiment, Socrates changed his opinion. This episode served as a catalyst for the following lessons, as the teacher used it as a point of reference when the children expressed their different opinions about the solution of a problem. In this sense, this was a critical moment in the development of children’s probabilistic thinking.

Estimating the outcomes in a problem

A third critical step was attained when the children could give the solution in a probabilistic situation by estimating the outcomes in a problem without the need of the realization of the experiment. This step arose from children's attempts to pose their own problems about the probability of an event, when they had to discuss probability situations with equally likely events.

The following episodes illustrate the previous remarks.

Paul: I want to give a notebook to Kostas.

Children: What color does Kostas like?

Paul: The red. (*He is putting in the box, 2 red, 2 green and 2 blue notebooks*).

Teacher: Which color is the easiest to come out?

Anna: The red.

Paul: No. It's equally easy to get the red and the green and the blue.

Teacher: (*To the class*) What do you say about Paul's opinion?

Thanasis: Yes. We have two red, two blue and two green notebooks.

Anna: Paul, put more red notebooks in the box!

Although the teacher had not talked with the children about probability situations with equally likely events until this episode, this problem arose from the children's attempts to make their own problems. In this way they provided us with the opportunity to discuss similar activities with them, like fair and unfair games.

The episode described below is representative of the dialogues that evolved later in the classroom. The children were engaged in the following activity.

The grandmother and the grandfather of the squirrel would like to make a jumper for him. The grandmother found a bag that contained balls of wool with different colors (1 blue, 2 white and 3 red). As the grandmother and the grandfather were in complete disagreement about the color of the jumper, they decided to pick a ball at random. The children were asked to answer which color was "the easiest to come out". The bag was painted on cardboard as in diagram 2. The black ball represents the blue color and the grey balls represent the red color.)



Diagram 2

Andreas: The red color, because there are three red balls.

Teacher: Basil, what do you want to say?

Basil: The blue color, because it is higher.

Andreas: If we mix the balls, we do not know which is higher up in the bag.

Teacher: Andreas said that we can mix the balls and then we cannot know which is higher up in the bag, the red, the white or the blue. What do you say now, Basil?

Basil: Can we do it? (*He means the experiment*).

The teacher asked the opinion of the other children and then they realized the experiment. After having reached the conclusion that the red color was the solution of the problem, the teacher asked the children the following question.

Teacher: If we would like all the colors to be "easy" to come out, the red and the blue and the white, what do we have to put in the bag?

Basil: 3 blue, 3 white and 3 red.

Teacher: Do you agree with Basil?

Anna: No.

Teacher: What do you believe? What must we do?

Anna: We will put 2 blue, 2 white and 2 red.

Teacher: Is there another opinion?

Thodoris: 4 blue, 4 red and 4 white.

Teacher: Socrates?

Socrates: 1 blue, 1 red and 1 white.

Teacher: OK. Finally, which of you is right?

Thodoris: They are all the same.

Teacher: Very good. All of you are right. If we would like all the colors to be “easy” to get out, the red and the blue and the white, then we have to put an equal number of red, blue and white balls in the bag.

5.3. AFTER THE TEACHING EXPERIMENT

The children were interviewed at the end of the teaching experiment and one month later. After both interviews, the results were the same, as described below.

In relation to the sample space, all the children could find the outcomes in a one-stage experiment and the majority (8 children out of 15) could report all the outcomes in a two-stage experiment, although without a systematic strategy.

In probability situations concerning the probability of an event and the probability comparisons, apart from two children, the rest of them used quantitative arguments to justify their answers. Eight children used both the words “more” or “less” and numbers. Five children used exclusively numbers for their justification.

Following one trial of a one stage experiment, all the children could give the right answer, that is a complete list of outcomes without being influenced by the outcome of the first trial of the experiment. Moreover, they recognized the changes that happened in an event (that is, if it was easier or more difficult to appear), according to the results of previous trials, in an experiment without replacement

So, we could infer that all the children had made real progress in their probabilistic thinking as a result of the teaching experiment. They had all developed a naïve quantitative thinking responding to probability tasks and they had approached the second level according to the cognitive model that Jones and his colleagues (1997) have constructed.

6. CONCLUSIONS

The results of the research presented in this document exhibit that kindergarten children are able to make considerable progress in their probabilistic thinking when they are involved in simple probabilistic tasks. They can overcome their subjective interpretations and develop a primitive quantitative thinking. The classroom teaching experiment reveals many critical steps in children’s development of probabilistic thinking. The steps that the children made in order to proceed from a subjective towards a naïve quantitative level of thinking are connected with the discussion of the notion of different outcomes in a two-stage experiment, the acceptance of the realization of the experiment for the solution of a probabilistic problem, and the estimation of the data in a probabilistic situation with equally likely events. It should also be noted that the acceptance of the notion of randomness, expressed in different ways from the children, in their various activities, constituted a hidden driving force in the children’s development of probabilistic thinking. However, more evidence from different cultural kindergarten classrooms’ settings is needed to investigate the critical steps that enhance the development of children’s probabilistic thinking at this age.

Moreover, we should mention the fruitfulness of the sequence of the activities that were used in the classroom teaching experiment. This manner of the organization of the instructional activities seemed to offer a lot of learning opportunities to the children. Thus, we could infer that this is a possible path for the incorporation of probabilistic activities in the kindergarten school.

ACKNOWLEDGEMENTS

The author wishes to thank Flavia Jolliffe and anonymous referees for helpful comments in preparing this paper.

REFERENCES

- Acredolo, C., O'Connor, J., Banks, L., & Horibin, K. (1989). Children's ability to make probability estimates: Skills revealed through application of Anderson's functional measurement methodology, *Child development*, 60, 933-945.
- Cobb, P., & Bauersfeld, H. (1995). *The emergence of mathematical meaning: interaction in classroom cultures*. Hillsdale, NJ: LEA.
- Cobb, P., Wood, T., Yackel, E., & McNeal, B. (1992). Characteristics of classroom mathematics traditions: An interactional analysis, *American Educational Research Journal*, 29(3), 573-604.
- English, L. D. (1993). Children's strategies for solving two and three stage combinatorial problems. *Journal for Research in Mathematics Education*, 24(3), 255-273.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M.C. Wittrock (Ed.), *Handbook of research on teaching*, 3rd ed. (pp. 119-161). New York: Macmillan.
- Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic situations? *Educational Studies in Mathematics*, 15, 1-24.
- Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic intuitively based misconceptions. *Journal for Research in Mathematics Education*, 28(1), 97-105.
- Greer, B. (2001). Understanding probabilistic thinking: the legacy of Efraim Fischbein, *Educational Studies in Mathematics*, 45, 15-33.
- Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1997). A framework for assessing and nurturing young children's thinking in probability, *Educational Studies in Mathematics*, 32, 101-125.
- Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A.T. (1999). Students' probabilistic thinking in instruction. *Journal for Research in Mathematics Education*, 30(5), 487-519.
- Kapadia, R., & Borovcnik, M. (1991). *Chance Encounters: Probability in Education*. Dordrecht, The Netherlands: Kluwer.
- National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA.
- Piaget, J., & Inhelder, B. (1975). *The Origin of the Idea of Chance in Children*. London: Routledge & Kegan Paul.
- Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 465-494). New York: Macmillan.
- Schroeder, T. L. (1988). Elementary school children's use of strategy in playing microcomputer probability games. In R. Davinson & J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics*, (pp. 51-56). Canada: University of Victoria.

SONIA KAFOUSSI

Department of Sciences in Pre-school Education
and Educational Design
University of Aegean
1, Demokratias Street
85100, Rhodes
Greece

UNDERSTANDING STUDENTS' EXPERIENCES OF STATISTICS IN A SERVICE COURSE

SUE GORDON

Mathematics Learning Centre, The University of Sydney
sueg@mail.usyd.edu.au

SUMMARY

In this paper we explore issues surrounding university students' experiences of statistics drawing on data related to learning statistics as a compulsory component of psychology. Over 250 students completed a written survey which included questions on their attitudes to learning statistics and their conceptions of statistics. Results indicated that most students were studying statistics unwillingly. A minority of students acknowledged that statistics was necessary for psychology, but statistics was seen by many as boring or difficult. Students' conceptions of statistics were analysed from a perspective developed from phenomenography (Marton & Booth, 1997). The aim of phenomenographic research is to describe the qualitative variation in the ways people experience or conceptualise a phenomenon — in this case students' interpretations of the topic statistics. The conceptions fell into five categories of description including: statistics as decontextualised processes and algorithms, statistics as a tool for professional life and statistics as a way to self-development and enhanced perspectives on our world. Excerpts from interviews with selected students indicate the diversity of experiences in learning statistics. The perceptions of two teachers flesh out the learning and teaching environment. The findings raise challenges for supporting the learning of "occasional users" (Nicholls, 2001) of statistics in higher education.

Keywords: Statistics education research; Service course; Affect, Conceptions; Phenomenography; Activity theory

1. INTRODUCTION AND BACKGROUND

This paper brings to the foreground important, and often invisible, issues surrounding statistics service courses. These are highlighted through a case study on university students' learning of statistics as a compulsory component of psychology. We examine empirical data on the students' feelings about learning statistics and the students' interpretations of the topic, statistics, they were studying. The data alert us to look beyond the statistics curriculum to the broader context of learning, including affective dimensions and students' interpretations of the subject, statistics.

In recent years there has been a much needed and well-focused series of publications researching the teaching of undergraduate or introductory statistics courses and making recommendations for the reform of these courses (Hogg, 1999; Garfield, Hogg, Schau, & Whittinghill, 2002; Moore 2001). Hogg (1999) suggests using continuous quality improvement to improve undergraduate programs and statistics courses. Garfield et al. (2002) propose that desired outcomes of introductory statistics courses include not only statistical learning and understanding, but also students' willingness to persist in their learning and application of skills, and positive attitudes and beliefs about statistics. Moore (2001) discusses the need for realism and inter-disciplinary cooperation in formulating statistics programs for undergraduates.

In the literature on reform of statistics courses, research on affective reactions and attitudes towards statistics has been, for the most part, neglected. Gal & Ginsburg (1994) found that "the body of research on students' attitudes, beliefs, and affect related directly to statistics education is very

small and problematic”. Further, the data relied on paper-and-pencil, Likert-type scales, in which items included statements using the word ‘statistics’. Hence understanding how students interpret the word ‘statistics’ is critical to interpreting the results of these attitude surveys. Gal & Ginsburg (1994) conclude that further attention by both statistics educators and researchers should be focused on beliefs about statistics, attitudes to learning it, and expectations that students bring into statistics classrooms or develop while studying statistics. These are challenges and opportunities for statistics education today.

The primary focus of the literature in undergraduate statistics education is on introductory statistics courses offered by mathematics or statistics departments. Published data on service courses in statistics, taught by other departments, are harder to access. Indeed Garfield et al. (2002) comment that teachers of statistics in the biological or social sciences are harder to reach and support. The literature provides valuable examples of course design, teaching approaches, curriculum outlines and assessment strategies in statistics courses in diverse areas such as chemistry and business (Zetterqvist, 1997; Yilmaz, 1996). These focus on ‘real world’ applications of statistics for non-specialists. However there are few data on affect and its impact on learning statistics as a service course.

The challenges facing statistics education at the tertiary level are particularly poignant in service courses for “occasional users” (Nicholls, 2001), such as for social sciences. Issues include: the training of and background of instructors; communication within and between departments; cramming of curriculum into decreasing time slots, as competition for student time increases; and support of statistics teachers in university departments where colleagues may include few, if any, statisticians. These issues also impact on curriculum reform in service courses in statistics. For example, the well documented neglect of variation in statistics courses is exacerbated in service courses (see Meletiou, 2002, for a review of the literature on variation and its neglect). Further, the availability of powerful and easy to use statistical software can enable “occasional users” of statistics to “undertake technically difficult analyses without the formal theoretical background to fully understand how the results are being derived” (Nicholls, 2001, p.13). Green (2002) notes that conference participation by instructors of statistics is essential, not only so that statistics educators find out what is going on elsewhere, but also for them to make contact with others in the field of statistics education. One concern for professional bodies in statistics education is how to encourage teachers of statistics in service courses to participate in statistics education conferences or sessions in statistics education — when their research interests may lie in other disciplines or fields — and there may not be institutional support for participation in statistics education conferences.

The challenges facing statistics education in universities and other tertiary institutions are embedded in the challenges to higher education in general. A recent review of higher education in Australia identifies factors transforming the context of higher education (Department of Education, Science and Training, 2002). These factors include a shift from an elite education system to a mass education system with increasing diversity of students, changing patterns of enrolment, such as increased part time study, impact of information and communication technologies, and expectations of graduates held by employers. Ongoing transformations are altering the focus and very nature of higher education. According to the Department of Education, Science and Training (2002, p. 2):

There is an ongoing debate over the ‘what’ of higher education. This is exemplified by the tension between whether higher education should provide students with a liberal, general education, with a focus on learning for its own sake, or a professionally focused, specialist preparation for the workforce.

Moreover, in the Australian higher education system, there is an increased dependence of universities on private funding, decreased resources, more emphasis on accountability, and substantial changes in the ratio between students and teaching staff - a 38.6% rise in this ratio the last eight years (Kent, 2002).

Systemic changes such as these are transforming higher education in many countries and affect statistics courses at universities. They translate into practical terms, defining the resources available to university teachers and impacting on the ways teaching statistics is organised or constrained. For example, a university teacher of statistics interviewed for this study (section 4.6) perceived that multiple-choice examinations are an undesirable method of assessment but “essential for financial

reasons". Engeström (1999) pointed out that if practitioners are able to identify and analyse contradictions and tensions in the system surrounding and shaping their activities they may be able to focus on resolving the contradictions and reorganising the activity, "instead of being victimised by changes that roll over them as if forces of a natural catastrophe" (Engeström, 1999 p. 68). To do this the tensions must be made visible - as when a practice is shared and common it becomes invisible.

The aim of this paper is to provide a basis for reflection on issues and challenges that could arise in service courses of statistics. The paper is organised in layers. The outer layer is the broader environment for statistics education at university - the political, philosophical and pragmatic issues that shape the experiences of teachers and students in statistics service courses. These were briefly outlined above and will be contextualised in an empirical study. The next section describes the theoretical and methodological lenses through which I interpret a case study on students' orientations to learning statistics as a unit of study of psychology. Activity theory provides a framework for linking students' orientations to learning statistics to the learning environment surrounding and shaping these orientations. In tandem with the relational view of an activity theory framework, I use a phenomenographic approach to exploring students' interpretations of statistics in the context of a psychology course. The core of the paper is an empirical study conducted at a major, metropolitan, Australian university. I report on psychology students' written responses to survey questions on their feelings about learning statistics and how they interpret the subject statistics. Excerpts from interviews with students are reported to supplement the written survey responses and interviews with two teachers of psychology statistics add to the context for the students' responses and flesh out the teaching and learning issues outlined in this introductory section. The data reported in this paper are aspects of my PhD thesis (Gordon, 1998).

The position of the researcher is important in any investigation. My interest in psychology students learning statistics arises from teaching at a Mathematics Learning Centre in a major Australian university. The Centre provides ongoing support for students of my university who seek assistance with basic level mathematics and statistics courses. I take responsibility for and joy in teaching the psychology students at the Centre who are studying statistics. As a teacher in the Mathematics Learning Centre I am an outsider to the psychology statistics course in that I have no input into the curriculum, teaching strategies or assessment of this course. My practice involves supporting the statistics learning of psychology students at the Centre and researching their learning (Gordon, 1995). This study aims to give the students a voice — how they felt about learning statistics and what their conceptions or interpretations of statistics were. Hence this paper is an attempt to balance personal experiences teaching statistics in the Centre and researching students' learning with issues in the broader context of statistics education.

2. THEORETICAL FRAMEWORK

Activity theory, based on the powerful ideas of Vygotsky and Leont'ev (for example, Leont'ev, 1981; Vygotsky, 1978), can be an effective tool for understanding learning. This theory explains that people actively develop knowledge on the basis of life experiences that are rooted in the ongoing practical and communal life by which societies organise and reorganise themselves. The activity of learning is a process in which people grapple with new information - to make it meaningful, to solve problems and to adapt to new conditions.

The importance of the theoretical framework to this investigation is twofold. Firstly, it centres our attention on students' actions, including mental actions, and their goals - what students do and why they so act. Secondly, it brings the context of learning to the foreground. When students learn statistics they interpret it within their own frameworks. The ways students learn statistics reflect their personal stories, their experiences, needs and goals. Moreover, students' activities are organised within and shaped by the histories and cultures of their institutions, as well as the social factors prevailing in their statistics classes. It is by looking at the context in which students learn statistics that we can make sense of their actions. How students think and act depends on how they position themselves - what the students believe to be strategic in the circumstances.

Activity theory emphasises the developmental nature of human actions and thinking. Activities progress and transform over time. Varela, Thompson and Rosch (1991, p. 205) describe cognitive capacities as: “paths that exist only as they are laid down in walking”. Their metaphor beautifully illustrates the inseparability of personal meaning and setting - the co-emergence of thought and action - knowing as doing. An implication of the developmental approach is to view students’ interpretations of statistics as having developed through their experiences and schooling over a long period. For example, experiences of learning mathematics at school may impact on learning statistics at university. This development is also ongoing. The results of this study are a snapshot - we can only capture a moment in an ongoing progression in the students’ thinking.

In sharp contrast to psychological theories in which cognitive processes are viewed as separate from the emotional domain, activity theory emphasises the inseparability of students’ feelings from their thinking. Vygotsky (1962, p. 150) describes this inseparability of intellectual processes and affective elements as follows:

Thought itself is engendered by motivation, i.e. by our desires and needs, our interests and emotions. Behind every thought there is an affective-volitional tendency, which holds the answer to the last ‘why’ in the analysis of thinking. A true and full understanding of another’s thought is possible only when we understand its affective-volitional basis.

Affective elements operate as invisible dimensions of student learning, neither acknowledged nor valued in statistics education. Lerman (1996) argues that the valuing of decontextualised, intellectual processes, divorced from personal elements, is expressive of oppressive discourse. It is this privileging of abstract thought, such as academic mathematics, that is disempowering for some students. Sierpinska & Lerman (1996) refer to the vested interest mathematicians have in maintaining the status of mathematics in society. This idea was passionately expressed by a statistics student. This student identified statistics with mathematics. She wrote in her survey (Gordon, 1998, p.18):

Maths is an exercise in agony, because the people who teach it make one feel as though maths belongs in a higher plane of evolution. Even though the number system is for everyone, and the concepts are there for everyone, the feeling (especially if you are doing pass options) that you do not deserve to know anything runs rampant.

Leont’ev (1978) makes a useful distinction between meaning as ‘cultural’ or ‘personal’. Cultural meanings are connected with the reality of the outside world, the life of society. Personal sense, on the other hand, is connected with the reality of the person’s own life, motives and goals — what is significant for the individual. Society endorses statistical information as scientific and useful but to some students statistics is neither relevant nor important. Hence sometimes there is a mismatch between societal and personal meaning. It is affect that determines the fit — how cultural meanings are translated into personal sense. What is culturally endorsed, for example, achieving high grades in mathematics, may not translate into personal motivation if the student hates mathematics or is anxious and apprehensive about learning it.

2.1 BACKGROUND TO PHENOMENOGRAPHY AS A RESEARCH TOOL

The approach taken in this study to analyse students’ conceptions of statistics is based on phenomenography. Phenomenography is a well-established, qualitative, research tool in education research that seeks to uncover participants’ own conceptions or experiences of a phenomenon in varied contexts (Marton, 1986; Marton & Booth, 1997). The results of phenomenographic analysis are categories of description that capture the essence of each conception. That is, the categories reflect critical differences in the participants’ expressions of their conceptions. For example, my colleagues and I found phenomenographic methods useful to explore the qualitatively different ways in which students experience learning mathematics (Crawford, Gordon, Nicholas & Prosser, 1994). In that study conceptions of mathematics were found to divide broadly into categories expressing fragmented conceptions of mathematics, such as mathematics being about numbers, rules and formulae, and categories expressing cohesive conceptions, where mathematics was seen as being about a complex logical system providing insights about the world. Reid and Petocz (2002) investigated the conceptions of statistics held by students enrolled for degrees in mathematical sciences. They

described six qualitatively different conceptions of statistics ranging from fragmented to inclusive views. The six conceptions were grouped in terms of the focus of the conception, namely a focus on techniques, on data or on meaning. In this paper the categories describe the qualitative variation in students' conceptions or interpretations of the topic statistics, as indicated by their responses to questions which were part of a written survey.

The set of categories that is the outcome of a phenomenographic investigation represents an awareness of the phenomenon which is a collective rather than individual. Phenomenographic analysis starts with accounts of individual experiences from which categories of description are developed. In developing these collective categories individual idiosyncrasies are ignored. Individuals are seen as "bearers of fragments" (Marton & Booth, 1997, p.114) of each conception category so that any one person's expression of a conception may not fully sum up the category. A strength of phenomenographic analysis is that the categories of description are developed from the data and are not preconceived. The researchers try to step back from their own perceptions or "bracket them" (Ashworth & Lucas, 2000) and attempt to see the phenomenon, in this instance the experience of learning statistics, through the eyes of the students.

A phenomenographic perspective is relational in the sense that a person's experience or awareness of the phenomenon constitutes an internal relationship between the person and the phenomenon (Marton, 1988). Svensson (1994, p. 12) argues that underpinning the phenomenographic "specialisation" is an assumption that knowledge is fundamentally a "question of meaning in a social and cultural context". Marton & Booth (1997) emphasise that conceptions or experiences are neither psychological nor physical but are descriptions of the internal relationship between persons and phenomena; "an experience is of its essence nondualistic" (Marton & Booth, 1997, p. 122). Hence phenomenography emphasises the relatedness of individuals, the context and the subject matter, statistics.

Any research tool develops as it is applied and in this paper the phenomenographic approach is developed in accordance with the activity theory framing the investigation. Students' conceptions of statistics are interpreted as understandings of statistics experienced by the students while studying statistics. That is, conceptions, rather than being purely mental images, in the head, are accompanied by and inseparable from emotions, goals and actions taken to study statistics.

3. METHODOLOGY

In this paper we explore students' written responses to part of a survey. These data are supplemented by excerpts from audio-taped interviews with selected students and two university teachers. The participating students were studying statistics as a compulsory component of second year psychology at a major and traditional metropolitan university in Australia. The survey forms were handed out to 279 students and completed in writing by these students during 20 minutes of a statistics lecture, near the end of the first semester, with the co-operation of the lecturer concerned. The theoretical position described above suggests that learning will be shaped by students' perceptions and the survey was designed to elicit students' own views and perceptions on learning statistics. The survey was completed by all students attending the statistics lecture on the given day and it is a matter of conjecture whether the students present at the statistics lecture differed from the absentees concerning the study of statistics. Hence the sample may not be representative of the cohort of second year psychology students.

The survey included open-ended questions on students' willingness to study statistics and their interpretations of the subject, statistics, which will be analysed in this paper, as well as a questionnaire on approaches to learning statistics (discussed elsewhere, see Gordon, 1999). Later in the year seven students were selected for interviews based on the ongoing analysis and categorisation of survey responses. These students were selected because their diverse survey responses typified, as far as possible, the different analytic categories that had emerged by that time. I also ensured a range of values on variables such as age, gender and background in mathematics for the students interviewed. Two teachers of the psychology statistics course were interviewed to shed some light on the course from the teaching perspective.

The open-ended survey questions are shown below. In what follows we focus mainly on students' responses to Question 1 and Question 3, although responses to Question 2 were taken into account in the analysis as will be explained in section 3.1.

Question 1. Would you study statistics if it were not a requirement of your psychology course?

Please give reasons for your answer.

Question 2. Think about the statistics you've done so far this year.

a) How do you go about learning it?

b) What are you trying to achieve?

Question 3. What in your opinion is this statistics course about?

Please explain as fully as possible.

Analyses of students' responses to Question 1 above are presented in sections 4.1 and 4.2. Section 4.3 describes students' performances in assessments. The categories of conception that emerged from the phenomenographic analysis are discussed in section 4.4 with quoted excerpts from the surveys that illustrate the categories as best possible. Interview data from students, which flesh out the written survey responses, are reported in section 4.5 and excerpts from interviews with two teachers of the course are provided in section 4.6.

Student learning is set within a particular context in space and time. At the time of data collection and analysis the statistics course consisted of one lecture and one tutorial per week, each an hour long, and the course continued for two semesters (that is, one academic year). The course was taught in the traditional lecture-tutorial format with assessments being primarily open-book quizzes, assignments and examinations. Students did not have access to computer workshops but were presented with SPSS outputs for analysis and interpretation.

Diversity is a hallmark of student groups at university and the participants of the study were diverse in educational background and experience, expertise in mathematics, age and other demographic variables and course of study undertaken. The participants were enrolled for various degree courses with the majority pursuing Bachelor of Arts degrees (183 students, 66%). The bulk of the rest were studying degrees in Science (48 students, 17%), Economics (19 students, 7%) or Education (15 students, 5%). Most of the students (184 students, 66%) were either 19 years old (133 students, 48%) or 20 years old (51 students, 18%) though a sizeable minority (46 students, 16.5%) was at least 25 years in age. About three quarters of the students were female. The prior level of mathematics studied by the surveyed students was higher than is often assumed for psychology students. Fifty eight percent (163 students) had studied mathematics at school to a level which included Calculus and a further 68 students, almost one quarter of those surveyed, had studied mathematics at university for at least one year.

All the participants had studied a short, five-week, introductory course in statistics during their first year study of psychology. Interestingly, most of the participants (57%) expressed an intention to continue studying psychology at the postgraduate level, suggesting that many were committed to studying psychology for professional reasons - to practise as psychologists. Only 7% (19) of the participants reported that they intended terminating their study of psychology after completing second year, to pursue other subjects.

3.1. ANALYSIS OF OPEN-ENDED QUESTIONS

Students' responses to the first survey question listed in the previous section were coded as "Yes" or "No" and then classified according to students' reported reasons for willingness or reluctance to study statistics.

Students' conceptions or experiences (Marton & Booth, 1997) of statistics were analysed using an approach developed from phenomenography, as outlined previously in section 2.1. The focus of this analysis was on responses to the third open-ended question. However, students' responses to Questions 1 and 2 were examined for clarification on their responses to Question 3. Consistent with the activity theory lens, the aim of the categorisation was to capture not only what the students

thought about statistics but how their mental interpretations and perceptions were expressed in action - what the students did to learn statistics and what the goals of these actions were. So students' overall responses to the three open-ended survey questions were taken into account in order to gain as comprehensive an understanding of the students' awareness of statistics as was feasible.

The method of categorising written survey responses, used in this study, was based on previous phenomenographic research conducted by colleagues and myself into students' conceptions of mathematics (Crawford et al., 1994). As for that study, the method used here involved attending to the similarities and differences in students' statements and selecting the themes or ideas that differentiated categories of conception. This involved cycles of analysis and review. In the first stages of analysis an independent researcher and I identified an initial set of categories. This was achieved by each of us independently reading and re-reading the entire set of 279 written responses and drawing up the critical statements that appeared to us to differentiate categories of description for students' conceptions of statistics. That is, the students' individual understandings or conceptions of statistics were grouped into sets of statements or categories. We then compared and discussed our interpretations and identified what Dahlgren (1984) describes as "core elements which make up the content and structure of a given category" (p. 24). In the next stages, each of us, and a third researcher, independently classified a selection of 30 survey papers in terms of the draft set of categories. The selection included responses which appeared to typify each category or, conversely, to be on the borderline between two categories. The individual classifications of the three researchers were compared, discussed and reviewed. In the ensuing discussions the categories were clarified and refined and a set of clear statements defining each category of conception was agreed upon. The different categories were delineated by the focus of each conception.

Finally, the distribution of students' responses into the categories was recorded. All the survey responses were classified into the categories by myself and one other researcher working independently. These classifications were compared and discussed in further cycles of analysis. The final distribution was a consensus between the two of us.

4. RESULTS

4.1. WILLINGNESS TO LEARN STATISTICS

In this section I report on students' responses to the first survey question:

Would you study statistics if it were not a requirement of your psychology course?

Please give reasons for your answer.

Seventy three percent of the students surveyed reported that they would not have studied statistics, if they had been given a choice. Students gave many reasons for their willingness or reluctance to study statistics. These can be broadly divided into reasons which relate to how learning statistics fits with the student's goals and interests or perceived abilities - the student's "personal" sense of what it means to learn statistics in Leont'ev's (1981) term - and reasons which have to do with the expectations of others, such as teachers or a societal view of the relevance of statistics, culturally endorsed meanings of statistics. The students' reasons as reported are accordingly grouped under the headings Personal Sense and Cultural Meaning in Table 1. Within these broad categories, Table 1 summarises the reasons that emerged most frequently from the surveys (top 6 categories). Short illustrative excerpts from students' scripts are provided to give examples of responses. These are divided into positive responses, favouring the learning of statistics, and those indicating unfavourable or negative appraisals. Many students gave multiple responses and some gave reasons that fell into more than one category, such as below.

... although I see it is necessary for psychology I find it boring and tedious.

This response was scored as positive in the category *Necessary For Psychology* and negative in the category *Interest*.

Table 1 shows, too, the number of students who gave a response in the indicated category. Each frequency is also shown as a percentage of 279 to indicate the proportion of students surveyed who

gave that response. For example, 5 students (2% of them) indicated that that they found statistics interesting, 19 students (7%) reported that statistics was generally useful in life; the categories were not mutually exclusive so, in this example, some students may have written that statistics was interesting and useful.

As can be seen, the most frequently cited reasons fell under the Personal heading and were negative, with 29% of the students reporting statistics as boring or tedious (*Interest*) and 13% of the students indicating that they disliked learning it (*Affect*). The category with the highest positive frequency was *Necessary For Psychology* with 16% of the students giving a positive response in this category. Responses in this category indicated that statistics was integral to psychology as a discipline but did not specify why this was so. Responses referring to a specified usefulness of statistics, in society or for a career, were categorised separately. Interestingly students' responses in the area of Personal Sense often referred to the students' attitudes to and experiences of learning mathematics, while responses concerned with usefulness to society or necessity for psychology (Cultural Meaning) tended to be specifically about statistics. These responses highlight the tension between statistical knowledge as culturally endorsed and the individual's personal appraisal of learning statistics.

Table 1. Reasons for willingness or reluctance to study statistics

PERSONAL SENSE		
Category	Favourable Responses	Unfavourable Responses
INTEREST	5 responses (2%)	80 responses (29%)
Excerpts	<i>It's interesting</i>	<i>cause I generally find it dull, boring and tedious</i>
AFFECT	13 responses (5%)	37 responses (13%)
Excerpts	<i>I sort of dig numbers</i>	<i>I dislike maths intensely</i>
PERSONAL RELEVANCE	0 responses	21 responses (8%)
Excerpts	—	<i>No, it is irrelevant to my life</i>
CONFIDENCE	0 responses	20 responses (7%)
Excerpts	—	<i>Maths of any sort immediately makes me cringe</i>
CULTURAL MEANING		
Category	Favourable Responses	Unfavourable Responses
NECESSARY FOR PSYCHOLOGY	46 responses (16%)	8 responses (3%)
Excerpts	<i>It is necessary in the study of psychology</i>	<i>I am not sure what relationship there is between statistics and psychology</i>
GENERAL USEFULNESS	19 responses (7%)	1 response (0.4%)
Excerpts	<i>statistics are used throughout our society i.e. newspaper reports & it is important to have an understanding of the way in which information is gathered, processed & manipulated.</i>	<i>No, it is not a practical subject</i>

4.2. DISTRIBUTION OF RESPONSES ACCORDING TO WILLINGNESS OR RELUCTANCE TO STUDY STATISTICS

Different reasons were given by students who responded that they would have studied statistics, even if it had not been compulsory to do so and those who stated that they would not have studied statistics, given a choice. Students reporting willingness to study statistics will be referred to as “Yes” students in what follows and those who would not have studied statistics, given a choice, will be termed “No” students. The most dramatic differences between these two groups can be seen in the bar graphs shown in Figure 1.

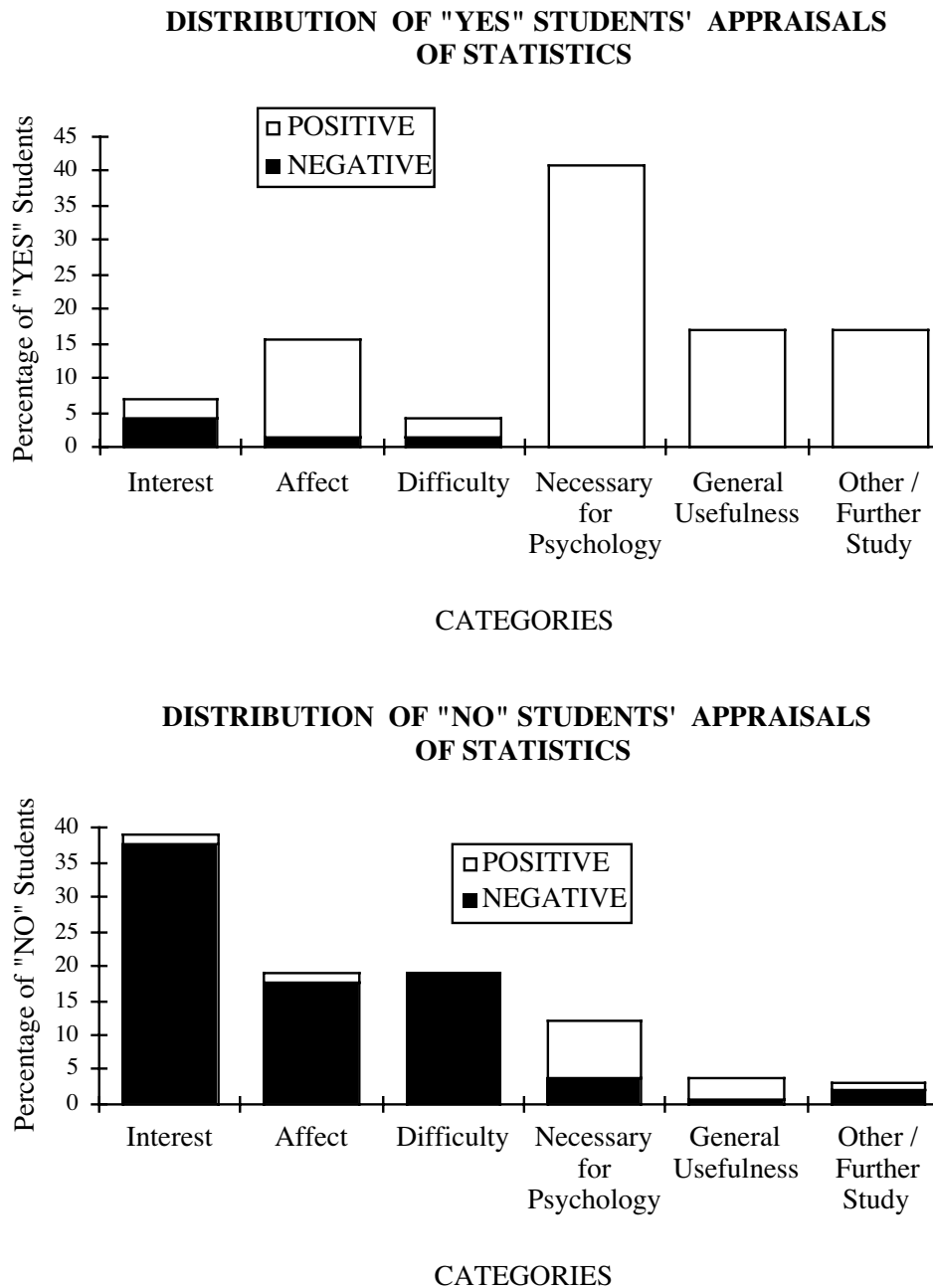


Figure 1. Reasons of “Yes” students and “No” students for studying statistics

Figure 1 compares the frequency of responses in the three most common categories for willing (“Yes”) students, namely *Necessary For Psychology*, *General Usefulness* and *Other or Further Study*, with those in the three highest frequency categories for reluctant (“No”) students, *Interest*, *Affect* and *Difficulty*. The comparison shows that, rather than reflecting opposite sides of the same coin, the reasons given by students willing to study statistics related primarily to the relevance or usefulness of statistics, while the reasons given by the reluctant majority related mainly to feelings and emotions - the “personal plane” (Leont’ev, 1978).

One interpretation of these findings is that the students who expressed positive reasons for studying statistics had “internalised” (Leont’ev, 1981) the institutionally endorsed reasons for doing so, for example the necessity of statistics to psychology as a discipline. Another possibility is that the willing students took a more mature or realistic view of the relevance of statistics to their professions or lives. However few students reported that statistics had “personal sense” (Leont’ev, 1978) for their long-term goals or self-fulfillment.

4.3. PERFORMANCE IN ASSESSMENTS

The final mark for statistics was the average of the following four components: the class mark, based on “open-book” tests or quizzes, the examination mark for semester one, and, similarly, the class mark and examination mark for semester two.

The students who expressed willingness to study statistics attained a higher average mark on the statistics tests and examinations than the reluctant learners. The mean final mark in statistics for the “Yes” students was 66% compared to 55% attained by the “No” students. This difference was not reflected in the overall performances of these two groups in psychology, excluding the statistics component (mean for “Yes students: 67%, mean for “No” students: 66%). No causal inferences should be drawn from these findings. Research on assessment procedures in statistics and student motivation shows the relationship to be complex and multi-faceted (Garfield, 1994; Hubbard, 1997; O’Connell, 2002).

4.4. CONCEPTIONS OF STATISTICS

Five categories describing students’ conceptions of statistics emerged from the phenomenographic analysis. They were: *No Meaning*; *Processes* or algorithms; *Mastery* of statistical concepts and methods; a *Tool* for getting results in real life and *Critical Thinking*.

The categories describe the qualitative variation in the ways these students viewed statistics, as interpreted by the researchers, and each category summarises the primary focus of the expressed conception. The categories range from the narrowest and most limited interpretations of statistics to the broadest and most inclusive view of statistics. That is, the structure of the categories is a hierarchy from what Reid and Petocz (2002) call “limiting to expansive views”. This means, for example, that an understanding of statistics as being about *Processes* or mechanical aspects reveals a more limited conception or experience of statistics than a focus on the application of statistics — its use as a *Tool*. Conceptions showing awareness of statistics as a tool to be used in research and professional life could include an awareness of the algorithmic, formula-based aspect of statistics, but is not limited to that aspect.

As indicated earlier, the students’ responses to all three open-ended questions (presented in section 3) were taken into account to flesh out what students had written in response to Question 3. Hence it was hoped that the categories captured the students’ holistic experiences of statistics including their awareness of statistics expressed in thought, emotion and action. If a student’s response indicated more than one of the conception categories it was classified in the ‘highest’ or most expansive category. The categories, the distribution of students’ responses in the categories, and brief excerpts from the surveys illustrating each category, are summarised below. The percentages given are out of 279, the number of students surveyed.

No meaning (4%)

Responses in this category indicated conceptions of statistics as an imposed and irrelevant subject. For example, one response to the question “What in your opinion is this statistics course about?” was:

You tell me, I just learn.

Processes (24%)

This category expressed a focus on mechanical techniques for solving statistical problems or algorithms. In this conception statistical knowledge is viewed as disconnected — for reproduction in assessments. Typical responses were:

... You don't have to understand how it works, just be able to get the right answer.

or ... computers do all the work.

Mastery (33%)

The focus of this category was on reading and understanding material as presented in class; statistics was understood as information to be accumulated and stored in order to meet the demands of assessments. For example, one student wrote:

... I am trying to achieve a basic understanding of the material & concepts & an ability to work out the problems. This will hopefully lead to a good result at the end of the year.

Tool (25%)

This category indicated a focus on statistics as a useful tool. Some students indicated statistics as personally useful, for example:

Using statistics to apply it to experiments we will use later on in careers in psychology. A practical course.

Others suggested that statistics could be applied in real life:

It is an attempt to give psychology students insight into stats & experimental method, to enable them to do psychological research if they choose to follow psychology as a career.

Critical thinking (3%)

This category was the most expansive. The category indicated conceptions of statistics as providing a perspective on the ways in which data is used to make decisions - a scientific way of thinking and communicating. Responses in this category showed insight into the complexity and limitations of statistical theory.

Stats is about methodology which is used as a comprehensive form of analysis to interpret and test theories & correlations psychologists create. Substantiated method.

The majority of students held conceptions of statistics that were fragmented; disconnected from other knowledge and disjoint from their goals and lives. Reid & Petocz (2002) assert that the “importance of developing learning environments that encourage students to use the broader conceptions - to look for meaning in the data and relate this meaning to their own personal situations - cannot be overstated”.

Table 2 shows the percentages of students whose conceptions of statistics were classified in the five categories broken down into “Yes” students (students studying statistics willingly) and “No” students (students indicating reluctance). The percentages shown are column percentages, that is out of 71, 204 and 279, respectively. They are rounded to whole numbers except where the percentage is less than 1%. Each student’s response was categorised into one and only one category. The final column summarises the overall frequency distribution for the participants in the five categories. Table 2 reveals the structural differences between the conceptions or experiences of students reporting that they were studying statistics willingly and those who reported reluctance to take the subject. The majority of the “Yes” students indicated awareness of the relevance of statistics to work as a

psychologist or even scientific thinking (*Tool* and *Critical Thinking* categories). On the other hand almost three quarters of the “No” students expressed interpretations of statistics in the first three categories, mostly indicating statistics as being about classroom learning.

Table 2. Number and percentage of “Yes” students and “No” students in each conception category

Category	“Yes” Students (N= 71)	“No” Students (N=204)	Total (*N=279)
NO MEANING	0	11 (5%)	11 (4%)
PROCESSES	7 (10%)	61 (30%)	68 (24%)
MASTERY	17 (24%)	75 (37%)	93 (33%)
TOOL	35 (49%)	36 (18%)	71 (25%)
CRITICALTHINKING	6 (8%)	1 (0.5%)	7 (3%)
Unclassified conceptions	6 (8%)	20 (10%)	29 (10%)

* Note four students could not be classified as “Yes” or “No”.

4.5. VIGNETTES FROM STUDENT INTERVIEWS

The five categories outlined in section 4.4 describe collective conceptions. That is, the categories have been “stripped” (Marton & Booth, 1997, p. 114) of the individual voices from which they were developed. The following excerpts are from interviews with selected students (reported in Gordon, 1999) and illustrate some of the individual interpretations of statistics captured by the conception categories.

The responses given by Narelle and Ben, below, are both indicative of the *Processes* category. The excerpt from Ruth’s interview illustrates the *Mastery* category, as her focus was on understanding and mastering the concepts for examination purposes. Tilly, in her interview, indicated an awareness of the use the statistics she was learning (*Tool* category), while Tessa’s interview responses illustrate the most expansive conception of statistics (*Critical Thinking* category).

Narelle (Processes)

Narelle indicated that learning statistics was associated with anxiety and a lack of confidence. Although she had studied a one-year, introductory statistics topic for Arts students the previous year at university this had added to her confusion rather than demystifying statistics. She reported her experience of studying statistics as follows:

I usually come home and have a look at what we’ve done in class and hope I understand it. If I don’t then I’ll perhaps contact my tutor, which I’ve only done once, or just keep doing the exercises until I work out what’s going on.

... I want to be able to do it automatically rather than it be such a difficult process. ... I’d find what I need out of the problem, the mean and whatever, and then work through that - rather than read so much into the problem; what’s going on with the rats or whatever.

... I’m trying to get the relevant information from the problem without being distracted by so much to think about, like the whole experiment. ...

I don’t enjoy it particularly. I enjoy it once I get an answer out I guess, but the working out is just not my thing. I’m not a very mathematical person.

Narelle’s comments indicate a focus on the mechanics of problem solving without considering the context of the problems, which fits with a conception of statistics as being about processes. Her negative feelings about learning statistical concepts appeared to be linked to her lack of confidence in mastering mathematics.

Ben (Processes)

Ben, like Narelle expressed a focus on getting the answers without much attention to the meaning of the statistical theory, but, unlike Narelle, he indicated a facility with learning statistical processes. He wrote in his survey that his aim was to “be able to get the right answers”. In his interview he said:

It doesn't so much interest me, but its easy going. You don't have to muck 'round doing reports and stuff. You can just learn it. And it's sort of half relevant to what I'm doing, and that's fine. It doesn't bother me a great deal if I don't understand, say, the exact theory behind different distributions and stuff like that. It doesn't really concern me a great deal. As long as I understand the basics and be able to just get through the questions.

Ruth (Mastery)

In her interview Ruth indicated that statistics was irrelevant to her, observing that “this manipulation seems to be a little bit pointless. We're never going to have to use it again.” However Ruth revealed that she was trying to achieve good grades by understanding the concepts and methods. Hence, Ruth's focus appeared to be on the mastery of ideas and skills she was taught in class.

It's more trying to understand. It's more trying to give an understanding rather than just answers. It's not just the answer on the page. It's understanding how you get to it. ... You can write down and explain it rather than just having to use formulas.

I was hoping that in the exams — that at least I'd have remembered something, that it would stay there. But other than that, it (the knowledge) is not doing anything.

I'd like to understand it for next year. I don't want to have wasted time. To have nothing to show for it. And have no added knowledge or no added systems of doing things. It would be nice to have learned some logical ways — some logical methods. But it would be nice to pass the exams. Just that.

Ruth explained that she seldom practised using the statistics, rather concentrated on “trying to get it to stick in my head”.

Tilly (Tool)

Tilly's focus on statistics was on the application of knowledge. This represents a shift in focus from the processes of statistics and mastery of skills and concepts to an awareness of statistics as being useful in the real world and relevant to professional psychologists. This focus on statistics as a tool was expressed in Tilly's remarks on how she learned statistics.

When I'm trying to summarise my lecture notes I try to integrate all the information I have on each set topic. And understand that and therefore know how to apply it. What we're doing this semester is a lot more real statistics, how you really apply statistics to what you're doing. I think that's how psychologists would go about testing and researching these things. I want to learn it because I'm going to need the knowledge. And also - in a lot of ways it overcomes my frustration I had with maths last year. So it's a triumph, almost, to have overcome that big barrier to statistics that I had built up last year.

Tilly appeared to be trying to make connections and see the big picture rather than learn isolated and decontextualised concepts.

Tessa (Critical Thinking)

Tessa expressed a high level of engagement with learning statistics and insight into statistical knowledge. In addition to indicating an awareness of statistics as a tool, which could be applied in professional life, Tessa's comments demonstrate an understanding of the basis and communication of scientific thinking. An excerpt follows:

It's very interesting - how statistics moulds itself into psychology. Right now I think psychology's trying to become a science because of the emphasis that people put on science being the main area of knowledge. The way statistics moulds itself into psychology kinda gives psychology a

basis. Like - how would you put it - like raw facts that can be analysed scientifically, I guess. In maths at school there wasn't that much application to things in real life. A lot of it was just formula based, following the formulas, plugging in the numbers. With statistics, however, you've got an aim. In a real life situation, in society - statistics is like a tool to analyse whatever happens when you do experiments. ... It's helped me think more logically. Like inductive reasoning. It's a lot like inductive reasoning. You come up with a hypothesis and you have to follow through in order to get an answer. ...

I guess that's one of the main reasons I chose psychology and philosophy. I wanted a broader view of life. Science and maths in high school was more a regurgitation of theories. I just thought that in psychology and philosophy I might be able to contribute some new ideas.

4.6. INTERVIEWS WITH TWO TEACHERS OF PSYCHOLOGY STATISTICS: WENDY AND CATHY

Teachers are arguably the most important influences on student learning and send messages instrumental in shaping the experiences of their students. Teaching is organised in an institutional context that may constrain or encourage preferred ways of teaching. The following excerpts are from two teachers of psychology statistics. The excerpts express what Cathy and Wendy perceived to be the purposes of including statistics in psychology, their methods of teaching and the constraints and institutional features which shaped their teaching.

Cathy saw statistics as a tool for research as well as a way for students to critically assess information. She described the aims of the statistics course in this way.

Cathy: *First of all, in reading the literature, understanding what other people have done in their experiments and secondly to help them if they carry out their own research as well, to understand how we have to make sense of numbers and what we do with statistics. To help us make sense of it.*

Cathy hoped that what students would get out of the statistics course was:

Well, just an understanding of the place of statistical procedures within the overall research approach in psychology. So, not to see it as something completely separate but to see how to use it as an aid to systematic thinking and detecting effects amongst all the noise, that sort of thing.

Wendy related statistics to further study but also drew on her own experiences to explain the place of statistics in students' ongoing lives.

Wendy: *I see it as giving little tastes of what you can do. Building up perhaps to the third year course. The purpose in any course is to lay the ground-work for:*

- A) those who want to go on further in your traditional, academic research oriented stuff and*
- B) for those who don't want to go on further but are basically getting a degree so that they can then move into a large company. To basically lay the framework so that when things do come across their desk they can understand it.*

It sounds trite but I've got friends who have done that - moved into banking or market research or whatever, and are having troubles in terms of colleagues who are not understanding:

'What's an average? What's a distribution?'

Even if the students don't remember how to do a chi-squared test, they'll know what a population is. And what the point is - we're trying to tell whether this is having an effect.

I asked Cathy and Wendy how they tried to achieve these educational aims for statistics in their teaching. Their responses showed that they both regarded concrete examples as of prime importance in helping students understand the concepts. Cathy drew on textbooks for her source of examples while Wendy evidently tried to motivate her students by linking her examples to students' everyday lives.

Cathy: *I suppose it depends which topic we're talking about. In general, I try and do it with examples, instead of doing it in the abstract. I relate it to actual examples of when these problems*

might arise and how we're going to cope with these sorts of numbers. What we're going to do with them - and understand why they're going this way or that way.

Wendy: *What I try to do is give them an overall feeling for what they're going to do. For example with chi-square tests. Not all data is continuous. Sometimes it's: 'Do you like Pepsi, or do you prefer Coke?' And that sort of data is just as interesting. However, you treat it differently because you can't build up a number line. It's just: 'Yes' or 'No', or maybe you've got more categories. Give them an overview like that. Give them an overall spiel about when would you use it. Try and come up with examples that are interesting. For example, to make multiple regression interesting you can try and predict who's going to win the Melbourne cup (horse race). And change the equation: 'This is a new jockey, this is his weight, this variable is: Has the horse won the Caulfield?'. And it's quite amazing then how you'll get to the people at the back of the room. And then move onto question and answer time.*

The above comments of the two teachers indicate that each had a different sense of what it means to learn statistics. This difference emerged more clearly from their reports about how a student's understanding of statistics was assessed. Cathy indicated her dissatisfaction with multiple choice examinations which were a resource issue.

Cathy: *I think our best feedback comes from the tests rather than the exams. The exams are multiple-choice and I don't really think that multiple-choice exams are a wonderful method of assessment. They're just essential for financial reasons. I think it's in the test that you get the most feedback about what they're understanding and what they're not understanding. What they can do with it. Whether they can draw appropriate conclusions. You see a lot of the little things that they're not understanding by looking at how they perform in tests.*

Well, things like in hypothesis testing - just whether they can - given a critical value, an observed value, can they work out what they had to do with it? And then what conclusion they can draw having decided to accept or reject the null hypothesis. That's a funny whole cycle that they can often be confused about. Three different parts. Once you can force them to write out their answers, you can see if they're understanding it. And if they can apply formulae. Which is tedious and will probably become less and less of a necessity as courses become more computerised. So the emphasis will become more on understanding output of the procedure I think.

Cathy's observations seem to show that her emphasis was on assessing skills and procedures, although she acknowledged that the time for such procedural learning was past. Wendy appeared to be more concerned with the future relevance of the concepts being learned and indicated that rote learning was a stepping-stone to conceptual understanding.

Wendy: *I think with stats, a lot of it is rote learning and it is very dull - until you actually get to do it yourself. I was no brilliant stats brain at university but once I started working it had a point. So that's what I try and do. I always try to bring it back to something tangible. Otherwise it does get to be: 'Okay, just write down the answers and we'll learn it later'.*

The two teachers' differing evaluations of the procedural part of statistics versus the interpretative component indicated by the above comments are also demonstrated in their observations, below, comparing the performances of students enrolled for Science degrees to performances of Arts students. Cathy reported that she had taught tutorial groups which, because of timetable constrictions, were predominantly made up of Science students.

Cathy: *You tend to get concentrations of science students in certain tutorials because of timetable constraints and they just stand out. They're so different to the Arts students because they're not scared of formulae. They can deal with it. I was astounded with this group I had last semester. They had it. They could do it. Then I went straight into a predominantly Arts tutorial group. It was a whole different story. It was just much harder for them. That was the main, obvious, difference.*

Wendy: *Very definite differences between Arts and Science students. The Arts, with 'Arts' in inverted commas - your heavily, language based people, tend to get a lot more into analysis at the language level. But your general Arts students will tend to be fazed by the sigma and other symbols and basically a bit number phobic at the beginning. However, they're very good at interpreting results. They're good at telling you what's actually going on. Whereas the Science*

students tend not to worry at all. It doesn't faze them - all the Greek letters, etc. etc. or the subscripts or the algebra, but they're not very good at drawing conclusions. I think that in your traditional physical sciences - what's there is there. You don't have to interpret it. You just report: 'We rejected', or 'This is from a different sample'. Full stop. Well what does that mean? So the different abilities of these groups tend to balance out.

These comments of Cathy and Wendy are particularly interesting in view of Ben's assessment of statistics (excerpt in section 4.5). Ben's responses indicate a focus on and an ease with the procedural part of statistics and he appeared to be content to report results without much thought about the interpretations.

The institutional setting shapes educators' activities as much as it does students' perceptions and actions. Cathy and Wendy agreed that a shortage of teaching time for statistics was a major constraint.

Cathy: *Well, I think in our course students are disadvantaged. In other courses, apparently, they have a lot more time to spend on statistics. They have, I know, at another university, two hours a week of lectures, an hour of tutorials and an hour using computers. They have a different structure to their degree so they can manage their time. Our students are asked to absorb a lot in a short space of time. They do reasonably well at it.*

I asked Cathy and Wendy whether they felt the statistics course could be improved by having more time spent on it.

Cathy: *Yes, but I don't see where the time is coming from. It's an impossible limitation.*

Wendy: *I think with the second year statistics course it could cover more ground than it does. Not in terms of the time we have now. But I think we do need two lectures a week and two hour tutorials. You don't get enough time. You don't get enough time to sit down You've got thirty minds in a tutorial and everyone's going in different directions. Now if you're lucky, fifteen of them will be with you. But of course it's hard when you're up there.*

In addition to lack of time, Wendy felt the lack of opportunity for students to apply statistics was a further constraint and hindered students from seeing statistics as useful and relevant to the field of psychology. Wendy saw this as partly due to students' insufficient statistical knowledge at that stage of the psychology course.

Wendy: *Stats to me was just something you did - until I was actually working. And then it became relevant because it was my stuff. It was my data. What I would like to see is an integration, perhaps, with other areas - social psychology, learning, etc. etc. In that you can do an experiment and apply your own statistics. Unfortunately, my experience, having taught 'Learning' in second and third year, is that, of course, experiments that you run are too complicated for the stats that they know. But I think that you could present them with a data set that could be analysed. Chi square tests are very useful. I think that's where it needs to be tied in, rather than having very discrete parts. I think in second year every unit is very separate: 'Here's your Statistics (topic). And here's Social Psychology. And here's the statistics we've done for you'.*

Cathy concurred with Wendy's view that a lack of statistical knowledge constrained students' understanding of its application:

Cathy: *I think that probably they don't have a clear idea (of statistics) yet. They haven't done any research yet and a lot of them aren't even reading the publications in the literature. They're just reading the general text-books or whatever, so I think they're not really yet able to have an overview of what it's all about. How they might ever be able to use it. Usually they haven't got a clue what to do anyhow, but also any analyses that come up in other areas of the course are more complicated than anything they've covered yet in their actual statistics course. It's hard to coordinate between members of the department.*

Wendy was explicit in stating her aims to students. One of her aims was that students should feel confident enough to use their statistics, irrespective of how it looked "on paper".

Wendy: *I tend to go in at the beginning of every semester and introduce myself and say:*

'Alright, we're going to do this in various ways. One way is to get you through the exams. One aim of mine is - if you are interested - to go further on with it. Another aim of mine is to get every

single one of you up to a level where, alright, on paper you might not be brilliant but you feel competent enough to sit down and understand other things that you might be reading’.

Both teachers expressed enthusiasm for teaching statistics. Cathy described her perceptions of the rewards of teaching statistics as follows.

Cathy: *I love it. I love teaching statistics. The rest of psychology is so waffly in many ways. I feel statistics is just less so. Every year I have at least one student come up to me - I had one this week - and say: ‘I don’t know how you can bear to make a career out of teaching statistics’. But I really enjoy teaching it and I enjoy the gratification when they understand it. You see that they’ve gone from not understanding it to understanding and their own satisfaction in mastering it.*

Wendy related to her students by recalling her own experiences of learning statistics:

Wendy: *The thing that I have to remember is how I learned it. And before you go into a class you’ve got to think: ‘Okay, where are people going to go wrong?’ And second guess the points at which they will. Spend a lot of time clarifying it. I think what’s changed is my perception of how useful it is. I think that I would have had a much easier time in fourth year and in my initial working years if I had learned it. And thinking about it now, I did waste a lot of time. Sure I got through. But it is a very easy subject. And I always point out when students say: ‘Oh I can’t learn this’. ‘You can learn it. The point is whether or not you will’. But of course it’s all mixed up in this aura of: ‘Oh it’s maths.’*

Wendy’s expressions are in accord with the view of activity theorists (Leont’ev, 1981) that goals are the primary elements in determining how and what learning takes place. Wendy clearly empathised with her students’ lack of enthrallment with statistics and was aware that some students identified statistics with mathematics, as was reported in section 4.1. However, her own experience of using statistics enabled her to see its potential for enhancing the perspectives of students in their future lives.

Wendy: *What do they take away from the second year statistics course? I mean the formulas are gone by Xmas. Right? What I think stays is the concept of the underlying populations — the difference between looking different and being significantly different. But I’m not sure you can measure that. I think where that comes out is in their third year. When they go and read papers and do have a feel for an experimental hypothesis and a null hypothesis. Unfortunately we can’t tap into that until further down the track. The students can’t tap into it until they sit down and have that ‘Aha!’ experience. The realisation that: ‘I do know what they’re talking about!’.*

The theoretical framework I have developed highlights the need to recognise that the classroom is not just a place where instruction is received, but a social structure in which students’ and teachers’ actions form and transform. Students’ and teachers’ actions relate to their histories, goals and experiences and the social and cultural environment in which teaching and learning take place. The interviews with Wendy and Cathy provide some insights into this environment.

5. CONCLUSION

The aim of this paper was to bring to the foreground the taken-for-granted context of statistics as a service course and students’ experiences of the subject. The report focused on psychology students’ feelings about learning statistics and their interpretations of statistics. The results could alert researchers and teachers of statistics to look at the ‘big picture’ or system that surrounds statistics service courses and the qualitative variation in students’ awareness of statistics - the diverse meanings of statistical knowledge to students.

The majority of the students who were surveyed in this study were unwillingly studying statistics at university. They reported learning mechanical procedures or mastering decontextualised statistical concepts and methods. A minority group of students expressed a greater willingness to participate in the statistics course and reported more thoughtful and personally meaningful conceptions of statistics. The latter group also achieved higher grades. Statistics educators, Cathy and Wendy observed that a lack of awareness of the functionality of statistical skills and processes made it difficult for the students to experience statistical thinking as personally meaningful. Their comments indicate the

tensions that underpin efforts to engage students in meaningful statistical learning for their profession and the contextual, organisational factors impacting on teaching statistics as a service course at university. Carvalho, (2002) reiterates that changing the focus from calculations and techniques, the “reality in the classroom”, remains a challenge for statistics teaching in the future.

An important component of the learning environment is the way statistics is taught. The interviews with Cathy and Wendy indicated their different perspectives on teaching and learning statistics in a psychology course. These data raise questions for further research. What are the diverse conceptions of statistics, attitudes to teaching statistics and teaching strategies of service teachers? How do teachers’ different beliefs about and positions on teaching statistics in service courses relate to student learning, including students’ ongoing developments of conceptions of statistics?

A further direction for research indicated by this study is the investigation of organisational frameworks that could promote active learning and help students develop positive attitudes to learning statistics. Cobb (1993, para. 82) asserts that:

Learning must be active if it is to build a student’s sense of responsibility for the process; lecture-based courses undermine the student’s sense of responsibility for learning. The teacher is neither producing a course for the student nor producing a student for an employer. Both these models make our students passive consumers rather than active constructors of their education.

Although the traditional lecture format and delivery of information model that Cobb (1993) deplores remains the mainstay of some statistics courses, alternative and diverse ways of thinking about statistical learning have been emerging. Research on problem-based learning methods in a graduate biostatistics course (Boyle, 1999) indicated that most of the students perceived that their overall understanding improved greatly with this approach, although more time and effort were required. There were also some indications that the problem-based course helped students develop suitable analysis plans for their research data and decreased their anxiety. Garfield (1993) reviewed the use of small group cooperative learning in statistics classes and advocated research in this area. Mohammad Yusof and Tall (1999) showed that a mathematics course encouraging cooperative problem solving and reflection changed students’ attitudes in a desired direction. That is, changes were from the negative attitudes teachers had come to “expect” from students towards attitudes “preferred” by their teachers.

This paper indicates some major issues for statistics education in the context of students learning statistics as a compulsory service course. Challenges include providing students with learning experiences which enable them to reinterpret statistics as personally meaningful knowledge rather than a culturally “necessary” body of skills and concepts and changing conceptions of statistics from classroom exercises in which “computers do all the work” to a tool for use in professional and individual life. Service courses in statistics could provide opportunities to connect professional statisticians and occasional users and to induct undergraduate students into the community of statisticians.

ACKNOWLEDGEMENTS

An earlier version of this paper was presented at the 10th Improving Student Learning Symposium (Brussels, September, 2002). I gratefully acknowledge the assistance of Peter Fletcher in this research. I would also like to thank the editor and the reviewers of this paper for their thoughtful and challenging comments and suggestions. Their time and effort is much appreciated.

REFERENCES

- Ashworth, P., & Lucas, U. (2000). Achieving empathy and engagement: a practical approach to the design, conduct and reporting of phenomenographic research. *Studies in Higher Education*, 25(3), 295-308.
- Boyle, C. (1999). A problem-based learning approach to teaching biostatistics. *Journal of Statistics Education*, 7(1). [Online: www.amstat.org/publications/jse/secure/v7n1/boyle.cfm]

- Carvalho, C. (2002). Accepting a challenge. *Statistics Education Research Journal*, 1(1), 18-20.
- Cobb, G. (1993). Reconsidering Statistics Education: A National Science Foundation Conference. *Journal of Statistics Education*, 1(1), [Online: www.amstat.org/publications/jse/v1n1/cobb.html]
- Crawford, K., Gordon, S., Nicholas, J., & Prosser, M. (1994). Conceptions of mathematics and how it is learned: The perspectives of students entering university. *Learning and Instruction*, 4, 331-345.
- Dahlgren, L. O. (1984). Outcomes of learning. In F. Marton, D. Hounsell and N. J. Entwistle (Eds.), *The Experience of Learning: Implications for teaching and Studying in Higher Education*, (pp. 19-35, Second edition). Edinburgh: Scottish University Press.
- Department of Education, Science and Training (2002). *Striving for quality: Learning, teaching and scholarship*. Department of Education, Science and Training, Canberra, Australia. [Online: <http://www.dest.gov.au/crossroads/pubs.htm>]
- Engeström, Y. (1999). Expansive visibilization of work: An activity theory perspective. *Computer Supported Cooperative Work*, 8, 63-69.
- Gal, I., & Ginsburg, L. (1994). The Role of Beliefs and Attitudes in Learning Statistics: Towards an Assessment Framework. *Journal of Statistics Education*, 2(2). [Online: www.amstat.org/publications/jse/v2n2/gal.html]
- Garfield, J. (1993). Teaching Statistics Using Small-Group Cooperative Learning. *Journal of Statistics Education*, 1(1). [Online: www.amstat.org/publications/jse/v1n1/garfield.html]
- Garfield, J. (1994). Beyond Testing and Grading: Using Assessment to Improve Student Learning. *Journal of Statistics Education*, 2(1). [Online: www.amstat.org/publications/jse/v2n1/Garfield.html]
- Garfield, J., Hogg, B., Schau, C., & Whittinghill, D. (2002). First Courses in Statistical Science: The Status of Educational Reform Efforts. *Journal of Statistics Education*, 10(2). [Online: www.amstat.org/publications/jse/v10n2/garfield.html]
- Gordon, S. (1995). A theoretical approach to understanding learners of statistics. *Journal of Statistics Education*, 3(3). [Online: www.amstat.org/publications/jse/v3n3/Gordon.html]
- Gordon, S. (1998). *Understanding Students Learning Statistics: An Activity Theory Approach*. Unpublished Doctoral Thesis, The University of Sydney. Sydney. Available online, Australian Digital Theses: <http://adt.caul.edu.au/>.
- Gordon, S. (1999). An Instrument for Exploring Students' Approaches to Learning Statistics. *Paper presented at the Annual Meeting of the American Educational Research Association (Montreal, Canada, April 1999)*. [Online: <http://ericae.net/ericdc/ED440142.htm>]
- Green, D. (2002). Lessons from the past. *Statistics Education Research Journal*, 1(1), 21-23.
- Hogg, R. (1999). Let's Use CQI in Our Statistics Programs. *The American Statistician*, 53(1), 7-14.
- Hubbard, R. (1997). Assessment and the Process of Learning Statistics. *Journal of Statistics Education*, 5(1), [Online: www.amstat.org/publications/jse/v5n1/hubbard.html]
- Kent, S. (2002). Crossroads@cross purposes? *Advocate: Journal of the National Tertiary Education Union*, 9(2), 16-18.
- Leont'ev, A. N. (1978). *Activity, Consciousness, and Personality*. (M. J. Hall, Trans.). Englewood Cliffs, New Jersey: Prentice-Hall.
- Leont'ev, A. N. (1981). The problem of activity in psychology. In J. V. Wertsch (Ed.), *The Concept of Activity in Soviet psychology*, (pp. 37-71). New York: M. E. Sharpe.
- Lerman, S. (1996). Intersubjectivity in mathematics learning: A challenge to the radical constructivist paradigm. *Journal for Research in Mathematics Education*, 27(2), 133-150.
- Marton, F. (1986). Phenomenography - a research approach to investigating different understandings of reality. *Journal of Thought*, 21(3), 28-49.
- Marton, F. (1988). Describing and improving learning. In R. Schmeck (Ed.), *Learning Strategies and Learning Styles*, (pp. 53-82). New York: Plenum Press.

- Marton, F. & Booth, S. (1997). *Learning and Awareness*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Meletiou, M. (2002). Conceptions of variation: A literature review. *Statistics Education Research Journal*, 1(1), 46-52.
- Mohammad Yusof, Y. & Tall, D. (1999). Changing attitudes to university mathematics through problem solving. *Educational Studies in Mathematics*, 37, 67-82.
- Moore, D. (2001). Undergraduate programs and the future of academic statistics. *The American Statistician*, 55(1), 1-6.
- Nicholls, D. F. (2001). Future directions for the teaching and learning of statistics at the tertiary level. *International Statistical Review*, 69(1), 11-15.
- O'Connell, A. (2002). Student Perceptions of Assessment Strategies in a Multivariate Statistics Course. *Journal of Statistics Education*, 10(1).
[Online: www.amstat.org/publications/jse/v10n1/oconnell.html]
- Reid, A., & Petocz, P. (2002). Students' conceptions of statistics: a phenomenographic study. *Journal of Statistics Education*, 10(2). [Online: www.amstat.org/publications/jse/v10n2/reid.html]
- Sierpinska, A., & Lerman, S. (1996). Epistemologies of mathematics and of mathematics education. In A. J. Bishop (Ed.), *International Handbook of Mathematics Education*, (pp. 827-876). Dordrecht: Kluwer Academic Publishers.
- Svensson, L. (1994). Theoretical foundations of phenomenography. In R. Ballantyne & C. Bruce (Eds.), *Phenomenography: Philosophy and Practice. Proceedings of the 1994 Phenomenography Conference*, (pp. 9-20). Brisbane: Centre for Applied Environmental and Social Education Research, Queensland University of Technology.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind*. Cambridge, Massachusetts: The MIT Press.
- Vygotsky, L. S. (1962). *Thought and Language*. Cambridge, Massachusetts: The M.I.T. Press.
- Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.
- Yilmaz, M. R. (1996). The challenge of teaching statistics to non-specialists. *Journal of Statistics Education*, 4(1). [Online: www.amstat.org/publications/jse/v4n1/yilmaz.html]
- Zetterqvist, L. (1997). Statistics for chemistry students: how to make a statistics course useful by focusing on applications. *Journal of Statistics Education*, 5(1).
[Online: www.amstat.org/publications/jse/v5n1/zetterqvist.html]

SUE GORDON
Mathematics Learning Centre
Carslaw Building, F07
The University of Sydney
NSW 2006, Australia

SOME ISSUES ABOUT THE STATUS OF STATISTICS TEACHING IN AGRICULTURAL COLLEGES IN ARGENTINA

MARÍA VIRGINIA LÓPEZ

Universidad de Buenos Aires, Argentina
mvlopez@agro.uba.ar

MARÍA DEL CARMEN FABRIZIO

Universidad de Buenos Aires, Argentina
mcfabrizio@sinectis.com.ar

MARIA CRISTINA PLENCOVICH

Universidad de Buenos Aires, Argentina
plencovi@agro.uba.ar

HERNAN GIORGINI

Universidad de Buenos Aires, Argentina
hernang@coneau.gov.ar

SUMMARY

The status of Statistics teaching has not been sufficiently explored in Agricultural Colleges in Argentina. Although Statistics is considered as an important subject in different academic institutions, there is very little information about the way that it is taught in different university curricula. The aims of this study were to (a) gather information about the place of Statistics in college programs through different indicators, and (b) explore different issues concerning the teaching of Statistics in agricultural colleges, such as epistemological views, academic organization, etc. For this purpose, a survey was conducted in the main agricultural colleges in Argentina. Twenty-three teaching teams from different university answered a questionnaire. The responses were analyzed and categories were built in order to draw some conclusions.

Keywords: Agricultural Colleges; Statistics teaching; Survey

1. INTRODUCTION

Argentina is an extended agricultural and livestock production country with many universities granting degrees in Agricultural Studies.

Graduates from Agricultural colleges are mainly Agricultural Engineers - a word that has the same etymology as “genius” – and they have an active intervention in the interface between natural and social systems. These professionals are, above all, decision-makers, no matter whether they manage agribusiness companies, do research, teach, or work for public agencies or in the private sector. Their decisions are based on predictions that are heavily conditioned by uncertainty. This uncertainty is remarkable in agricultural activities as they rely mainly on climate factors, and biological and ecological natural processes, as well as on socio-economic variables. However, information reduces this uncertainty and provides decision makers with a framework that contributes to a better professional intervention.

In the 20th century, Statistics was considered a fundamental scientific tool, and the basis for the experimental scientific method. Presently, Statistics may be found as a subject in different kinds of academic programs. Nowadays it is unconceivable to think of research projects without statistical support, thus this discipline is part of the curriculum in different programs at the undergraduate level studies- post-secondary programs leading to four-year degrees- and in post-graduate studies- more than one-year programs leading to specialist, master's or doctoral degrees. A sound, scientific culture demands training in statistical and probabilistic thinking, and rejects a single, deterministic data interpretation.

Statistics is an ever-changing scientific realm, and its teaching is mainly influenced by computer availability, and by the development of specific software. Neither the process of specific content learning, nor the development of calculation skills is the most important teaching aim. The teaching focus is placed on a specific way of reasoning that is deeply rooted in the idea that reality is viewed as a set of randomized phenomena. Data collection becomes useful only after a proper statistical analysis. Teaching statistical reasoning has to be considered only within this scope. Further, in Statistics the context of application (in this case, the agricultural sector) is the source of meaning for collecting, processing, and interpreting information.

This way of understanding reality enables professionals from different areas to manage sound concepts. Their decision-making processes are then supported by inductive inferences and statistical information, and are based on the critical reading of scientific publications, and on interdisciplinary research. Thus, teaching Statistics becomes a formidable challenge. Teachers should focus on thinking processes, interactive work, interdisciplinary projects, learning assisted by computing, etc. A primary goal of teaching Statistics is to enable students to render reasoned judgments about data, or about data interpretation, using multiple tools (Gal & Garfield, 1997, Cobanovic, 2002).

This study was carried out to explore and collect information about the status of Statistics teaching in Argentine agricultural programs. More specifically, the study aimed to (a) gather information about the place of Statistics in college programs, analyzing its resources, weaknesses, and strengths; and to (b) explore different issues concerning the teaching of Statistics in agricultural colleges, such as epistemological views, academic organization, etc. Smeeton (2002) carried out a similar study to collect information about teaching in undergraduate courses in dental statistics.

2. BACKGROUND

Towards the end of the 19th century, there was only one higher education institution in Argentina that specialized in the Agricultural Sciences, the "Escuela Superior de Santa Catalina" (1882), which then merged into the Agronomic and Veterinary Provincial School of La Plata (1890). In 1904, the "Instituto Superior de Agronomía y Veterinaria" was founded in Buenos Aires and then was annexed to the University of Buenos Aires. It was in this institution- now Facultad de Agronomía, Universidad de Buenos Aires- where formal studies in Statistics started in 1937. They were introduced as a second course of Mathematics in the Agricultural Engineering Program (Marotta, 1943).

In Argentina the teaching scenario is affected by the students' lack of previous statistical knowledge, course size, infrastructure, curriculum, and administrative structure of programs.

Teaching any discipline implies some underlying epistemological assumptions. In the case of Statistics, there is not a general agreement as to its epistemological nature, and authors have different approaches as regards its definition. The Merriam-Webster's dictionary (1995) defines to Statistics "as a branch of Mathematics dealing with the collection, analysis, interpretation and presentation of masses of numerical data". In her book, Batanero (2001) agrees with this idea: "*Notwithstanding its axiomatic foundation, Statistics may well be the only branch of Mathematics where debate is still going on about the interpretation of basic concepts*" (p. 9). Other authors consider Statistics as a set of methods or tools ready to be used in any context. Iversen (1992) states that Statistics is different enough from Mathematics, and that any Statistics course needs to be taught by a person having considerable experience with data and with the issues of statistical methods. In his opinion, such a course should have the aim to convey the nature of uncertainty and how seemingly random events have patterns to them; patterns that Statistics is created to look for. In addition, he establishes that it

may be possible to define Statistics as a set of methods to study regularities in the face of variability. Some others, as Moore (1992), define Statistics as a science that has emerged as a discipline in its own right, that it is a subject in its own right. Even though Statistics is a mathematical science, it is neither a subfield of Mathematic nor merely a collection of methods that can be understood as ancillary to any substantive discipline such as psychology, business, or engineering. For these authors, Statistics has its own substance, its own distinctive concepts and modes of reasoning.

These definitions represent epistemological views that give rise to different teaching and research organizations within the academic institutions. These epistemological views can be analyzed not only through the definition of Statistics given by the teaching teams, but also through the bibliography they work with, the didactic strategies they use, and some other indicators. In addition, they can be indirectly studied through the faculty career paths, their research lines, and teaching profiles.

3. MATERIALS AND METHODS

A survey was conducted of all the universities having colleges that grant degrees in Agricultural studies in Argentina. Both public and private universities were involved in the study. The survey was conducted from August to December 2001. Due to great distances between the universities and our research center (Facultad de Agronomía, Universidad de Buenos Aires) most of the questionnaires were answered by e-mail. We surveyed the 27 agricultural colleges (see Table 1) by using a questionnaire that had 18 items relating to characteristics of the teaching teams, their academic and research activities, course organization, teaching methodologies, students' difficulties and teachers' definitions of Statistics.

Table 1. Agricultural Colleges in Argentina

Kind of University	ID ⁽¹⁾	University	Province
Public	1	Universidad Nacional de Catamarca	Catamarca
	2	Universidad Nacional del Comahue	Río Negro
	3	Universidad Nacional de Córdoba	Córdoba
	4	Universidad Nacional de Cuyo	Mendoza
	5	Universidad Nacional de Entre Ríos	Entre Ríos
	6	Universidad Nacional de Jujuy	Jujuy
	7	Universidad Nacional de La Pampa	La Pampa
	8	Universidad Nacional de La Plata	Buenos Aires
	9	Universidad Nacional de La Rioja	La Rioja
	10	Universidad Nacional del Litoral	Santa Fe
	11	Universidad Nacional de Lomas de Zamora	Buenos Aires
	12	Universidad Nacional de Mar del Plata	Buenos Aires
	13	Universidad Nacional del Nordeste	Corrientes
	14	Universidad Nacional de Salta	Salta
	15	Universidad Nacional de San Luis	San Luis
	16	Universidad Nacional de Santiago del Estero	Sgo. del Estero
	17	Universidad Nacional del Sur	Buenos Aires
	18	Universidad Nacional de Tucumán	Tucumán
	22	Universidad de Buenos Aires	Buenos Aires
	23	Universidad Nacional de Luján	Buenos Aires
	24*	Universidad Nacional de Río Cuarto	Córdoba
	25*	Universidad Nacional del Centro	Buenos Aires
	26*	Universidad Nacional de Rosario	Santa Fe
	Private	19	Universidad del Salvador
20		Universidad Católica de Córdoba	Córdoba
21		Universidad de Concepción del Uruguay	Entre Ríos
27*		Universidad de Morón	Buenos Aires

*ID stands for identification number. * These agricultural colleges did not respond to the survey.*

Twenty three out of the twenty seven chairs replied. The college academic unit is the chair, which contains both professors and teaching assistants. The questionnaire was addressed to the Statistics Chair Heads, who are responsible for the epistemological, teaching, research, and extension lines developed by their teams. The item dealing with Statistics definition was used to get information about the epistemological ideas supported by the chairs.

At first, the relations among variables were explored using exact chi-square tests. Agricultural Colleges were classified in groups by using hierarchical cluster methods using the quantitative variables. Cluster analysis was performed using the complete linkage method with Euclidean distances. No variable transformation was done. Further, a canonical discriminant analysis provided a representation of the populations found in the previous stage. SAS (1988) software was used for the analyses.

4. RESULTS

We first present the statistics teaching team characteristics as regards their academic background and activities in research and consulting work. Secondly, we place the subject Statistics within the Agronomy curriculum, its workload, and class size. Then, we discuss teaching methodologies, and in the fourth section we deal with common teaching difficulties. In the fifth section, teaching team opinions about Statistics definitions are analyzed, and finally university characteristics are presented by clusters.

4.1. CHARACTERISTICS OF STATISTICS TEACHING TEAMS IN AGRICULTURAL COLLEGES

Three Argentine colleges (among which there are two private ones) do not have teaching assistants, and one has no professors. Colleges with Agricultural Engineering Programs have 5 teaching members per chair in average, ranking from a minimum of 1 to a maximum of 10. Private colleges have a smaller number of professors and teaching assistants: in two cases there is only one professor per chair, and only three in the remaining one. In our opinion, the low number of professors is due to low salaries, which are paid by both public and private institutes.

In this study, an average of 43.6% faculty members per chair holds the degree of Agricultural Engineer. The number of these professionals per chair varies from 0 to 5 with a median of 2. This profession is the most frequently found across Statistics chairs, even though there are other related professionals, holding other degrees mainly in Biology, Mathematics, and Statistics. In only two of the chairs the professors have a degree in Statistics. As regards postgraduate studies, an average of 6% of the teaching team have Ph.D. degrees, minimum 0, maximum 4 (median = 0) and 28% have Master's degrees, minimum 0, maximum 3 (median = 1). Their postgraduate programs do not belong exclusively to the statistical area. In eight colleges there are no postgraduate degree holders and in ten, teachers having no postgraduate degrees lead teaching teams. Twelve out of the twenty three carry out consulting work and seventeen are involved in research activities.

The scarce number of teachers having a degree in Statistics is remarkable. The number of postgraduate degree holders is low, however, there is an increasing number of teachers taking postgraduate programs. Consulting work and research activities are closely related to full-time or part-time activities.

4.2. CHARACTERISTICS OF STATISTICS AS A CURRICULUM SUBJECT

Most of the students have their first contact with Statistics at the university level. In general, Agricultural College programs have one or two compulsory Statistics courses. In general, the first course deals with Statistics basic knowledge and the second with experimental design. Eighteen out of 23 colleges have their first Statistics course in the second year of their programs, comprising an

average of 100 hours of work load (minimum: 64, maximum: 168). Only 7 colleges have a second Statistics course in their programs, comprising an average of 66 hours (minimum: 45, maximum: 90). This second course is taught in the second year in five colleges and in the fifth year in the rest. On the average, 76 students attend the first course (minimum: 10, maximum: 200). The second course has less students, averaging 39 (minimum: 12, maximum: 80). Personal computers are used only in 11 out of 23 colleges in the first course, and 3 out of 7 in the second one. Software programs mentioned were Infostat, Macanova for Windows, Excel, SAS, Statistica, Statistix, Statgraphics and SPSS.

In our opinion we consider that teaching Statistics before university may contribute the strengthening of concepts which can be later applied in professional careers.

4.3. TEACHING METHODOLOGY

Ten of the surveyed chairs teach theory and practice in an integrated way. For example, in the same teaching period they devote the first part of the class to the theoretical approach, and then proceed to practical aspects, such as solving problems, exercises, etc. The rest separate theoretical and practical classes.

All the chairs have more than one compulsory reading text. Fifteen chairs use textbooks, and at least one of them is related to Applied Agriculture or Biology. Fourteen chairs publish their own reading texts and one of the colleges has developed its own software. Seventeen different texts are mentioned. Di Rienzo et al (2000), Steel and Torrie (1986), Mendenhall (1990) and Walpole and Myers (1999) are given more than once.

Twelve out of 23 chairs work interdisciplinarily in the statistics course with other chairs, and 12 use overhead projectors. Only one chair organizes field trips.

From our standpoint, interdisciplinary work contributes to improve content integration by interpreting and solving real, agricultural problems. At Buenos Aires University, in our Statistics courses we worked together with the Plant Physiology Chair in designing and controlling different experiments.

4.4. TEACHING WEAKNESSES

Teachers often mention curricular and infrastructure difficulties as their main obstacle to teaching. As their courses are placed too “early” in the curriculum, students cannot understand the practical use of Statistics, and of the biological models being taught. Some teachers mention that other subjects demand too much from the students, so they cannot devote much time to Statistics. They also mention as obstacles uncomfortable classrooms, and the lack of availability of computing hardware and software. None of the private universities mentioned any infrastructure problems. Some faculty members consider that overcrowded classes, the students’ lack of interest, and of previous basic mathematical skills hamper their learning.

Sixteen chairs mention “probability” as one of the hardest topics, followed by the “tests of hypotheses”.

4.5. STATISTICS DEFINITION

Definitions were grouped in broad categories. Statistics is considered as:

- (a) A branch of Mathematics;
- (b) A set of methods or tools ready to be used in different contexts;
- (c) An autonomous science;
- (d) Other.

It may be of interest to illustrate some of the definitions given by the teaching teams corresponding to the three main categories:

Definition. (a): Statistics as a branch of Mathematics

La Plata College defined Statistics as: "...a branch of Mathematics applied to information analysis".

Definition (b): Statistics as a set of self-governing procedures for decision-making

Córdoba College defined Statistics as: "...a discipline strongly based on Mathematics, having an engineering spirit, which approaches the problem of the collection and analysis of data, and the development of techniques to obtain information about systems generating a response with random errors".

Catamarca College considered it as "...a tool to provide the correct information, analyze it properly, and give inputs for decision-making (not setting aside common sense, though)".

Similarly, Comahue pointed out: "...is a set of methods backed up by probability theory for the treatment of information and decision-making".

Salta stated "Statistics is a discipline that provides tools, which are used not only to process large volumes of data and make them understandable, but which integrate the research process. It is also a tool that helps take more objective decisions. It provides a way to gather and analyze data accurately, contrast hypotheses, estimate parameters, and draw valid and reliable conclusions and recommendations within uncertainty".

San Luis indicated: "... is a set of methods that allows us to present and analyze information that may enable us to take decisions. It is related to mathematics in terms of probability theory".

Universidad del Salvador considered it as: "a set of criteria and tools useful for creating, analyzing, and interpreting valuable information for the development of production systems".

Definition (c): Statistics as an autonomous science

Cuyo pointed out: "from an applied perspective (...) it is a science that has an identity of its own, and provides tools for capturing real data, analyzing variability, drawing general conclusions based on probabilistic theory, and the communication of results".

Mar del Plata stated: "Statistics is an important aspect of research methodology as it contributes to: 1. Clarify the research problem. 2. Set the research objectives. 3. Plan the search for reliable information suited to the research problem and objectives 4. Process information through specific techniques 5. Interpret results within the context of the methodology used".

Two of the chairs view Statistics mainly as a branch of Mathematics applied to information analysis -definition (a)- and eleven conceive it as a set of self-governing procedures for decision-making- definition (b)-. Nine teachers envisage it as a scientific dealing extracting information from numerical data -definition (c)-. The remaining chair has a contradictory answer. Agricultural Engineers lead the two chairs that gave definition (a). On the other hand, the two chairs led by Statisticians define Statistics as an autonomous science -definition(c)- and the three chairs led by Biologists consider it a set of tools- definition (b). No differences were observed as regards postgraduate studies (Chi-Square = 0.8, p = 0.82), neither in relation with the position of Head of the Chair (Chi-Square = 3.93, p = 0.17), nor with undertaking interdisciplinary courses (Chi-Square = 0.90, p = 0.82).

4.6. UNIVERSITIES CLUSTERING

Faculty teams were grouped by characteristics regarding the number of Ph.D. and Master's degrees, and the number of teachers working as advisers and researchers. Cluster analysis identified three clusters of universities in the dendrogram in Figure 1. The first cluster was made up by 13 colleges, including the private ones. They have a low number of members and scarce postgraduate

background (only one Ph.D. within the group), and only a few of them provide technical advise and work on research projects. The second cluster comprises only one college that has a high number of teachers holding Ph.D. and Master's degrees, and all of them are members of advising and research programs. The third cluster, made up of 9 colleges, presents a medium level of postgraduate degrees, working in consulting activities and doing research work. The first canonical variable in the canonical discriminant analysis can explain 83% of the total variance. This variable appears to be largely influenced by the variables related to the degree reached by the members of the chair (Table 2). The second canonical variable is positively influenced by the number of professors and inversely related to the number of Ph.D. degrees. The three clusters can be easily discriminated by the first and the second canonical variables. The plot of scores of canonical variable one against canonical variable two is shown in Figure 2. This confirms the same grouping identified in the dendrogram. There is a clear separation of the second cluster (Universidad de Córdoba) with high values on canonical variable one and high negative value on canonical variable two.

Table 2. Raw canonical coefficients

Variable	Canonical variable	
	1	2
Number of professors	0.732	0.690
Number of assistant teachers	-0.079	0.246
Number of Ph.D.	1.759	-1.973
Number of Master	0.954	0.497
Number of teachers doing consulting	0.462	0.030
Number of teachers doing research work	0.313	0.260

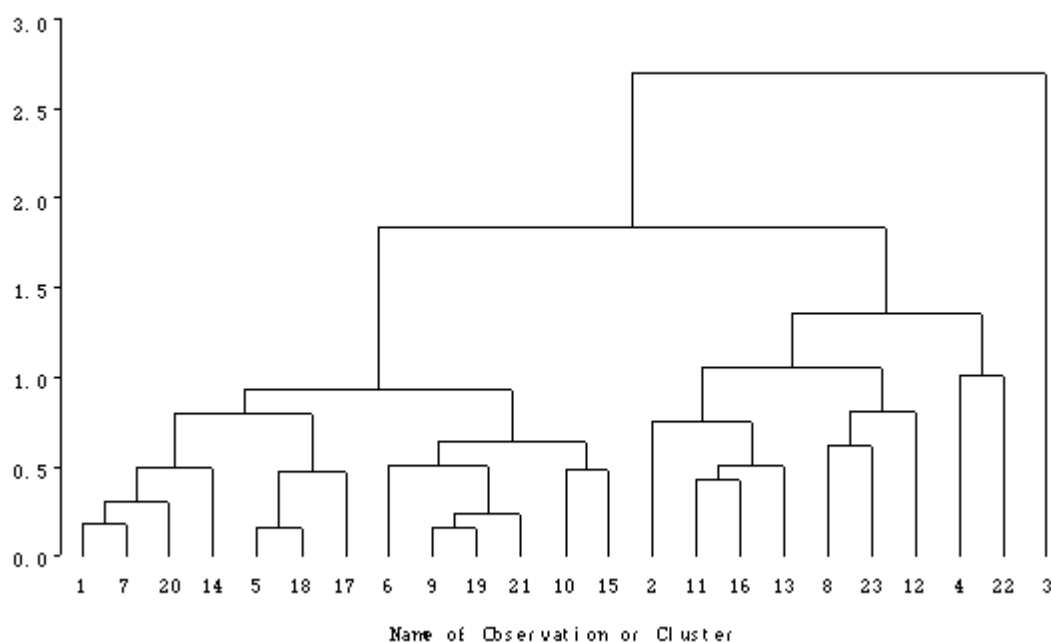
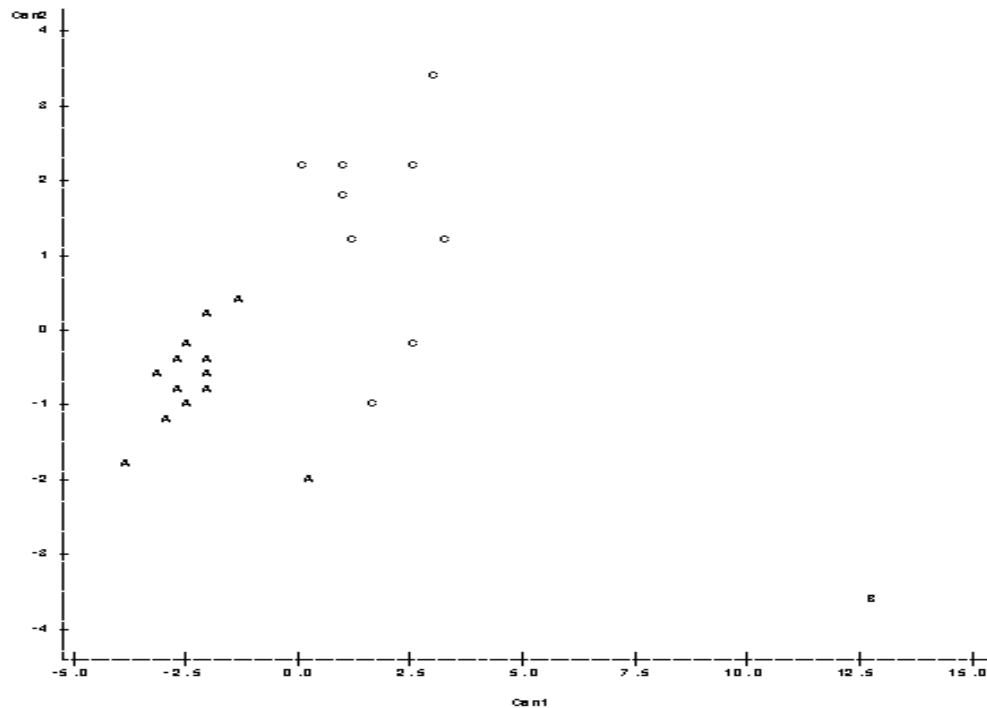


Figure 1. Dendrogram for the twenty-three universities



First, second and third clusters are indicated by the letters A, B and C, respectively.

Figure 2. Plot of the twenty three universities for the canonical variables

5. DISCUSSION

The need to get information about the status of teaching statistics in Argentine Agricultural Colleges was the main purpose of this study. In analyzing the results, we notice that some characteristics are homogeneous, such as the use of more than one compulsory reading texts, some related to Agriculture or Biology. In almost all colleges inadequate infrastructure was mentioned as a major barrier. In some cases, it is not possible to use computers in class. Teachers also mentioned that as the courses are too early in the curriculum, students cannot understand the practical use of Statistics. Cobanovic (2002) points out the same problem. Other countries having no infrastructure restrictions seem to have the same problems. Magel (1996) states that student motivation is perceived to be a problem in the class. However, she mentions some techniques which pose interesting challenges for teachers, and improve teaching results without affecting costs. The teaching team characteristics did not seem to have any impact in teaching methodology. Ten chairs are managed by teachers having no postgraduate degrees.

In most of the colleges, probability and random variables are the topics that present most difficulties for students. In a further study, it would be interesting to explore the didactic strategies used in teaching them. Only twenty-three colleges were differentiated as regards the number of teachers, the postgraduate degrees, and the number of teachers involved in technical advice and research. These characteristics do not reflect the teaching methodologies, the difficulties reported, nor the epistemological ideas underlying the Statistics definition. The definition of Statistics as a branch of Mathematics was less popular as only two colleges held it. Curiously enough, the professors of those chairs of Statistics are not Mathematicians but Agricultural Engineers. The rest of the universities mainly consider Statistics either as a set of tools or an autonomous science. All the chairs led by Biologists defined Statistics as a set of tools, whereas all the chairs led by Statisticians defined it as a science. We did not observe any differences as regards the teaching team's undergraduate and postgraduate studies.

Córdoba University differed from the rest since it has a chair with many professors and assistants, and all its members but one have a Master's or Ph.D. degree, all provide advisory services and do research, and have developed statistic software and a manual.

In a future study, it would be interesting to contrast these findings, which are mainly concerned with teaching processes and organization, with the students' learning processes to assess students actual achievements.

ACKNOWLEDGEMENTS

We would like to thank Carol Blumberg for her assistance in the final English version.

REFERENCES

- Batanero, C. (2001). *Didáctica de la Estadística* [Didactics of Statistics]. Granada, Spain: Grupo de Investigación en Educación Estadística.
- Cobanovic, K (2002). Role of statistics in the education of agricultural science students. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town: International Statistical Institute. CDROM.
- Di Rienzo, J, Casanoves, F., González, L., Tablada, M., Diaz, M. P., Robledo, C. W., & Balzarini, M. (2000). *Estadística para las Ciencias Agropecuarias*. [Statistics for the Agricultural Sciences]. Córdoba, Argentina: Editorial Triunfar.
- Gal, I., & Garfield, J. B. (1997). *The Assessment Challenge in Statistics Education*. Amsterdam: IOS Press & International Statistical Institute.
- Iversen, G. R. (1992). Mathematics and statistics: An uneasy marriage. In F. Gordon & S. Gordon (Eds.) *Statistics for the Twenty-First Century* (Notes, No. 26, pp. 37-44). Washington D.C.: Mathematical Association of America.
- Magel, R. (1996). Increasing student participation in large introductory statistics classes. *The American Statistician*, 50, 51-56.
- Marotta, P. (1943). *La Facultad de Agronomía y Veterinaria de la Universidad de Buenos Aires*. [Agriculture and Veterinary Schools of Buenos Aires University]. Buenos Aires: Imprenta de la Universidad.
- Mendenhall, W. (1990). *Estadística para administradores*. [A Course in Business Statistics]. México: Grupo Editorial Iberoamericana.
- Merriam-Webster's Collegiate Dictionary* (1995). Tenth Edition. Springfield: Merriam-Webster, Incorporated.
- Moore, D. S. (1992). Teaching statistics as a respectable subject. In F. Gordon & S. Gordon (Eds.) *Statistics for the Twenty-First Century* (Notes, No. 26, pp. 14-25). Washington D.C.: Mathematical Association of America.
- SAS Institute (1988). *SAS/STAT User's Guide*. Release 6.03 Edition. Cary, North Carolina: SAS Institute.
- Smeeton, N. (2002). Undergraduate courses in dental statistics in Britain and Ireland. *Statistics Education Research Journal*, 1(2), 47-50.
- Steel, R. & Torrie, J. (1986). *Bioestadística: Principios y procedimientos*. [Biostatistics: Principles and Procedures]. 2da. Edición. México: McGraw Hill.
- Walpole, R., Myers, R. (1999). *Probabilidad y estadística para ingenieros*. [Probability and Statistics for Engineers and Scientists]. 6ta. Edición. México: Prentice Hall.

MARIA VIRGINIA LOPEZ
Avenida San Martín 1447
(1417) Buenos Aires, Argentina

**APPENDIX: QUESTIONNAIRE FOR STATISTICS TEACHERS
(UNDERGRADUATE AGRICULTURAL PROGRAMS)**

1) University:

Type: (check) Public Private:

City:

Province:

College(or academic unit):

Program:

Chair /Department:

2) Does the Chair/Department provide classes only to this program? (check)

Yes No

3) If your answer is negative, state whether it provides classes only to the agricultural area.

Yes No

4) Faculty:

Position	Full-time (F) Part-time (P) Simple (S) (1)	Grade	Post-graduate

(1): Full-time (about 40 hs per week); Part-time: (about 20 hs per week); Simple (about 10 hs per week).

5) How many teachers in the Chair/Department have taken Postgraduate courses about teaching Statistics?.....

Mention the courses:

6) How many teachers in the Chair/Department work as agricultural consultants?

.....

7) How many teachers in the Chair/Department carry out interdisciplinary research activities in agricultural areas?

8)

	Course (1)	Year (2)	Hours (3)	Number of teachers	Course size (4)
1					
2					
3					
4					

(1): Mandatory courses related to Agricultural Sciences

(2): Year in the curriculum

(3): Total hours in the term

(4): Mean number of students per class.

9) Teaching Methodology:

Classes may be described mainly as (check)

Course	Theoretical-Practical	Theoretical and Practical (in different sessions)	Only theoretical	Only practical
1				
2				
3				
4				

10) Students have to read (check)

Course	No reading	Manual *	Notes prepared by Chair	Papers	Other (state)
1					
2					
3					
4					

* State author/s name:

11) State the percentage of class exercises which need a PC to be solved?

Course	%	Number of Students per PC	Software
1			
2			
3			
4			

12) Other teaching media: (overhead projector, videos, etc.) State:

.....

13) If you work interdisciplinarily, state the disciplines you work with:

.....

14) State the main teaching drawbacks:

.....

15) List three topics causing learning difficulties:

.....
.....
.....

16) Statistics is ... (check):

- (a) a branch of Mathematics applied to information analysis.....
- (b) a set of self-governing procedures for decision-making
- (c) a scientific method for extracting information from numbers
- (d) Other (state)

STATISTICAL EDUCATION FOR PHD STUDENTS IN UK MEDICAL SCHOOLS

PAULA R. WILLIAMSON
University of Liverpool
prw@liv.ac.uk

GILLIAN A. LANCASTER
University of Liverpool
g.lancaster@liv.ac.uk

SUMMARY

Little is known about the provision of statistics teaching for PhD students in UK medical schools. A recent survey found that statistics courses were available to PhD students in 13 of 21 schools responding. The provision across these 13 schools was variable in terms of contact hours and content. At a meeting of 27 medical statistics teachers, consensus was reached that such teaching should be undertaken by a subject specialist, however there was no consensus as to the best mode of delivery. We describe the rationale for, content of, and student feedback from our newly developed course programme which emphasises aspects of both design and analysis of research projects.

Keywords: *Statistics education research; Teaching; Medicine; Postgraduates; Survey*

1. INTRODUCTION

The importance of statistical education for PhD students undertaking medical research is becoming increasingly recognised. Medical Research Council guidelines (MRC, 2002, Section 5.2 Skills training) state that

The MRC expects each student to have access to training in generic research skills (eg. experimental design, data handling, statistics, intellectual property rights, exploitation), and access to transferable skills training, (eg. oral and written communication, IT skills and time management) during their MRC-funded training period, in addition to subject specific training.

Whilst some discussion has taken place regarding statistical teaching in undergraduate courses for both medical students (Lancaster, 2002), and dental students (Smeeton, 2002), there is little known about the provision of statistics training for PhD students in UK medical schools.

Researchers may also improve their understanding of statistical issues through consultation with a statistician. Statistical consultancy for researchers has been the subject of several papers (Svensson, 2001; Ospina & Ortiz, 2001; Shia, 2001; Jolliffe, 2001; Godino, Batanero, & Jaimez, 2001; Belli, 2001; Mji & Glencross, 2002; Bangdiwala et al., 2002). However there is no information available concerning the provision of such support for PhD students in UK medical schools.

In order to make informed decisions regarding the content and delivery of statistical teaching for PhD students in the Faculty of Medicine at Liverpool, we undertook a survey of UK medical schools. In this paper we report the results of the survey, together with a discussion held at the Burwalls annual meeting of 'Teachers of Medical Statistics', and present our plans and rationale for the teaching of statistics to our PhD students.

2. METHODS

Medical statisticians working in each of the 26 UK medical schools were identified from either the Burwalls 'Blue Book' (<http://www.mas.ncl.ac.uk/~ndw/bluebook.htm>) or personal contact. A questionnaire was sent to either a single statistician or a group working within each medical school in February 2003. If they were unable to help, they were asked to forward the questionnaire to the Director of Postgraduate Studies.

The first author led a discussion workshop at the 2003 annual meeting of 'Teachers of Medical Statistics' at Burwalls in April. Prior to seeing the results of the survey, the audience was split into five groups with four to five people per group. Each group was asked to consider the following questions relating to the statistical teaching of PhD students in a Faculty of Medicine: Which students should be taught? When should they be taught? Who should teach them? What should they be taught? How should they be taught? Should attendance be compulsory? Should the course be assessed? The groups discussed these issues during a 20-minute period and were asked to report back during a feedback session.

3. RESULTS

3.1 SURVEY

Questionnaire replies were received from 21 (81%) of 26 UK medical schools. Of these, statistics teaching was provided specifically for PhD students in 13, under review in three, and not provided specifically for PhD students in five although they could apply to join Masters modules in two. Table 1 describes the content and delivery in those 13 schools where such teaching is provided.

The shortest courses lasting six hours covered data description, hypothesis testing and some elements of study design. This contrasted with the longest course of 50 hours which included these topics plus ANOVA, correlation and regression, multiple regression, logistic regression, survival analysis, and laboratory statistics. The content with respect to analysis was classified as introductory level when it included only topics from the following: data description, estimation, hypothesis testing, ANOVA, correlation, and simple linear regression. The classification was intermediate when topics covered included the above plus some or all of the following: multiple regression, logistic regression, survival analysis, multivariate analysis. The methods used for formal assessment in four institutions were examination (3), written assignment involving data analysis (3) and protocol development (1), and critical appraisal (1).

The question 'Is there any one-to-one statistical consultancy provided for PhD students?' was also asked. In three cases, the answer was not known. Of the remaining 18 schools, there was provision in 15. Of the 15 schools where consultancy was offered, free advice was available to all Faculty PhD students in 10, to those students able to access an advisory service funded by the National Health Service in three, and to students in departments employing a statistician in one. In one school, consultancy was available but had to be paid for.

Overall, there was some level of statistical teaching and consultancy in 9 schools, teaching alone in 3 and consultancy alone in 5. Of the remaining four schools responding, consultancy was provided but the teaching was under review in one. Whilst it was not known whether consultancy was provided in the other three, teaching was provided in one and was under review in two. Thus in 19/21, some level of statistical teaching and/or consultancy was provided.

Table 1. Survey results

Question: responses	Number of schools
Specialty of person responsible for teaching:	
Statistician	11
Statistician plus molecular biologist	1
Not known	1
Year of PhD in which teaching provided:	
First	7
Any	5
Not known	1
Mode of delivery:	
Lectures only	1
Computing sessions only (STATA)	1
Lectures plus tutorials	1
Lectures/computing or lectures/tutorials/computing (MTB 2, SPSS 4, MTB/SPSS 1, STATSDIRECT 1)	8
Not known	2
Total hours: 6,6,6,7,10,10.5,13,16,19,20,30,50, not known in one	
Brief content of teaching:	
Analysis (introduction)	4
Analysis (intermediate)	1
Analysis (intro) plus design	1
Analysis (inter) plus design	3
Design plus analysis (intro)	1
Design plus analysis (inter)	1
Not known	2
Formal assessment:	
Yes	4
No	7
Not known	2
Compulsory attendance:	
Yes	4
No	7
Not known	2
Funding provision:	
Yes	4
No	7
Not known	2

3.2 BURWALLS WORKSHOP

The audience of the Burwalls meeting were all university medical statistics teachers. In total 27 statisticians attended the meeting comprising a mix of approximately 40% senior academic (professor, reader, senior lecturer) and 60% academic staff (lecturer, research fellow, research associate). The

majority of statisticians were from Departments of Public Health, Centres for Medical Statistics or Mathematics and Statistics Departments at their various institutions.

The audience varied in terms of their experience of teaching medical statistics. However all groups were consistent in identifying the need to teach general statistical principles of design and analysis, and computing. Four groups felt that all PhD students should be taught whilst the fifth group thought that course attendance should be by negotiation. All agreed that the students ought to be taught early on in their first year. One group thought the students ought to learn when to involve a statistician and what to expect from a statistician.

All groups thought the teaching should be undertaken by a subject specialist, although two groups suggested the teaching could also be done by an experienced non-statistical researcher. No group identified a single best mode of delivery: suggestions included lectures, problem-based learning, Computer Aided Learning (CAL), workbooks, small group work, computing, and practicals.

Three of the five groups thought attendance ought to be compulsory, one at least for core subjects, another as part of a research training package. Four groups thought the teaching should be assessed, although proposed methods of assessment varied from self assessment to examination to a mini-project within the PhD area of the student.

Several further comments of interest were noted during the discussion. Firstly, teaching may create a demand for individual advice that cannot be met later. Secondly, joint PhD supervision by a statistician is a good idea. This was agreed in principle but several people commented about whether the workload would be realistic. Finally, the remark was made that it would be beneficial if supervisors also attended the courses.

3.3 PROPOSED TEACHING

At Liverpool University, teaching in statistics for PhD students within the Faculty of Medicine is being provided for the first time in the academic year 2003/2004 by the newly established Centre for Medical Statistics and Health Evaluation. The students are predominantly non-medical health scientists. Funding has been provided for a new statistics lectureship within the Centre with specific responsibility for this teaching. In practice, the teaching will be undertaken by the four statistics lecturers within the Centre. Table 2 describes the course programme in statistics.

Table 2. Details of the PhD course programme in statistics

Course title	Maximum number of students	Credits
PhD courses and attendance for November 2003:		
Statistical issues in the design and analysis of research projects	30	25
Design and analysis of randomised controlled trials	20	5
Design and analysis of diagnostic test and method comparison studies	20	5
Design and analysis of laboratory-based studies	20	5
Proposed new courses:		
Introduction to logistic regression	15	7
Introduction to longitudinal data analysis	15	7
Survival analysis	15	7

The teaching schedule consists of four courses, a week-long introductory course entitled '*Statistical issues in the design and analysis of research projects*', and three specialised one-day courses entitled '*Design and analysis of randomised controlled trials*', '*Design and analysis of diagnostic test and method comparison studies*', and '*Design and analysis of laboratory-based studies*', for which prior knowledge of the content of the introductory course is assumed. A higher number of places are allocated to the introductory course to ensure that we meet demand as this is viewed as the core course. Places on the specialist courses are more restrictive to allow more specific discussion and interaction with the students about their particular projects during the teaching sessions. In the proposed new courses where computing practicals are to be introduced then places are restricted again to allow for closer interaction and computer support.

PhD students are required to gain 90 credits in compulsory and optional university training courses, as suggested by their supervisors and heads of department. The statistics courses are classified as optional training. They run twice a year in November and April to allow for student registration at different times of the year, and priority is given to first-year students. In the academic year 2002/2003 approximately 150 first year PhD students were registered in the Faculty.

The courses emphasise aspects of both design and analysis. The introductory course consists of six teaching sessions, three of which focus on basic statistical methods: variability and types of data; populations, sampling and confidence intervals; hypothesis testing (parametric and non-parametric), and two of which give an overview of simple linear regression, correlation and analysis of variance; and multiple regression. One additional session is devoted to an overview of study designs commonly used in health research together with design concepts such as sample size and bias. The practical work, which is designed as self-learning material, involves an SPSS analysis in which a basic two-group comparison is performed to construct a baseline characteristic table for an observational study. The students must identify appropriate summary measures, construct confidence intervals and perform relevant hypothesis tests. Lectures for this course are given on Monday, Wednesday and Friday mornings and afternoons, with self-learning on Tuesday and Thursday. The teaching sessions are organised as two-hour lectures, sometimes including a video, with an extra half hour 'question-time' where the students can discuss the practical work or their own work in the context of what they have just learned.

The one-day courses focus on specific subject areas. They take the same teaching format as morning sessions with self-learning using the STEPS programmes (STEPS TLTP consortium 1992-95) in the afternoon. There are no formal assessments for any of the courses. Additional resources, including recommended text books and CAL resources, are listed on the centre website and university intranet. Within the next five years it is envisaged that course materials will be made available on the university intranet using the new university VITAL (Virtual Interactive Teaching at Liverpool) learning management system and Blackboard virtual learning environment software (Blackboard Inc. 1998), to facilitate administration and to enable overseas students to have on-line access to the course, in line with university policy. The VITAL system is a very flexible system allowing not only storage of materials and resources but also messaging, on-line assessment, and evaluation.

At the end of each course students are requested to complete a course evaluation form asking how useful they found the course. Student feedback from the first delivery of the courses is shown in Table 3 and indicated that the courses were well received. Constructive comments on how to improve the courses are also requested on the evaluation forms. Comments made by more than one student included the following. For the '*Statistical issues in the design and analysis of research projects*' course, six students requested more time to be spent on correlation and regression and multiple regression, and two would have liked less mathematical formulae, which were in fact kept to a minimum. On the '*Design and analysis of RCTs*' course two students requested that more time be spent on sample size. All the students attending the '*Design and analysis of diagnostic test studies*' found the course very useful and those on the '*Design and analysis of laboratory-based studies*' course requested examples in SPSS.

Table 3. Student feedback on usefulness of course

PhD Statistics Course	How useful did you find the course?				
	Not at all	Not much	Adequate	Fairly useful	Very useful
Statistical issues in the design and analysis of research projects:					
Session 1: Variability & types of data		2	2	5	8
Session 2: Confidence intervals			3	7	7
Session 3: Hypothesis Testing		1	2	7	7
Session 4: Overview of study design		1	3	7	7
Session 5: Correlation & regression	1		4	7	6
Session 6: Multiple regression	1	3	2	7	5
SPSS self-learning practical			1	11	6
Design and analysis of Randomised Controlled Trials			1	2	3
Design and analysis of diagnostic test studies					5
Design and analysis of laboratory-based studies				4	6

4. DISCUSSION

The majority of UK medical schools responding to the survey provide some level of statistical teaching to PhD students, usually by a statistician. However the total contact hours was highly variable, from six to 50. Rather surprisingly, design issues were covered in only half of the courses. The five medical schools not responding to the survey all offer a PhD programme. The results presented here may reflect the best scenario in the sense that non-respondents may be schools where statistical teaching and/or consultancy is less likely to be provided.

A further unexpected result was the high rate of consultancy provision, since this is usually perceived to be more resource-intensive than teaching and there is generally thought to be a shortage of medical statisticians to undertake this. It is unclear whether the decision to provide either courses or consultancy has been made on the basis of evidence of effectiveness or on the basis of resources available.

Our own teaching plan has been devised as a minimum but crucial package for PhD students. Due to limited resources we decided to provide teaching rather than consultancy, but to allow discussion and question time at the end of each session. Further courses that are being proposed are listed in Table 2. They are being introduced as a result of student feedback requesting more specific training in multiple regression techniques, and will include ‘hands-on’ computing sessions. In addition, we promote joint supervision of PhD students by a statistical supervisor where appropriate, in keeping with recommendations made in the MRC guidelines for studentship funding (MRC, 2003, Section 3.1 Quality assurance – standards of student supervision), which state that:

In some scientific areas e.g. Health Services Research, there may be grounds for joint supervision of individual students by, for example, both “basic” (e.g. psychologist, statistician) and “clinical” scientists (e.g. Professor of Nursing), to ensure that training spans both the methodology and its application in the field.

Our paper has been concerned with statistical education for PhD students in UK medical schools. Whilst we are aware of some courses in other countries (Harraway & Sharples, 2001), we have not undertaken a systematic search for non-UK reports. In particular, we are not aware of any recent survey of medical schools in other countries. Such a comparison would be interesting but would require further research.

ACKNOWLEDGEMENTS

The authors would like to thank all those who responded to the survey and the participants in the workshop at Burwalls. We also thank Dr Graham Kemp, Director of Postgraduate Research, Faculty of Medicine, University of Liverpool, for helpful discussions around this topic.

REFERENCES

- Bangdiwala, S.I., Amarillo, M.L., Ughade, S., Rodriguez, M.N., Singer, J.M., Munoz-Navarro, S.R., Komoltri, C., & Cumsille-Garib, J.F. (2002). Teaching consultancy through direct experience in research: the approach of the Department of Biostatistics at the University of North Carolina at Chapel Hill. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town: International Statistical Institute. CDROM.
- Blackboard Inc. 1998, Blackboard Learning System (release 6), Washington. [Online: <http://www.blackboard.com>]
- Belli, G. (2001). The teaching/learning process in university statistical consulting labs in the United States. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 325-338). Voorburg, The Netherlands: International Association for Statistical Education.
- Godino, J.D., Batanero, C., & Jaimez, R.G. (2001). The statistical consultancy workshop as a pedagogical tool. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 339-353). Voorburg, The Netherlands: International Association for Statistical Education.
- Harraway, J.A., & Sharples, J. (2001). A first course in biostatistics for health sciences students. *International Journal of Mathematical Education in Science and Technology*, 32(6), 873-886.
- Jolliffe, F. (2001). Learning from experience. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 355-370). Voorburg, The Netherlands: International Association for Statistical Education.
- Lancaster, G.A. (2002). Problem based learning. Presentation made at Burwalls Meeting of 'Teachers of Medical Statistics' in Bristol.
- Medical Research Council (2002). Postgraduate studentships handbook 2002-2003, Section 5.2. London: MRC. [Online: http://www.mrc.ac.uk/index/funding/funding-personal_awards/funding-studentship/pdf-handbook02-03.pdf-link]
- Medical Research Council (2003). Assessment criteria for MRC studentships, Section 3.1. [Online: http://www.mrc.ac.uk/index/funding/funding-personal_awards/funding-studentship/funding-assesment_criteria.htm]
- Mji, A., & Glencross, M. (2002). Statistical consultancy in a South African university: using a research resource centre to empower social science researchers. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town: International Statistical Institute. CDROM.
- Ospina, D. & Ortiz, J.E. (2001). Statistical research and consulting in universities from developing countries: the Colombian case. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 295-310). Voorburg, The Netherlands: International Association for Statistical Education.
- Shia, B.C. (2001). How to think about statistical consultation? Learning from data. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 371-377). Voorburg, The Netherlands: International Association for Statistical Education.
- Smeeton, N. (2002). Undergraduate courses in dental statistics in Britain and Ireland. *Statistics Education Research Journal*, 1(2), 45-48.
- STEPS TLTP consortium (1992-95). STEPS (Statistical Education through Problem Solving), Glasgow University. [Online: <http://www.stats.gla.ac.uk/steps>]
- Svensson, E. (2001). Important considerations for optimal communication between statisticians and medical researchers in consulting, teaching and collaborative research, with a focus on the

analysis of ordered categorical data. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 23-35). Voorburg, The Netherlands: International Association for Statistical Education.

PAULA R. WILLIAMSON

Centre for Medical Statistics and Health Evaluation

Faculty of Medicine

University of Liverpool

L69 3BX

FORTHCOMING IASE CONFERENCES

1. IASE 2004 RESEARCH ROUND TABLE ON CURRICULAR DEVELOPMENT IN STATISTICS EDUCATION

Lund, Sweden, June 28 - July 3, 2004

The International Association for Statistical Education (IASE) and the International Statistical Institute (ISI) are organizing the 2004 Roundtable on Curricular Development in Statistics Education, which will be held at Lund Institute of Technology at Lund University in Lund, Sweden from 28 June to 3 July 2004. The Roundtable will bring together a small number of experts, representing as many different countries as possible, to discuss one another's views and approaches to curriculum for teaching statistics. The Roundtable Conference will provide opportunities for developing better mutual understanding of common problems and for making recommendations concerning the statistics curriculum. A main outcome of the Roundtable will be a monograph containing a set of papers, which have been prepared for and discussed during the conference. The monograph will present a global overview of the conference that can serve as starting point for further research on issues related to the statistics curriculum.

The need for processing the increasing amount of data people receive in the course of their work and lives has made it imperative that students leave elementary and secondary schools prepared to make reasoned decisions based on sound statistical thinking. Countries and communities have approached this problem in different ways. The Roundtable will provide the opportunity for sharing what works and to highlight the challenges and potential solutions researchers and teachers have faced as they design and implement curricula to produce statistically literate citizens. The Roundtable will be held immediately prior to the Tenth International Congress on Mathematical Education to be held in Copenhagen, Denmark in 2004, July 4-11.

The IASE Scientific Program Committee will prepare the program and schedule for the Roundtable. The Committee has agreed on a list of topics that will form the basis of the discussions and invites those interested to send in a three-page summary of their proposed paper. The major topics to be addressed at the primary, secondary, tertiary, or inservice levels are: Relationship between curriculum and assessment; Role of research in shaping curriculum; Impact of technology on the statistic and probability curriculum; Innovative curricular practices; Teacher preparation and Statistical literacy.

- Theoretical papers should include; a) the statement of the problem, b) background or appropriate previous work, c) discussion of main arguments, d) implications for curricular development, e) references.
- Descriptions of experimental research should include; a) the statement of the problem b) background or appropriate previous work; c) methodology, data analysis and discussion of main results; d) implications for curricular development; e) references.
- Descriptions of curriculum innovations should include; a) focus and philosophy of the curriculum, b) background and development process, c) description, d) pilot and implementation results, e) sources and references.

Lena Zetterqvist (lena@maths.lth.se) and Ulla Holt will be local organisers. More information from Gail Burrill (Division of Science and Mathematics Education, College of Natural Science, Michigan State University), (burrill@msu.edu).

Web site: http://hobbes.lite.msu.edu/~IASE_2004_Roundtable/

2. STATISTICAL ACTIVITIES AT THE 10TH INTERNATIONAL CONGRESS ON MATHEMATICS EDUCATION Copenhagen, Denmark, July 4-11, 2004

As a part of the 10th International Congress on Mathematical Education to be held in Copenhagen, Denmark July 4-11, there will be a set of sessions set aside in TSG-11 to address issues related to research and development in the teaching and learning of probability and statistics, as well as the series of Regular Lectures in Statistics Education.

The ICME-10 venue will be the Technical University of Denmark, located in a northern suburb of Copenhagen. Chair International Programme Committee: Mogens Niss (ICME10-IPC@ruc.dk). Chair Local Organising Committee Morten Blomhøj (ICME10-LOC@ruc.dk). Web page: <http://www.icme-10.dk/>

2.1. TSG-11 RESEARCH AND DEVELOPMENT IN THE TEACHING AND LEARNING OF PROBABILITY AND STATISTICS

Team Chairs: Li, Jun, Department of Mathematics, East China Normal University, China; Joseph M. Wisenbaker, Educational Psychology, University of Georgia, USA

Team Members: Dani Ben-Zvi, Faculty of Education, University of Haifa, Israel; Manfred Borovcnik, Mathematics, Economics and Informatics, University of Klagenfurt, Austria; Maxine Pfannkuch, Mathematics Education Unit, Department of Mathematics, The University of Auckland, New Zealand.

Teachers at all levels find that teaching statistics and probability is immensely challenging. Not only are there new developments in and approaches to the subject matter, but there are constantly opportunities afforded by access to new instructional materials and methods and more advanced educational technology. At the same time, the difficulties that students have in learning statistics and probability pose major difficulties to teachers. While developments in statistical software and hand calculators have eliminated much of the computational burdens associated with applying statistics and probability, the difficulties posed by the basic worldview inherent in those subjects are just as challenging as ever. This session will address some of these issues by concentrating on research and developments in the teaching and learning of probability and statistics. Our TSG is intended to represent the diversity of the work being done with students of all ages and contexts, and that advances our knowledge of the possibilities and challenges facing us as educators. Papers and supporting materials for all presentations and papers chosen for presentation by distribution are available on the TSG-11 website at <http://www.icme-organisers.dk/tsg11/>.

2.2. ICME-10 REGULAR LECTURES IN STATISTICS EDUCATION

The following three Regular Lectures in Statistics Education will be delivered at ICME-10.

Rolf Biehler, University of Kassel, Germany, *Variation, co-variation, and statistical group comparison. Some results from epistemological and empirical research on technology supported statistics education.*

Margarida Cesar, Universidade de Lisboa, Lisboa, Portugal, *Come away with me: Statistics learning through collaborative work.*

Jane M. Watson, University of Tasmania, Hobart, Tasmania, Australia, *Assessment in Statistics education: Obstacle or Opportunity?*

**3. IASE ACTIVITIES AT THE 55TH SESSION OF THE INTERNATIONAL STATISTICAL
INSTITUTE
Sydney, Australia, April 5-12, 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. The sessions approved for ISI 55 in Sydney that were sponsored or co-sponsored by IASE are as follows (titles may change slightly). More information from Chris Wild at c.wild@auckland.ac.nz.

- Reasoning about Variation.
- The use of Simulation in Statistics Education
- Teaching Statistics Online
- Statistics for Life: What are the Statistical Ideas or Skills that Matter most and why?
- Research in Statistical Education
- Teaching Bayesian Statistics
- Challenges in the Teaching of Survey Sampling
- Using History of Statistics to Enhance the Teaching of Statistics
- Promotion of Statistical Literacy among Students
- Quality Assurance in Statistics Education
- Educating the Media on how best to Report Statistics
- Ethical Standards in Statistics Education

**4. IASE SATELLITE CONFERENCE - STATISTICS EDUCATION AND THE
COMMUNICATION OF STATISTICS
Sydney, Australia, April, 2005**

This conference focused on Statistics Education and the Communication of Statistics is jointly organised by the IASE and the Victorian Branch of the Statistical Society of Australia and will immediately precede the International Statistical Institute Session in Sydney. The approach will be non-technical, suitable for both a specialist and non-specialist audience who would like to learn how to better communicate the statistical ideas which occur in their everyday and working lives. This meeting is intended to be of interest to a wide cross section of society including teachers, educational administrators, researchers in statistical education and in probabilistic reasoning and others who want to gain a better grasp of how to communicate statistics in general and who would like to broaden their knowledge of statistics applications. It should also be of interest to people concerned with interpreting sociological, economical, political, scientific or educational reports, predicting sports results, by policy makers, journalists, health professionals and others from the general population. More information from Brian Phillips, bphillips@swin.edu.au. Web site:

<http://www.stat.auckland.ac.nz/~iase/conferences.php?show=iase2005>

**5. SRTL-4 THE FOURTH INTERNATIONAL RESEARCH FORUM ON STATISTICAL
REASONING, THINKING AND LITERACY
Auckland, New Zealand, April 2-7, 2005**

The Fourth International Research Forum on Statistical Reasoning, Thinking, and Literacy, to be hosted by the Department of Statistics, The University of Auckland, New Zealand, July 2-7, 2005. 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. Having emerged from the three previous forums, the topic and focus of SRTL-4 will be *Reasoning about Distribution*. The Forum is co-chaired by Dani Ben-Zvi (University of

Haifa, Israel) and Joan Garfield (University of Minnesota, USA), co-organized by Maxine Pfannkuch and Chris Wild (The University of Auckland, New Zealand), and planned by a prestigious international advisory committee.

Based on the SRTL tradition, we plan to keep the number of participants small to facilitate a working research forum. There are three possible roles for participants in this Forum. The first role is to present current research on reasoning about distribution, the second is to discuss and react to research presentations, while the third is to be a small group moderator, which is ideal for doctoral students who are not yet ready to present research but want to participate. Participants will be strongly encouraged to use videotape and written transcripts of students in classroom and interview settings to provide illustrations of what the researchers are learning about how students reason about distribution. As with the previous SRTL Research Forums, we encourage the participation of young promising scholars. One outcome of the Forum will be a publication summarizing the work presented, discussions conducted, and issues emerging from this gathering.

The SRTL-4 Research Forum organizers invite anyone interested in participating in this forum to contact them as soon as possible. The first deadline for submission of interest is June 1, 2004. More Information from Maxine Pfannkuch, m.pfannkuch@auckland.ac.nz.

Web site: <http://www.stat.auckland.ac.nz/srtl4/>.

6. ICOTS-7: WORKING COOPERATIVELY IN STATISTICS EDUCATION Salvador (Bahia), Brazil, July 2-7, 2006



The International Association for Statistical Education (IASE) and the International Statistical Institute (ISI) are organizing the Seventh International Conference on Teaching Statistics (ICOTS-7) which will be hosted by the Brazilian Statistical Association (ABE) in Salvador (Bahia), Brazil, July 2-7, 2006.

The major aim of ICOTS-7 is to provide the opportunity for people from around the world who are involved in statistics education to exchange ideas and experiences, to discuss the latest developments in teaching statistics and to expand their network of statistical educators. The conference theme emphasises the idea of *cooperation*, which is natural and beneficial for those involved in the different aspects of statistics education at all levels. Some examples are given below.

- *Cooperative learning in statistics education.* Recent trends in educational psychology emphasise the role of student activity and social interaction in learning. These developments are particularly important in the case of statistics where students are taking a more active role in working on cooperative projects and studies.
- *Cooperation between statistics teachers and researchers.* Real life applications generated by working with a researcher in another area help motivate the teaching of statistics. The subject is more enjoyable for students when a teacher can call on such real applications. At the same time, teachers are an essential part of a research team in statistics education, since they collaborate both in collecting data from the students and in helping with the design and evaluation of action-research programs.
- *Cooperation between statistical agencies and statistics educators.* Statistical agencies need the cooperation of the population at large when collecting their data. They are also interested in improving the statistical literacy of their citizens. Consequently, the agencies are communicating statistical ideas to their populace as well as providing official data for research on different topics, including teaching. Statistical offices and educators collaborate in the development of teaching resources based on official data and set up workshops and conferences on the teaching of statistics.

- *Interdisciplinary cooperation for research.* Interdisciplinary research is natural both in applied statistics and statistics education. Many central statistical concepts and procedures arose from research problems in other subjects. At the same time the researcher, whatever subject he or she is working in, benefits by having problems actually solved. Statistics education is based on many different disciplines, such as psychology, pedagogy, statistics and sociology, which all contribute in their own unique way to the study and solution of teaching problems.
- *International cooperation in statistics education.* Global communication and increasing interest and respect for complementarity in education are leading to an increasing number of successful international research or educational programs at different levels: e.g., Large scale statistical literacy comparative studies; Regional, National or International funded projects; International statistical education centres; International training programs or conferences in statistics education.
- *Globalization and diversity in statistics education.* Cooperation requires both global and local approaches to research and teaching. There is a contrast and a complementarity of global and local approaches in statistics education; e.g., large sample, quantitative studies versus qualitative and ethnographic research; the need to recognise global tendencies, and at the same time being sensitive to specific difficulties or talents of special and gifted students, minorities, etc.

The Conference will include keynote speakers, invited speakers, contributed papers, workshops and forums, demonstration lessons, roundtable sessions, poster sessions, book and software displays. People interested in organising a session or in presenting a paper are encouraged to contact the appropriate Topic Convenor. More information is available from the ICOTS7 Web page at <http://www.maths.otago.ac.nz/icots7/> and from Carmen Batanero, batanero@ugr.es.

6.1. TOPICS AND TOPIC CONVENORS

Topic 1. *Working cooperatively in statistics education.* Lisbeth Cordani, lisbeth@maua.br and Mike Shaughnessy, mike@math.pdx.edu

Topic 2. *Statistics Education at the School Level.* Dani Ben-Zvi, benzvi@univ.haifa.ac.il and Lionel Pereira, lpereira@nie.edu.sg

Topic 3. *Statistics Education at the Post Secondary Level.* Martha Aliaga, martha@amstat.org and Elisabeth Svensson, elisabeth.svensson@esi.oru.se

Topic 4. *Statistics Education/Training and the Workplace.* Pedro Silva, pedrosilva@ibge.gov.br and Pilar Martín, pilar.guzman@uam.es

Topic 5. *Statistics Education and the Wider Society.* Brian Phillips, BPhillips@groupwise.swin.edu.au and Phillips Boland, Philip.J.Boland@ucd.ie

Topic 6. *Research in Statistics Education.* Chris Reading, creading@metz.une.edu.au and Maxine Pfannkuch, pfannkuc@scitec.auckland.ac.nz

Topic 7. *Technology in Statistics Education.* Andrej Blejec, andrej.blejec@uni-lj.si and Cliff Konold, konold@srri.umass.edu

Topic 8. *Other Determinants and Developments in Statistics Education.* Theodore Chadjipadelis, chadji@polsci.auth.gr and Beverley Carlson, bcarlson@eclac.cl

Topic 9. *An International Perspective on Statistics Education.* Delia North, delian@icon.co.za and Ana Silvia Haedo, haedo@qb.fcen.uba.ar

Topic 10. *Contributed Papers.* Joachim Engel, Engel_Joachim@ph-ludwigsburg.de and Alan McLean, alan.mclean@buseco.monash.edu.au

Topic 11. *Posters.* Celi Espasandín López, celilopes@directnet.com.br

6.2. LOCAL ORGANISERS

Pedro Alberto Morettin, (Chair; pam@ime.usp.br), Lisbeth K. Cordani (lisbeth@maua.br), Clélia Maria C. Toloí (clelia@ime.usp.br), Wilton de Oliveira Bussab (bussab@fgvsp.br), Pedro Silva (pedrosilva@ibge.gov.br).

6.3. IPC EXECUTIVE

Carmen Batanero (Chair, batanero@ugr.es), Susan Starkings (Programme Chair, starkisa@vax.sbu.ac.uk), Allan Rossman and Beth Chance (Editors of Proceedings; arossman@calpoly.edu; bchance@calpoly.edu), John Harraway (Scientific Secretary: jharraway@maths.otago.ac.nz), Lisbeth Cordani (Local organisers representative; lisbeth@maua.br).

OTHER FORTHCOMING CONFERENCES

1. THE 33RD CARNEGIE SYMPOSIUM ON COGNITION “THINKING WITH DATA” Pittsburgh, USA, June 4-6, 2004

In today's information-rich environment, people use data in their daily lives in a variety of ways and for a variety of reasons. School children learn to create graphical representations of data to help solve problems; scientists collect and analyze data to address research questions, and consumers evaluate presented and experienced data to make everyday decisions. Although these different populations and tasks have been studied by different researchers (even by different fields), the basic, underlying question is the same: How do people think with data? The 33rd Carnegie Symposium on Cognition “Thinking with Data” will be held from June 4 to June 6, 2004, at the Carnegie Mellon University, Pittsburgh, PA, USA. More information from Marsha Lovett: lovett@cmu.edu. Web page: <http://www.psy.cmu.edu/thinkingwithdata/>

2. HAWAII INTERNATIONAL CONFERENCE ON STATISTICS, MATHEMATICS AND RELATED FIELDS Honolulu, Hawaii, USA, June 9-12, 2004

The 2004 Hawaii International Conference on Statistics, Mathematics and Related Fields will be held from June 9 to June 12, 2004 at the Sheraton Waikiki Hotel in Honolulu, Hawaii. The conference will provide many opportunities for academicians and professionals from statistics and/or mathematics related fields to interact with members inside and outside their own particular disciplines. Cross-disciplinary submissions with other fields are welcome. More information from statistics@bogus.example.com. Web page: <http://www.hicstatistics.org>

3. MATHEMATICS EDUCATION INTO THE 21ST CENTURY Ciechocinek, Poland, June 26-July 1, 2004

The next Mathematics Education into the 21st Century Project conference is to be held in Ciechocinek, Poland from June 26 to July 1, 2004. Future conferences will be held in Egypt, Jordan, Poland, Australia and Sicily. This conference theme is *The Future of Mathematics Education*. Papers are invited on all innovative aspects of mathematics education. More information from Alan Rogerson, arogerson@vsg.edu.au

4. MERGA-27 Townsville, Queensland, Australia, June 27-30, 2004

The Mathematics Education Research Group of Australasia Incorporated (MERGA) is an association that aims to promote, share, disseminate, and co-operate on quality research on mathematics education for all levels particularly in Australasia. It also aims to provide permanent means for sharing of research results and concerns among all members through regular publications and conferences. The Annual Conference of the Mathematics Education Research Group of Australasia is to be held at Townsville, Queensland, Australia, June 27-30, 2004. The conference theme is *Mathematics Education for the Third Millennium: Towards 2010*. More information from Ian Putt, Ian.Putt@jcu.edu.au. Web page: <http://www.merga.net.au/index.html>.

**5. ALM-11
Sweden, June 29-July 2, 2004**

The Adults Learning Mathematics group is an international research forum bringing together researchers and practitioners in adult mathematics/numeracy teaching and learning in order to promote the learning of mathematics by adults. The 11th International Conference on Adults Learning Mathematics is to be held in Sweden from June 29 to July 2, 2004. More information from Anette Strandberg, alm11@alm-online.org. Web page: <http://www.alm-online.org/>

**6. HISTORY AND PEDAGOGY OF MATHEMATICS 2004
SATELLITE CONFERENCE OF ICME-10
Uppsala, Sweden, July 12-17, 2004**

The HPM satellite conferences have taken place every 4 years since 1984. They provide a unique occasion to attend lectures, workshops, research reports from all over the world about the use of history in mathematics education and history of mathematics. The participants in the HPM meetings are researchers in history and in mathematics education, and teachers who have experimented with the use of history in their teaching. HPM 2004 is the sixth quadrennial meeting of the International Study Group on the Relations between History and Pedagogy of Mathematics (HPM). The study group is devoted to understanding and promoting the use of history of mathematics in teaching. Besides, it is the fourth European Summer University on History and Epistemology in Mathematics Education, which is a movement to bring together teachers from many countries to develop their knowledge and share their experiences of history and epistemology in mathematics education. The conference HPM 2004 will be held in Uppsala, from Jul 12 to July 17, in the week following ICME 10 in Copenhagen. The purpose of the conference is to bring together mathematics teachers, educational researchers and historians of mathematics to a meeting where they will share their insights and experiences of using the history of mathematics in teaching. More information from Fulvia Furinghetti, furinghe@dima.unige.it. Web page: <http://www-conference.slu.se/hpm/welcome/>

**7. PME-28
Bergen, Norway, July 14-18, 2004**

The 28th Annual Meeting of the International Group for the Psychology of Mathematics Education, will take place in Bergen University College, July 14 to July 18 2004. The theme of the conference is "Inclusion and Diversity". The conference will highlight a vision of mathematics for all. This perspective includes all people and a variety of mathematics as relevant for their use in their different needs. Challenges would be to make mathematics available and relevant; to support personal use of mathematics, and to organize for and stimulate people's learning. The following Research Forums are being organized for PME28:

RF1: Affect in Mathematics Education – Exploring Theoretical Frameworks. Co-ordinators: Markku Hannula, Jeff Evans, Rosetta Zan and George Philippou

RF2: Algebraic Equations and Inequalities: Issues for Research and Teaching. Co-ordinators: Lucianna Bazzini and Pessia Tsamir

RF3: An International Perspective on the Nature of Mathematical Knowledge for Secondary Teaching: Progress and Dilemmas. Co-ordinators: Helen M. Doerr and Terry Wood

RF 4: Contrasting Comparative Research on Teaching and Learning in Mathematics. Co-ordinators: Jonas Emanuelsson and David Clarke

RF 5: Researching Mathematics Education in Multilingual Contexts: Theory, Methodology and Teaching Mathematics. Co-ordinators: Richard Barwell and Philip Clarkson

Conference Secretariat Bergen University College Faculty of Education Landåssvingen 15 N-5096 Bergen, Norway. Email: secretariat@pme28.org. More information from Marit Johnsen Høines mjh@pme28.org. Web page: <http://www.pme28.org>.

**8. RELME-18 LATIN AMERICAN MEETING OF MATHEMATICS EDUCATION
Tuxtla Gutiérrez, México, July 19-23, 2004**

This series of conferences are organised by the Latin American Committee for Mathematics Education CLAME, which was created in the X Centro American and Caribbean Meeting of Training Teachers and Research in Mathematics Education in Puerto Rico, 1996. The aim was to consolidate Mathematics Education in the Region, favouring the teaching and learning of mathematics at all the educational levels. RELME-18 will be held in the city of Tuxtla Gutiérrez, Chiapas (México). July 19-23, 2004. More information from Miguel Solís Esquinca, solise@unach.mx, or the CLAME President Rosa M. Farfán, rfarfan@mail.cinvestav.mx.

Web page: <http://www.clame.org.mx/relme18/primera.html>

**9. JOINT STATISTICAL MEETINGS
Toronto, Canada, August 8-12, 2004**

This meeting is to be held at the Metro Toronto Convention Centre and Royal York Hotel, Toronto, Canada, August 8-12. It is sponsored by the American Statistical Association, ENAR, WNAR, IMS, and SSC. More information from American Statistical Association, meetings@amstat.org. Web page: <http://www.amstat.org/meetings/jsm/2004/>.

**10. PME-NA 26
Canada, October 21-24, 2004**

PME came into existence at the Third International Congress on Mathematical Education (ICME-3) in Karlsruhe, Germany in 1976. It is affiliated with the International Commission for Mathematical Instruction. The twenty-sixth conference Annual Conference of the North American Chapter of the International Group for the Psychology of Mathematics Education (PME-NA) is to be held in Toronto, Ontario in October 21 to 24, 2004. The theme of the conference is *Building Connections Between Communities*. We especially encourage proposals that highlight examples of the interplay between research and practice—practice that has been shaped by research and research that grows out of practice. The conference is sponsored by the Ontario Institute for Studies in Education at the University of Toronto. The conference provides a forum for the exchange of research information on the psychology of mathematics education. More information from Doug McDougall, Dmcdougall@oise.utoronto.ca. Web page: <http://www.pmena.org/2004/>

**11. ASIAN TECHNOLOGY CONFERENCE IN MATHEMATICS 2004
Singapore, December 13-17, 2004**

There is little doubt that technology has made an impact on the teaching and Mathematics. The Asian Technology Conference in Mathematics on the theme *Technology in Mathematics: Engaging Learners, Empowering Teachers and Enabling Research* is hosted by the National Institute of Education, Nanyang Technological University of Singapore (December, 13-17, 2004). In this conference, we shall go beyond justifying the use of technology in Mathematics to discuss and examine the best practices of applying technology in the teaching and learning of Mathematics and in Mathematics research. In particular, the conference will focus on how technology can be exploited to enrich and enhance Mathematics learning, teaching and research at all levels. The conference will cover a broad range of topics on the application and use of technology in Mathematics research and teaching. More information from Wei-Chi YANG, wyang@radford.edu, and Tilak de Alwis, tdealwis@selu.edu. Web page: <http://www.atcminc.com/mConferences/ATCM04/>

12. EPISTEME – 1
Goa, India, December 13-17, 2004

Over the last thirty years science, technology and mathematics education have emerged as lively new research areas. This research, inspired by issues of learning and teaching, has clear uniting themes in the cognitive, pedagogical, historical, philosophical and socio-cultural aspects of the sciences. An international conference to review research on Science, Technology and Mathematics Education is to be held at the International Centre Dona Paula, Goa, India, in December 13-17, 2004. The conference will survey the global progress of research in this field and will aim to identify promising directions for future work. More information from Jayashree Ramadas, episteme@hbcse.tifr.res.in. Web page: <http://www.hbcse.tifr.res.in/episteme>.

13. CERME-4
Girona, Spain, February 17-21, 2005

The Fourth Congress of the European Society for Research in Mathematics Education (ERME) will be held in Platja d'Aro, Girona, Spain, from 17 to 21 February, 2005. The conference will focus mainly on work in Thematic Groups in a style similar to that developed in previous conferences. Details of the groups from CERME3 can be found on the website <http://www.dm.unipi.it/~didattica/CERME3/second.html>. Many of the previous groups will continue work, and we expect a few new groups. CERME4 will also include plenary activities and poster presentations. More information from the President of ERME: Paolo Boero, boero@dima.unige.it, the Chair of the CERME4 Programme Committee: Barbara Jaworski, barbara.jaworski@hia.no, or the Chair of the CERME4 Organising Committee: Marianna Bosch, mbosch@fundemi.com.

**14. ICMI STUDY 15, THE PROFESSIONAL EDUCATION AND DEVELOPMENT OF
 TEACHERS OF MATHEMATICS**
São Paulo, Brazil, May 15-21, 2005

The ICMI Studies are working conferences dealing with specific themes in mathematics education, meaning that a substantial amount of the time is dedicated to discussions, although paper and other presentations also take place. The focus of this Study will be the professional education and development of mathematics teachers around the world. The premise of this Study is that the education and continued development of teachers is key to students' opportunities to learn mathematics. The curriculum of mathematics teacher preparation varies around the world, both because of different cultures and educational environments, and because assumptions about teachers' learning vary. Countries differ also in the educational, social, economic, geographic, and political problems they face, as well as in the resources available to solve these problems. A study focused on mathematics teacher education practice and policy around the world can provide insights useful to examining and strengthening all systems. A Study Volume will be produced, representing and reporting selected activities and results of the Study Conference and its products. This Report will be useful to the mathematics education community, as well as for other researchers, practitioners, and policymakers concerned with the professional education of teachers. Location is Águas de Lindóia, São Paulo, Brazil. More information from the Study Co-Chairs: Deborah Loewenberg Ball (USA), dball@umich.edu, and Ruhama Even, ruhama.even@weizmann.ac.il (Israel). The Discussion Document and other general information is available from the Web page: <http://www-personal.umich.edu/~dball/icmistudy15.html>

15. ICTMA 12
London, United Kingdom, July 10-14, 2005

Mathematical modelling and applications, the transition, freely between real world problems and mathematical representations of such problems, is an enduring and important feature of industry, business and commerce. Teaching mathematical modelling, through tasks, projects investigations and applications embedded in courses and of mathematics itself through applications helps learners to understand relationships between real world problems and mathematical models. The 12th International Conference on Mathematical Modelling and Applications (ICTMA12) will be hosted by the School of Engineering and Mathematical Sciences City of London, Sir John Cass Business School, London, UK. This conference brings together international experts in a variety of fields and from different sectors to consider: modelling in business and industry, evaluating effectiveness, pedagogic issues for learning, applicability at different levels, research: education and practice, innovative practices and transitions to expert practice. More information from Chris Haines, ictma12@city.ac.uk.

Web site: www.city.ac.uk/conted/reseach/ictma12/index.htm

16. PME-29
Melbourne, Australia, July 10-15, 2005

The PME 29 conference will be held on July 10-15, 2005 in Melbourne, Australia. The First Announcement will be available in September 2004. More information from Helen Chick, h.chick@unimelb.edu.au.