Newsletter of the International Study Group for Research on Learning Probability and Statistics

Volume 4, Number 2

May 1991

Joan Garfield, Secretary and Editor
340 Appleby Hall
128 Pleasant St. S.E.
Minneapolis, MN 55455 USA

E-Mail: pqa6031@ca.acs.umn.edu or pqa6031@umnacca.bitnet
Fax: (612) 626-7848

Notes and Comments

This is the first attempt at a "theme" newsletter. The theme is assessment, and I will appreciate any comments you have on the newsletter and any suggestions regarding future themes. But first, some news and updates. We have two new members to the study group.Ted Hodgson at the University of Indiana is currently working on his doctoral dissertation in mathematics education which centers on the learning of procedures and the misconceptions that result when procedures are mislearned. He is also interested in the means by which students establish connections between representational systems, especially in relation to the learning of formal concepts within probability. He would like to establish contact with people who are working in these areas. His address is:

Ted Hodgson, Department of Mathematics Education
School of Education, #309, W.W. Wright Education Bldg, 3rd and Jordan,
Bloomington, Indiana, 47405, USA.

Our other new member is Marcelo Borba, who is a graduate student in mathematics education at Cornell University, in Ithaca, NY. Marcelo is interested in research on averages: balance beams, weighted and arithmetic means, median and mode. He is interested in establishing contact with people doing work in this area. His e-mail address is y2q@cornella.bitnet and his phone is 607-277-1728.

I still have several sets of the collected research papers from ICOTS III held in New Zealand last summer. Please let me know if you want to purchase copies of this collection. So far, I have only had a few orders, although several of you originally expressed interest in purchasing the collected papers.

Recent Papers of Interest

I have been sent copies of the following papers from study group members:

"Task variables in statistical problem solving using computers" by J. Diaz, M.C. Batanero Bernabeu, and A. Estepa Castro. This paper was presented at the NATO Advanced Research Workshop on Information Technologies and Problem Solving Research, held in April in Portugal.

Abstract:
In this work an analysis of some task variables of statistical problems which can be proposed to the students to be solved on the computer, are presented. The objective of this didactical-mathematical analysis is to pride criteria of selection of the said problems, directed to guiding the student's learning towards the adequate meanings of the statistical notions and to the development of their ability to solve problems.

The probability and statistics research group at the Scientific Reasoning Research Institute (SRRI) at the University of Massachusetts, Amherst have put together a list of 16 papers they have written. Anyone who would like a copy of this list should write to the SRRI , Hasbrouck Laboratory, University of Massachusetts, Amherst, MA 01003. I have written a summary of the group's current research projects for the fall issue of *Teaching Statistics* and would be happy to send a copy to anyone interested. Two of the most recent papers from this group are:

"Understanding Students' Beliefs about Probability" by Clifford Konold, will appear in E. von Glasersfeld (Ed.) *Constructivism in mathematics education*, to be published by Reidel Publishers.

"The Psychology of Learning Probability" by Ruma Falk and Clifford Konold, to appear in F.S. Gordon and S. P. Gordon (Eds.). *Statistics for the twenty-first century*, to be published by the Mathematical Association of America.

I received this paper from Heinz Steinbring:

"Mathematics in teaching processes - the disparity between teachers and student knowledge", published in RDM, Vol 10, no. 3. 1990.

Abstract:
When analyzing episodes of mathematics instruction from an epistemological perspective, it is seen that the disparity between teachers and student knowledge is not simply due to their knowing more or their knowing less. The independent and frequently incompatible levels of understanding knowledge which are peculiar to teachers and to students show how essential it is to make allowance for *conceptual* as opposed to *material* aspects, and how the conditions of classroom processes nevertheless always tend to regress to a form of mathematical knowledge strongly determined by subject matter and

method. (Note: the teaching episode used in this paper is from a 6th form lesson on probability.)

Andee Rubin sent me a draft copy of her paper presented at the American Educational Research Association (AERA) in Chicago: "TapeMeasure: Video as Data for Statistical Investigations". This paper reports how a videotape-based data system was developed and used to involve students in posing and solving real world statistical problems.

I presented a paper at AERA entitled "Reforming the undergraduate statistics course." This paper showed how the current reform movement in mathematics education (in the US) and the calls for reform of statistical education at the college level (to keep up with changes in the practice of statistics) lead to a new type of undergraduate statistics course. A model of such a course is offered.

Papers presented at AERA and NCTM

Other AERA papers of interest are:

"Statistics and Computer Attitudes and Achievement of Students Enrolled in Applied Statistics: Effect of a Computer Laboratory" by Patricia B. Elmore and Ernest L. Lewis at Southern Illinois University at Carbondale.

"Critical Concepts in Data Modeling" by Chris Hancock, at Technical Education Research Centers (TERC) in Cambridge, MA.

"Average and Data: Examining Mathematical Relationships" by Janice Mokros and Susan Jo Russell at TERC.

Susan Jo presented another paper at NCTM in April:
"Students' understanding of average: research results and implications."

Another paper presented at NCTM was:
"Children's understanding of averages n everyday contexts" by Karen J. Rothschild, Iddo Gal, and Daniel Wagner, all at the University of Pennsylvania.

I recently received a copy of a paper by Julian Weissglass and Deborah Cummings, titled "Dynamic Visual Experiments with Random Phenomena." This paper describes one approach to using visualization in learning statistics. The approach utilizes the ability of microcomputers to perform statistical experiments and display the results dynamically as they occur.

And now, for our theme: **Assessment**

Based on a few papers sent to me by members, Flavia Jolliffe's ICOTS III paper on assessment, and my own review of the related literature, I have put together the following brief review. Part of this review will be included in a paper I will be presenting at PME next fall in Virginia, which describes the assessment project I am currently working on as part of Cliff Konold's Chance-Plus project.

Assessment in Classrooms

Historically, most classroom testing has focused on students' computational skills and few tests have measured higher order thinking. Statistics items that appear in traditional tests typically test students' ability to correctly calculate the mean and median for a set of numbers, or to construct or interpret a graph of a data set. Tests composed of these items not only test skills in isolation of a problem context, they do not test whether or not students understand how these statistical measures are interpreted or know when one is a better summary measure to use than another. They also fail to assess students' ability to integrate statistical knowledge to solve a problem and their ability to communicate using the statistical language.

Because traditional assessment methods communicate little about the quality or substance of students' learning, and because the type of learning actually measured seems artificial or meaningless when compared to the actual use of that learning in the real world, there is currently a movement to reform classroom assessment. The NCTM standards also call for a new vision of assessment and the development of alternative assessment methods capable of measuring students' higher order thinking and reasoning skills. The Standards describe alternative forms of assessment that will provide information on how well students can communicate mathematics, understand mathematics as an interrelated set of ideas, and how well students are able to gain meaning from a particular situation (Webb & Romberg, 1988).

In interpreting the NCTM standards for assessment, Webb and Romberg (1988) provide criteria for new assessment instruments. Such instruments should:

1. Provide information that will help make decisions for the improvement of instruction.
2. Be aligned with instructional goals.
3. Provide information on what students know.
4. Be able to supplement other assessment results to provide a global description of what mathematics students know.

An alternative to traditional assessment that is currently being discussed is the idea of "authentic" assessment, a method of obtaining information about students' understanding in a context that reflects real-life situations, and challenges students to use what they have learned in class in an authentic context (Archbald & Newmann, 1988). Other alternative methods for assessing mathematics learning are described in a booklet prepared by the EQUALS staff and the Assessment Committee of the California Mathematics Council (Stenmark, 1989).

The Need for Common Assessment Methods

If the traditional tests do not determine how well students are understanding basic probability and statistics concepts and how well they are able to solve problems, then new assessment instruments and methods are indeed needed. The need for new instruments is particularly apparent when new curriculum and software are developed for teaching probability and statistics and need methods to assess the impact of these materials on student learning. For example, evaluations of the Quantitative Literacy Project and Reasoning Under Uncertainly Project required the

development of new tests and methods. However, the tests and assessment methods constructed reflect the specific program content rather than more general conceptual understanding and higher order reasoning skills.

In his forthcoming chapter on research on learning probability and statistics, Shaughnessy (in press) describes the need for some standard, reliable tools to assess students' conceptions of probability and statistics. If fact, he lists the development of assessment instruments as the first item on his "wish list" for future research in this area. Although a variety of items and tasks have been used by researchers or evaluators in the past, it is difficult assemble these items and tasks into a test because of their different purposes and formats (e.g., paper and pencil, clinical interview). Shaughnessy stresses the need for new instruments which incorporate and build on the ideas of previous research but which have greater applicability. Ideally, these instruments will help us determine if the new standards for learning probability and statistics (as described in the NCTM report) are being achieved.

Research Related to Assessing Statistical Understanding

A review of research by Jolliffe (1990) organizes the relevant literature in the categories of classification schemes for assessment tasks, newer methods of assessment, attitude scales, and studies of understanding. After my own literature review, I have modified these categories into five groups of studies relevant to the assessment of statistical understanding. These are: 1) students' attitudes and anxiety towards learning statistics, 2) students' computational skills in using probability and statistics, 3) students' misconceptions of probability and statistics, 4) conceptual frameworks for assessing statistical learning, and 5)methods of assessing mathematical learning and problem solving. Each is described below.

Student Attitudes and Anxiety. The Statistics Attitude Survey (SAS) scale (Roberts & Saxe, 1982), the Attitudes Toward Statistics test (ATS) (Wise, 1985), and the Statistical Anxiety Rating Scale (STARS) (Cruise, Cash, & Bolton, 1985) are Likert-type scales written for college students in statistics courses. The SAS was designed to assess various components of statistical attitudes, such as students' perceptions of their own statistical competence and the usefulness of statistical analysis. The ATS was developed specifically to measure attitude changes during statistics courses and is designed to be given as a pre- and post-test. Two scores are calculated: attitudes towards the course and attitudes toward the field of statistics. The STARS measures students' attitudes towards six areas: worth of statistics, interpretation anxiety, test/class anxiety, computation self-concept, fear of asking for help, and fear of statistics teachers. None of the three instruments assess student understanding of and beliefs about what the field of statistics is, what it means to "do" statistics and solve statistical problems. Instead, they deal with the more specialized attitudes and anxiety faced by college students.

Students' Computational Skills. Tests written to accompany commercial textbooks are the most common form of assessment for measuring students' ability to perform statistical calculations. Items on standardized tests and the National Assessment of Educational Progress (NAEP) tend to be of this type. One example from the NAEP asks students to calculate the mean, median and mode for a set of data consisting of inches of snowfall (Brown & Silver, 1989). Although these items typically test whether or not students can use formulas and come up with a single, correct answer, they do not assess whether or not students understand the concepts and can use them to analyze and interpret data. For example, students may be able to correctly calculate the median and mean but not know when one is a better average to use than another. This type of skill is best assessed in classes through assignments and quizzes, and does not need to be on a general test of statistical understanding and reasoning.

Students' Conceptions and Misconceptions about Probability and Statistics. The only test written and used on a large scale to assess students understanding of probability concepts was developed and administered by Green (1983) to 3000 students in Great Britain. Some of these items and other items appearing in the research literature have been found to be useful in detecting misconceptions and helping researchers to understand how student think about probability and statistics. See for example the work of Falk, delMas, Garfield, Jolliffe, Konold, Pollatsek, Shaughnessy, and Well. Research reviews by Garfield and Ahlgren (1989) and Shaughnessy (in press) refer to many of these studies. Items used are often open-ended and many have been used in clinical interviews to probe students' beliefs. Many have been used with adults or college students and involve a substantial amount of reading. Although these items are good at detecting student conceptions and misconceptions, many need to be revised and adapted for high school students.

Frameworks for developing assessment tasks. There have been at least two attempts to design frameworks for developing tasks for assessing statistical learning. Chervaney, et al., (1977) used a model of the problem solving process to develop a three stage model of assessment (comprehension, planning and execution, and evaluation and interpretation). These three stages contain 10 different steps in statistical reasoning which can be used to guide item development. Although this framework was designed to evaluate innovative college courses and was successfully used to design tests for a college level course (Garfield, 1981) it does not appear to have been used in other studies. Nitko and Lane (1990) also designed a framework for generating assessment tasks that provide a richer description of students' thinking and reasoning than just giving them problems to work out. This framework was developed for college and graduate level statistics courses and can be used to assess relationships among knowledge and whether or not important principles and concepts are understood by students. Three interrelated categories are used to classify statistical activities: problem solving, modeling, and statistical argument. Although developed for students at a level higher than secondary school, these models are useful in providing frameworks for organizing statistical knowledge and skills.

Assessment of mathematical learning and problem solving. There are 13 standards for evaluation included in the NCTM curriculum standards. These standards describe the assessment of students' mathematical knowledge, conceptual understanding, procedural knowledge, problem solving, reasoning, and mathematical disposition. Assessment is viewed as the process of understanding the meaning which students give to mathematics; it should be dynamic and involve a variety of approaches (Webb & Romberg, 1988). Recently, more attention has been given to assessment of higher order mathematical thinking (Kulm, 1990). Educators are encouraged to move away from using single number summaries to represent

students' knowledge, and using two dimensional frameworks for developing assessment measures, to instead explore alternative models of assessment and ways of building on more recent models of learning mathematics.

Current Developments in Statistical Assessment

The Statistical Reasoning Assessment currently being developed as part of the Chance-Plus project is designed to assess students' beliefs about statistics, their understanding of basic concepts of probability and statistics, and their ability to use these concepts in interpreting information, reasoning, and solving problems. After reflecting on the the research literature reviewed, previous tests and test items, teaching experience, and much group discussion, the ChancePlus project team of psychologists, educators, and statisticians outlined a framework of important beliefs, ideas, concepts, and reasoning skills. Components of the instrument were then collected, revised, or written from scratch to assess these ideas and skills. Although some parts of the instrument look like traditional test items, others appear unique in their format and ability to capture students' thinking and reasoning. Four parts of the test were created to be used at various times and in various combinations:

Part 1 assesses general beliefs about the nature of statistics and statistical work. Two formats are used for these items. One part contains statements about statistics and statistical work (e.g., there may be more than one way to correctly solve a statistical problem). The other part asks students to rate the skills that someone would need to have in order to analyze and interpret data (e.g., types of communication and mathematical skills).

Parts 2 and 3 assess general ideas about probability and statistics students would have before a course of instruction. These items do not use specialized vocabulary with which a student might be unfamiliar. Items are designed to see how students interpret information and make judgements about different situations. Items are also designed to assess students' intuitions and misconceptions about probability and statistics that interact with or are resistant to instruction.

Part 4 assesses students' ability to reason about and solve probability and statistics problems. In order to develop realistic contexts for solving these problems, a research study conducted by a high school class is described. All questions are based on the analysis of this project. One version of this test describes results from a survey of how students spend their money. Questions involve a decision, interpretation, or conclusion about some aspect of the data analysis.

Another project I am working on is the use of a structured statistical project as one type of authentic assessment. This is a structured project that students can complete in a reasonable length of time that allows students to demonstrate their ability to use statistical methods and ideas and communicate their understanding and reasoning about these methods and ideas. I have developed a scoring method to use for evaluating these projects that is easy for teachers to use, gives students feedback to improve their learning, and guides the teacher in improving instruction and modifying activities. I have been unable to locate any literature on evaluating student statistical projects and would appreciate information on this topic. If anyone is interested in more information on my project I will be happy to share a copy of a draft paper.

Here are some of the references cited in my literature review:

Archbald, D. & Newmann, F. (1988). Beyond standardized testing: Assessing authentic academic achievement in the secondary school. Reston, VA: National Association of Secondary School Principals.

Brown, C. and Silver, E. (1989). Data organization and Interpretation. In M. Lindquist (Ed.). Results from the fourth mathematics assessment of the National Assessment of Educational Progress (pp. 28-34). Reston, VA: National Council of Teachers of Mathematics.

Chervaney, N., Collier, R., Fienberg, S., & Johnson, P. (1977) A framework for the development of measurement instruments for evaluating the introductory statistics course. American Statistician , 31 (1), 17-23.

Cruise, R. Cash, R., and Bolton, D. (1985). The development and validation of an instrument to measure statistical anxiety. Paper presented at the annual meeting of the American Statistical Association, Las Vegas, Nevada.

Garfield, J. (1981) In investigation of factors influencing student attainment of statistical competence. (Doctoral dissertation, University of Minnesota, 1981). Dissertation Abstracts International, 42, 5064A.

Garfield, J. and Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: implications for research. Journal for Research in Mathematics Education. 19, 44-63.

Green, D. (1983). A survey of probability concepts in 3000 pupils aged 11-16 years. In D. Grey, P. Holmes, V. Barnett, & G. Constable (Eds.). Proceedings of the first international conference on teaching statistics (pp. 766-783. Sheffield, UK: Teaching Statistics Trust.

Jolliffe, F. (1990). Assessment of understanding of statistical concepts. Paper presented at the Third International Conference on Teaching Statistics, Dunedin, New Zealand.

Kulm, G. (1990). Assessing Higher Order Thinking in Mathematics. Washington, DC: American Association for the Advancement of Science.

Nitko, A. & Lane, S. (1990) Solving problems is not enough: assessing and diagnosing the ways in which students organize statistical concepts. Paper presented at the Third International Conference on Teaching Statistics, Dunedin, NZ.

Roberts, D. & Saxe, J. (1982). Validity of a statistics attitude survey: a follow-up study.Educational and Psychological Measurement, 42, 907-912.

Shaughnessy, J. M. (In press). Research in probability and statistics: reflections and directions. in D. Grouws (Ed.), Research on Teaching and Learning Mathematics.

Stenmark, J. K. (1989). Assessment Alternatives in Mathematics: An overview of assessment techniques that promote learning. Berkeley: Lawrence Hall of Science, University of California.

Webb, N. & Romberg, T. (1988). Implications of the NCTM standards for mathematics assessment. Paper presented at the AERA annual meeting, New Orleans.

Wise, S. (1985). The development and validation of a scale measuring attitudes toward statistics. Educational and Psychological Measurement. 45, 41-405.

Well, that's all I have so far on assessment. If you know of other relevant papers please let me know.

Have a good summer, and be sure to send me papers, comments, and suggestions for our next newsletter, due in your mailboxes next September.