CAN AUTHENTIC ASSESSMENT HELP IN DELIVERING COMPETENT CONSUMERS OF STATISTICS FOR NON-ACADEMIC PROFESSIONS?

SCHUYTEN, Gilberte & FERLA, Johan
Ghent University
Belgium

Since quantitative research skills become more and more relevant for non academic professions, the four courses research curriculum at the department of educational sciences of Ghent University aims to deliver competent consumers of statistics who possess quantitative research skills and attitudes needed to produce and use research in their professional careers.This study focuses on the impact of authentic assessment with group project work on student self-efficacy beliefs and attitudes towards statistics. About 180 students, enrolled at the fourth course, are engaged in collaborative project work during 8 weeks on a given data-base. Students' perceptions of self-efficacy, attributions for academic success, assessment expectations and attitude towards quantitative research as a field and as a course are measured after the presentations of their projects. The control group consists of students enrolled at the third course.

SCIENTIST-PRACTITIONER MODEL

An important goal of education graduate programs is to deliver graduates who are interested in research as well as practice and who are able to integrate science and practice into a body of skills, knowledge, and professional attitudes.

In counselling psychology, a growing body of literature indicates that graduate students and professionals lack interest in research and maintain low levels of research productivity (Bishop & Bieschke, 1998). Also Davis et al. (2006) discuss the research-practice gap as a dilemma facing many applied fields, from medicine to psychology to family students. These authors describe this often-cited gap as 'the tendency of practitioners (often clinicians) to ignore research findings and of scientists to study questions irrelevant to practice or to fail to disseminate findings in an accessible manner'. This behaviour of practitioners to ignore research findings is also occurring in professional education careers.

Gelso (1979) identified ingredients of the research training environment that he believed could be manipulated to enhance graduate students' interest in and positive attitudes toward research and also their eventual research productivity. Theoretical support for the importance of interest and positive attitude is given by Bandura's social cognitive theory (Bandura, 1986). Subsequent research has provided additional empirical support for Gelso's model. Phillips and Russell (1994), Kahn & Scott (1997) and Bishop and Bieschke (1998) highlighted the role of research self-efficacy and research interest.

It is clear that in order to close the research-practice gap quantitative research education programs should aim at two different types of study outcomes: (1) with respect to research skills, graduate students should be able to produce scholarly work, (2) with respect to beliefs and attitudes, desired outcomes for graduate students are positive judgments about their ability to complete quantitative research and positive attitude towards quantitative research in general.

COMPETENT CONSUMERS IN AN EVIDENCE-BASED SOCIETY

An evidence-based society is described by Smith (1996, p.367) as a society "in which informed quantitative reasoning is the dominant modality in public debate as well as in the decision-making processes of government, business and individuals". He also describes statistics as "the science of doing science, whose role is to provide theory and protocols to guide and discipline all forms of quantitative investigatory procedure" (Smith, 1996, p.368).Following Gal & Garfield (1997) the goal of statistics instruction is that students become informed citizens who are able to: (1) Comprehend and deal with uncertainty, variability, and statistical information in the world around them, and participate effectively in an information-laden society, (2) Contribute to or take part in the production, interpretation, and communication of data pertaining to problems they encounter in their professional life. Based on this broad vision, they discuss eight interrelated

basic subgoals for statistics instruction which are made up of two overlapping clusters, the 'doing' statistics cluster and a second cluster concerned with sense-making and communicative skills as well as with reflection and questioning. The 'doing statistics' cluster contains six subgoals, the second cluster consists of following three subgoals: Develop interpretive skills and statistical literacy (overlapping subgoal with the doing cluster), Develop ability to communicate statistically and Develop useful statistical dispositions.

This framework can be used for describing competences needed by competent consumers in professional education careers. The subgoals in the second cluster are particularly relevant for our target group when phrased in more demanding terms of 'actively engaged in using, producing and communicating quantitative evidence in their jobs'.

A graduate education student who possesses the competences and dispositions needed to get involved in quantitative research in his or her professional career is prepared to enter professional life as a competent consumer of statistics. These competences and dispositions are not only needed in an information and evidence-based society but are necessary to close the science-practice gap.

SELF-EFFICACY AND ATTITUDES

Self-efficacy

Academic self-efficacy beliefs refer to students' personal judgments about their capabilities to organize and execute the courses of action required to attain designated types of educational performances (Zimmerman, 1995). Ferla and Valcke (submitted) mention that a strong sense of self-efficacy induces students to undertake challenging tasks (Bandura & Schunk, 1981), to extend greater effort in accomplishing a given task (Salomon, 1984; Schunk, 1983), to persist longer in the face of adversaries (Bandura & Schunk, 1981; Schunk, 1984), to self-regulate the learning process better (Zimmerman et al., 1992) and to use more effective cognitive and meta-cognitive strategies resulting in deeper processing of the learning material and a higher level of understanding (Pintrich & De Groot, 1990). Researchers agree that self-efficacy beliefs are strong determinants of learning and achievement, even after prior achievement and cognitive skills are taken into consideration (Pintrich & Schunk, 1996; Schuyten & Dekeyser, accepted).

Attitudes

Attitudes towards statistics and towards mathematics are a much researched topic. It is generally accepted that "attitudes and beliefs can impede (or assist) learning statistics, and may affect the extent to which students will develop useful statistical thinking skills and apply what they have learned outside the classroom" (Gal et al., 1997, p.37).

Although earlier research has given evidence that in introductory statistics courses in education programs affective variables have no impact on students' study outcomes in statistics (Schuyten et al., 1999), 'students' attitudes and beliefs regarding statistics deserve attention. Gal et al. (1997) give three types of reasons: (1) process considerations such as feel safe, feel comfortable, believe in their ability, motivated; (2) outcome considerations after students leave the classroom; and (3) access considerations to enrol in a statistics course later on.

All three types of reasons are particularly relevant in educational research programs aiming at the delivery of competent consumers.

AUTHENTIC ASSESSMENT WITH GROUP PROJECT WORK

Different goals require different assessment methods. Mastering of isolated specific research skills such as 'choosing the appropriate statistical technique for answering the research question' can be assessed by traditional item formats such as multiple choice or filling in short answers, but assessing competences with respect to the ability to integrate and coordinate knowledge, skills, and attitudes, and the capacity to apply them in new situations requires authentic assessment tasks. It is widely acknowledged that in order to meet the goals of education, a constructive alignment between instruction, learning and assessment is necessary (Biggs, 1996).

Increasing the authenticity of an assessment is expected to have a positive influence on student learning and motivation (Herrington & Herrington, 1998). Gielen et al (2003) give two important reasons for using authentic competency-based assessments: their construct validity and

their impact on student learning, also called consequential validity. Construct validity means that tasks must reflect the competency that needs to be assessed, should represent real-life problems and should require thinking processes that experts use to solve the problem in real life. With respect to consequential validity Gibbs (1992) emphasized that student learning is largely dependent on the assessment and on student perceptions of the assessment requirements.

Researchers have different descriptions of authentic assessment. We take the definition as given by Gulikers et al. (2004, p.69): authentic assessment is "an assessment requiring students to use the same competencies, or combinations of knowledge, skills, and attitudes that they need to apply in the criterion situation in professional life".

One of the favoured authentic assessment methods used in statistics and mathematics education is project work. Project work is an interesting assessment method in a quantitative research education program since students have to handle compound aptitude complexes (Corno and Snow, 1986), especially assembled for performance in these situations.

Project work is often associated with collaborative small group project work. From the social constructivist point of view learning can be considered as constructing knowledge, active, situated, and collaborative, i.e. meaning is negotiated from multiple perspectives (Merrill, 1991). Definitions of collaborative learning differ but following concepts tend to be important (Davies, 2006): learning together and building an emerging pool of knowledge, learning from each other, working in partnership, creating learning communities, shared responsibility for product or outcome, sharing information and opinions, negotiation of roles, methodology, task and assessment. Collaborative learning implies that students go through all learning processes together; as such focus is more on process than on product.

CONCEPTUAL FRAMEWORK, DESIGN AND RESEARCH HYPOTHESES

Conceptual framework and research design

Taking into account that

- an evidence-based society needs professionals in education who are competent consumers of statistics willing to become actively engaged in statistical investigations in their professional career;
- two types of study outcomes are important to deliver competent consumers of statistics: mastery of research skills (statistical knowledge, meta-cognitive skills) and positive dispositions and attitudes;
- social cognitive theory of Bandura (1986) stresses the importance of self-efficacy and interest in research (Bishop, Kahn & Scott, 1997; Lent et al. 1994);
- the integrated nature of interwoven cognitive and meta-cognitive skills and students' dispositions requires authentic assessment;
- theories about collaborative learning focus more on process than on product; and
- self-efficacy beliefs, perceived assessment demands and attribution of success are closely related and have an impact on students' study approach (Ferla & Valcke, submitted)

The central aim of this study is to provide empirical evidence about the impact of collaborative small group project work on two types of study outcomes (1) statistics achievement (study outcomes on practical exam) and (2) self-efficacy, attitude towards quantitative research as a field and a course.

This study is a comparative group survey design with two groups of education students: third year students (undergraduates) without project work in the course and fourth year students (graduates) with project work.

Research Hypotheses

1. Collaborative project work increases students' confidence in research skills, a positive attitude toward quantitative research and toward quantitative research courses.
2. Students with low prior knowledge of statistics gain more from project work in respect to study outcomes on the practical exam and in respect to self-efficacy and attitudes toward the field and the course than students with high prior knowledge of statistics.
3. Students' with low statistics achievement scores (practical exam) gain less from project work with respect to self-efficacy and attitudes than students' with high statistics achievement scores.
4. Self-efficacy, study outcomes on practical exam, attitude towards the field and attitude towards the course material are positively related.
5. Interrelations between self-efficacy, study outcomes on the practical exam, attitude toward field and attitude toward course are stronger for fourth year students than third years

METHOD

Participants and procedure

The research education curriculum of students in education consists of four consecutive courses: two statistics courses and two courses quantitative research in education (Models of Empirical Research 1 and 2). The three first courses are at undergraduate level, the fourth and last course is at graduate level. The third course is a second semester course, the fourth a first semester course.

The comparative group survey design consist of two groups of about 180 students each from three education disciplines (instructional sciences, special education and social work) enrolled at a Belgian (Flemish) university.

According to the ECTS the fourth course provides a study of methods used to conduct research in education. It is a 12 week course with about 20 hours lectures, 8 hours practicals and 8 hours group project work. Assessment methods are group presentation of the project and individual practical exam on computer with open book aligned with the tasks during the practicals. The third course is also a 12 week course with 30 hours lectures and 15 hours practicals (with pen and paper 6 hours an on computer 9 hours) and is assessed by an open book examination on paper with open answers on questions closely aligned with the tasks presented during the practicals. A practical introduces or reinforces a particular statistical technique and has limited objectives; the project links a number of topics and is investigative and open ended and integrates technical and non-technical skills/abilities.

The teaching-learning environment of both courses differs only in the added project work. Support for project work was given during regular meetings between the instructor and groups of 5 students and informal meetings with students. Assessment method used for the third course is an individual pen and paper exam (open book) whereas assessment methods in the fourth course consist of group presentation of project and individual practical exam on computer.

To participate in this study, students were required to give their consent by signing an informed consent document; no student declined to participate.

Fourth course students were surveyed in December 2006, after the group presentation of their project, with respect to self-efficacy in research skills and attitudes toward quantitative research as a field and as a course. The individual practical exam was taken next day. Third course students were surveyed in May 2007 after the last practical with respect to self-efficacy in research skills and attitudes toward quantitative research as a field and as a course.

Achievement scores of both courses are given by scores got by the open book examination on paper in the third course and by the individual practical exam on computer in the fourth course.

Instruments

Current Self-efficacy in research skills

The following three instruments in the literature were studied: the 33-items SERM Self-efficacy in research measure (scale of 0-9; Phillips & Russell, 1994), the 51-items RSES Research Self-efficacy scale (scale from 0 to 100; Greely et al., 1989; Bishop & Bieschke, 1998; Bieschke

et al. 1996) and the 14-items CSE Current Statistics Self-efficacy constructed for introductory statistics students (scale 1 to 6; Finney and Schraw, 2003)

As the first two scales were too long and not specific enough for our purpose of measuring more advanced students in self-efficacy with respect to quantitative research skills, we developed a 15 item scale. We adopted the definition of current self-efficacy as given by Finney and Schraw (2003, p.164) "confidence in one's ability to solve specific tasks related to statistics" and respondents are asked to rate their current confidence to complete 15 specific tasks related to quantitative research using a 1 (no confidence at all) to 6 (complete confidence) response scale.
We retained six CSE items which were suited for more advanced students. These are in our numbering 1, 2, 4, 5, 10 and 12. Following 9 items, aligned to the list of competences aimed at, were developed and added 'formulate research questions on a given data-base', 'identify the structure (causal, latent, direct-indirect effect, discrimination …) in research questions, 'construction of a scale given a number of items', 'analysis of an effect in direct and indirect effects', 'reading of the output of a statistical technique', 'interpret the output of a statistical technique', 'identify factors that might have an impact on statistical results', 'identify the best fit model' and 'present project work'.

Attitude towards data-analysis

Many instruments assessing attitudes towards statistics are available in literature and conference proceedings (Gal et al., 1997; Marquez, 2004). Earlier research by Schuyten et al. (1999) used a 14-items instrument based on items of the NLSMA Reports of the SMSG (1968) and on items of by Beenen en Van Der Werff (1976). This instrument has been used in Ghent since 1978 for capturing freshmen attitudes towards statistics.

In this study we decided to use items based on the 'Attitudes Toward statistics (ATS)' of Wise (1985). This ATS 29-item attitudinal scale is constructed for introductory statistics students and several items needed to be adapted to more advanced statistics students. The term statistics was replaced by data-analysis. The scale consists of two subscales: attitude toward the field of statistics (20 items, 1, 4, 6, 8(R), 11, 13(R), 14, 17, 18(R), 21(R), 23, 25, 26(R), 27, 28, 29, 30, 33, 35(R), 36) and attitude toward the course (9 items, 2(R), 5(R), 9(R), 10, 16(R), 20(R), 24(R), 32(R), 34(R)). Students were asked to respond to how they feel about a statement using a 1 (strongly disagree) to 5 (strongly agree) response scale. Wise reports a two-factor solution and high internal consistency (~.90) for both subscale scores. The distinctiveness of Wise's subscale scores was further supported by correlations with the course grade. Specifically, the attitude toward the course subscale was a positive predictor of course grade, unlike the attitude toward the field subscale, which had a null relationship with course grade.

Achievement in research skills

For third year students achievement scores are given by exam scores at the end of the third course. The exam is an open book pen-and-paper exam aligned with tasks of the practicals.

For fourth year students achievement scores are given by exam scores at the end of the fourth course. These scores consist of two parts: a score on an open book exam on computer and a group score on presentation of project work.

STATISTICAL ANALYSES

Data will be analyzed with different techniques such as: exploratory and confirmatory factoranalysis, analysis of variance and multiple regression.

RESULTS & DISCUSSION

Results and discussion will be presented at the satellite meeting in Guimarães.

REFERENCES

Bandura, A. (1986). *Social Foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41,* 586-598.

Beenen, F. en Van Der Werff, J.J. (1976). Statisticophobie. Een onderzoek naar de weerstanden tegen Statistiek. *Nederlands Tijdschrift Psychologie, 22,* 209-217

Bieschke, K.J., Bishop, R.M., & Garcia, V.L. (1996). The utility of the Research Self-Efficacy Scale. *Journal of Career Assessment, 4*, 59-75

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*, 347-364

Bishop, R.M., & Bieschke, K.J. (1998). Applying social cognitive theory to interest in research among counselling psychology graduate students: A path analysis. *Journal of Counseling Psychology, 45,* 182-188

Corno, L. and Snow, R.E. (1986). Adapting teaching to individual differences among learners. In M.C. Wittrock (ed.) *Handbook of Research in Teaching.* New York: Macmillan.

Davies, S. (2006, May 26). *Definition of collaborative learning*. Message posted to http://legacy.uwcm.ac.uk/ltsn/forum/1085047704/index_html

Davis, Jonathan C., Dan M. Sandifer-Stech (2006) Wade Into the Water: Preparing Students for Successful Quantitative Research. *Family Relations, 55*, 56-66

Ferla, J. and Valcke, M. (submitted).Relationships between student cognitions and their effect on students' learning and study approach

Finney, S. J., & Schraw, G. (2003). Self –efficacy beliefs in college statistics courses. *Contemporary educational Psychology, Volume 28,* 161-186

Gal, I., Garfield, J. (1997). 1. Curricular Goals and Assessment Challenges in Statistics Education. In Gal,I. & Garfield, J.B. (editors). *The Assessment Challenge in Statistics Education* (pp. 1-13) IOS Press, 1997 (on behalf of the ISI). Freely available on the Internet at http://www.stat.auckland.ac.nz/~iase/publications/assessbk

Gal, I., Ginsburg, L., Schau, C. (1997). 4. Monitoring Attitudes and beliefs in Statistics Education. In Gal, I. & Garfield, J.B. (editors). *The Assessment Challenge in Statistics Education* (pp. 37-51). IOS Press, 1997 (on behalf of the ISI). Freely available on the Internet at http://www.stat.auckland.ac.nz/~iase/publications/assessbk

Gelso, C. J. (1979). Research in counselling: Methodological and professional issues. *The Counseling Psychologist, 8,* 7-35.

Gibbs, G. (1992). *Improving the quality of student learning*. Bristol, UK: Technical and Educational Services.

Gielen, S., Dochy, F., & Dierick, S. (2003). The influence of assessment on learning. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of quality and standards* (pp.37-54). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Greely, A. T., Johnson, E., Seem, S., Braver, M., Dias, L., Evans, K., Kincade, E. & Pricken, P. (1989). *Research Self-Efficacy Scale.* Unpublished scale, Pennsylvania State University, University Park.

Gulikers, J. T. M., T. J. Bastiaens, P. A. Kirschner (2004). A Five-Dimensional Framework for Authentic Assessment. *ETR&D, 52, 3*, 67-86.

Herrington, J., & Herrington, A. (1998). Authentic assessment and multimedia: How university students respond to a model of authentic assessment. *Higher Educational Research & Development, 17(3),* 305-322.

Kahn, J. H. & Scott, N. A. (1997). Predictors of research productivity and science-related career goals among counselling psychology doctoral students. *The Counseling Psychologist, 25,* 38-67.

Lent, R. ., Brown, S. D., & Hackett, G. (1994). Towa a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior, 45,* 79-122.

Marquez, J. C. (2004). Una Revision de las Evidencias de Fiabilidad y Vadlidez de los Cuestionarios de Actitudes y Ansiedad Hacia la Estadistica, *Statistics Education Research*

*Journal,3, 1,*5-28 Freely available on the Internet at http://www.stat.auckland.ac.nz/~iase/serj/SERJ3(1)_marquez.pdf

Merrill, M. D. (1991). Constructivism and Instructional Design. *Educational Technology, 31,* 45-53

Phillips, J.C., & Russell, R.K. (1994). Research self-efficacy, the research training environment, and research productivity among graduate students in counselling psychology. *Counseling Psychologist, 22,* 628-641

Pintrich, P. R. & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology***,** 82, 33-40.

Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in education: Theory, research, and applications*. Englewood Cliffs, NJ: Prentice Hall.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions, *Journal of Educational Psychology, 76(4),* 647-658.

Schuyten, G., Dekeyser, H. and Goeminne, K. (1999). Towards an electronic independent learning environment for statistics in higher education. *Education and Information Technologies, 4, 4,* 409-424

Schuyten, G. & Dekeyser, H.M. (accepted). Preference for textual information and acting on support devices in multiple representations in a computer based learning environment for statistics. *Computers in Human Behavior*

Schunk, D. H. (1983). Developing children's self-efficacy and skills: The roles of social comparative information and goal setting. *Contemporary Educational Psychology, 8,* 76-86

Schunk, D. H. (1984). Enhancing self-efficacy and achievement through rewards and goals: Motivational and informational effects. *Educational Research, 78,* 29-34.

Smith, A.F.M. (1996) Mad cows and ecstacy: chance and choice in an evidence-based society, *Journal of the Royal Statistical Society,* 159 (3), pp. 367-383

SMSG (1968). School Mathematics Study Group. No 3. Z-Population Test Batteries. *NLSMA Reports.* Leland Stanford Junior University

Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement, Volume 45,* pp. 401-405

Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies (pp. 202-231)*. New York: Cambridge University Press.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29,* 663-676.