TEACHING STATISTICS AND RESEARCH METHODS: AN INTEGRATED APPROACH

Hans van Buuren
Open Universiteit Nederland, The Netherlands
hans.vanbuuren@ou.nl

Traditional curricula in the social sciences result in students having statistical knowledge that is inert and consequently of low transferability. This is in part because these curricula separate mathematical and probabilistic content (present in statistics service courses) from the context in which the collection of observational and experimental data is designed (present in courses about research methods). This paper proposes a curriculum that removes this separation by merging the two domains into the research competency, in line with emergent pedagogical insights. This study describes the new curriculum and compares some preliminary learning outcomes of students following the proposed integrated competency-based curriculum with that of students following the traditional curriculum. The results suggest a higher level of understanding is achieved through the integrated approach.

INTRODUCTION

Most statistics service courses in the social sciences today are facing a number of problems encompassing negative attitudes towards statistics, low motivation, and statistical knowledge that is inert and results in low transferability of the knowledge and skills (cf., Barab, Squire, and Dueber, 2000). One obvious cause seems to be the mathematical nature of statistics requiring abstract thinking of the students. Another cause seems to be the function of statistics in a social science curriculum: students do not see the necessity of statistics – in the way it is usually taught - as a tool for data-analysis in research processes. For this reason, Snee (1993) suggested that “the ‘content side’ of statistics education should move away from the mathematical and probabilistic approach and place greater emphasis on “data collection, understanding and modeling variation, graphical display of data, design of experiments, surveys, problem solving, and process improvement” (p. 151). Both Snee and Moore (1997) emphasize what they call ‘statistical thinking.’ According to Snee, the collection and analysis of data is at the heart of statistical thinking, because data collection promotes learning by experience and connects the learning process to reality.

The suggestions made by these researchers to focus on statistical thinking and, thus, on the processes involved in data production, encroach upon the domain of subjects like research methods. The challenge would be to merge these subjects and include them within, for example, an empirical research-based competency. This paper proposes a curriculum in which thinking about data, production of data, investigations, modeling, analysis, and interpretation are not taught within the boundaries of separate statistics service courses, but will be treated coherently with research methods.

This approach is in line with emergent pedagogical insights, which include a number of paradigms shifts in the educational field, notably the shift from behaviorism to cognitivism—in particular, constructivism (Pappert, 1993), and the shift from individual learning to collaborative learning (Johnson and Johnson, 1994). These shifts imply moves from teacher-centered learning to learner-centered learning and from ‘passive’ learning to ‘active’ learning. Educational researchers conjecture that active learning in collaborative groups of students increases individual learning performances, such as a deeper understanding of statistics (cf., Johnson, Johnson, and Stanne, 2000). If the new pedagogies are applied, then may deeper understanding be achieved, but low motivation problems and a negative attitude towards statistics may also be alleviated.

The present study aims to explore the way in which individual statistics courses can be integrated in a research competency curriculum, and what the effects are on statistics learning outcomes. The next section of this paper describes the design of such a curriculum at the School of Psychology of Open Universiteit Nederland. Then the paper sketches the settings of the research project and the characteristics of the students participating in it. The last two sections analyse the data and conclude this study.

EMBEDDING STATISTICS IN A COMPETENCY-BASED CURRICULUM

Though statistics is a part in the process of doing research in the social sciences, until now specialist statistics educators have taught statistics relatively separately from the other subjects in the research process. The same goes for the research methods courses too. The advent of advanced computer technology and the insights of constructivism have stimulated researchers to look for a solution that no longer splits up complex cognitive skills like doing research. A solution would be the ‘whole-task approach’ (see Van Merriënboer, 1997) which focuses on training the complete complex cognitive skill of doing research (from simple to complex versions). This approach emphasizes the coordination and integration of the constituent skills from the very beginning, and stresses that students should quickly develop a holistic view of the whole task that is gradually embellished and detailed during instruction. The holistic view implies that students have to be grasp the necessity to align the successive steps systematically. Students should be stimulated to develop awareness of the inextricable connectedness of the subsequent steps. So within the teaching design emphasis must be put on general methodological principles and theories, and students have to be stimulated to experience and learn the complete research process with all relations and connections between the theoretical, methodological, and statistical levels.

Constructivism holds that the acquisition of complex skills is context-dependent and occurs most effectively in relevant contexts (Lovett and Greenhouse, 2000). Hence, the learning environment must provide realistic situations where learning through meaningful practices takes place. By doing research in a great variety of situations, students are more likely assured of a better transfer of the academic subject (Van Merriënboer, 1997). Computer technology and timesaving computer programs make such an approach feasible.

As Snee already stated in 1993: the goal should be to integrate statistical thinking into the subject matter on which the students are working (Snee, 1993). So the challenge is to arouse the interest and personal commitment of students, and develop their appreciation for statistical thinking by embedding the statistical methodology into the subject matter. In 1998 we took up this challenge and made a start to integrating the psychological subject matter, research methods and statistics into a new research curriculum for psychologists at Open Universiteit Nederland. This university is an institution for distance education, which means that students have little or no opportunities to meet, work and learn together. Students get their learning materials by mail, or via the computer-supported collaborative learning environment (CSCL) of Open Universiteit Nederland.

In the redesign of the research curriculum, the books and their contents of the individual subjects are no longer central, but according to the constructivist paradigm, the focus is aligned to (research) assignments. To emphasize the difference from the earlier curriculum with its separate research methods and statistics courses, in the case of the integrated courses in the new curriculum we will speak about *competencies*. In this case we are concerned with the bachelor psychologist’s research competency that features statistical knowledge and skills as vital constituents.

A research competency-based curriculum could be designed to replace seven courses of each comprising 120 hours of study time: two statistics service courses, three service courses about research methods; one course on applications of the computer program SPSS; and one course on literature research. Including the bachelor thesis (240 hours), a total of 1080 hours could be rearranged and reallocated (see Table 1).

The rearrangement and reallocation of the service courses into the new curriculum consisting of 7 research practicals is specified in Table 2. This table shows that in order to make empirical investigations manageable the research process has been broken down according to the research stages and the rules of the American Psychological Association (APA) for journal articles (Introduction (‘Problem stage’), Methods, Results, Discussion, the last ‘stage’ is omitted in the table). The Methods and the Results columns of the table show the contents of the two subjects that were formerly taught separately: research methods and statistics respectively. Each practical contains at least three different research assignments. The first assignment is a worked-out example in order to reduce the ‘mental’ or cognitive load (Lovett and Greenhouse, 2000); the

Table 1: Scheme of the separate service courses concerning the research curriculum Bachelor of Psychology OUNL (Old curriculum)

No	Service course	Content	Examination
1	Statistics 1 (120h)	Descriptive statistics (up to bivariate association/regression/correlation)	Multiple choice: independent items
2	Statistics 2 (120 h)	Probability theory and test theory (up to ANOVA)	Multiple choice: independent items
3	Research Methods 1 (120 h)	Methodological concepts (survey) Questionnaire: construction; interviewing	Multiple choice: independent items; Interview protocols
4	Research Methods 2 (120 h)	Methodological concepts (experiment) Observation	Multiple choice: independent items; Observation protocols
5	Research Methods 3 (120 h)	Concepts of qualitative research, case study	Compose a research proposal for a case study
6	Literature study (120 h)	Searching recent articles about a subject in order to identify the state-of-the-art	Compose a research proposal on the basis of a literature search
7	SPSS (120 h)	Operating program on existing data-files (without statistical interpretation of the results!)	Report
8	Bachelor thesis (240 h)	Doing a survey based on an existing data-file	Paper, research report (APA)

Table 2: Scheme of the research competency curriculum Bachelor of Psychology OUNL (New curriculum)

Nr	Research practical	Problem stage	Methods stage	Results stage
1	Parametric data analysis (120 h)	4 cases Psychology	- data gathering methods are given in cases - path model 2 variables - secondary analysis (on existing data)	*Comparison*: Means/standard deviation t-test/ANOVA *Association*: linear regression/correlation
2	Survey (120 h)	4 cases Psychology	- path model ≥ 2 variables questionnaire construction (survey) and administration (n ≥ 20 respondents) Likert scale - Cronbach's α	linear regression/correlation + moderation / mediation / interaction
3	Literature search (120 h)	Problem Theoretical justification Hypotheses	Latest research articles about a psychological phenomenon	
4	Non-parametric data analysis (120 h)	4 cases Psychology	- (path)model 2 variables - secondary analysis (on existing data)	*Comparison* : Most commonly used non-param. test measures in Psychology *Association*: Mostcommonly used non-param. measures for association, logistic regression loglinear analysis
5	Experimental design (120 h)	4 cases Psychology Causal relationships	Experimental designs Confounding influences	t-test / ANOVA + repeated measures + MANOVA
6	Observation and interviewing (120 h)	4 cases Psychology	Observation, Content analysis Depth interview Focus groups	Parametric statistics Non-parametric statistics + Cohen's kappa
7	Bachelor thesis (480 h)	Theoretically justified problem	All treated methods	All treated measures

second a less worked-out assignment (fill-in) where students have to contribute more, and the final assignment is the task where students have to carry out their own research in accordance with the learning goals. Within each practical the intensity of the coaching decreases and the contribution of the individual student increases (scaffolding), according to the Four Component / Instructional Design (4C/ID) model (Van Merriënboer, 1997). Not all elements of the research process are emphasized equally in the practicals. In the first practical, called Parametric Data-Analysis, the Results stage is emphasized, which means that although examples and cases of complete investigations are given, most attention is given to the data analysis. In the second practical attention shifts to the second stage of the research process (Methods). Now we expect students to apply correctly the statistics knowledge and skills acquired (in this case regression and correlation), and to be competent in extending this ‘prior’ statistical knowledge and skills to several specific features (moderation, mediation). We should point out that in their bachelor’s education students now get acquainted with statistical techniques and aspects of statistical thinking in far more opportunities than had been the case when students received only their two service courses. Note that in the first and second practicals we limited the statistics to the mean, standard deviation and extensions of these to analysis of variance and regression/correlation. As a consequence students can get a better understanding of the mean and its characteristics, without confusing these with other (e.g., non-parametric) statistics. In addition, all the statistics material learned comes back in all later research assignments. Thus students discover the necessity, functionality and utility of statistics, through experiencing them, and become more motivated as a result of real data and task authenticity. Finally, their knowledge of statistics should be interiorized properly by repeatedly performing data analysis; their knowledge becomes transferable, and deep understanding should grow as a consequence of the use of a variety of (psychological) cases (Van Merriënboer, 1997; Lovett and Greenhouse, 2000).

It has taken several years to gain insights into the consequences of a curriculum redesign, to evaluate the definitive psychological research programme, and to select the appropriate delivery, sequence and content of statistical topics in every practical. During this period of tryouts students’ motivation, affects, uses of self-regulating learning strategies (deep learning or superficial learning and their components) and learning outcomes have been monitored. Based on the positive results in the quasi-experiments, Open Universiteit Nederland decided in 2003 to change the curriculum of statistics and research methods definitively. At this moment, however, the whole curriculum has not yet been redesigned. In September 2004 the last opportunity for students to complete the classic statistics service courses was discontinued and research practical 1 (Parametric data analysis) was officially introduced. Practicals 2, 3 and 5 were introduced in 2003. Practicals 4 and 6 will be realized in September 2006.

To align learning goals, instruction, and assessment, the examination in each practical comprises an (empirical) investigation. So statistical (and methodological) knowledge and skills will be tested indirectly.

In this paper we focus on the effects of the different curricula on learning outcomes: are student performances in statistical problem solving in an integrated curriculum design better than in the case of individual classic statistics courses? In this study, we only compare the very first courses (Statistics 1 and/or Statistics 2) of the old curriculum and the first practical of the new curriculum (Parametric Data Analysis).

METHODS

Participants

Data were collected from 468 psychology students in distance education, in two different samples: 340 students who participated in the non-integrated classic statistics service courses, and 128 students who have been confronted with the whole-task teaching design. In both samples approximately 5% of the participants were male and 95% female, which is an adequate reflection of the total population of Psychology students at the Dutch Open University.

Procedures

Students were invited to participate in the study via e-mail or a letter two weeks before the questionnaire was presented. Participation was voluntary, i.e., it did not carry any rewards. The questionnaire was offered via a closed website or by letter for participants without an e-mail address. After a month students received a reminder to stimulate them to submit the questionnaire.

Materials

All participants were asked to fill in an extensive questionnaire. For the statistics performance test, seven open questions and one closed question were developed (for examples see for instance Reading, 1996). The answers and arguments on the statistics questions were evaluated and classified in accordance with the Structure of Observed Learning Outcomes (SOLO) taxonomy (Biggs and Collis, 1982). The SOLO classification describes the growth and complexity of learning outcomes in a widely applicable language across a range of subjects at any level of education, from the earliest engagement in the task to expertise. The SOLO taxonomy provides for five different levels of understanding, from 0. prestructural (incompetency at the bottom of the hierarchy in which responses are unstructured and irrelevant); 1. unistructural (mentioning a single relevant feature); 2. multistructural (listing some relevant features); 3. relational (relating items in list) to 4. the deepest level, known as extended abstract understanding. A major strength of the taxonomy is its validated ability not only to identify how much has been learned, but also how well students link and integrate what has been learned (Hattie and Purdie, 1998).

Two independent raters, using statistical and methodological criteria developed by experts, expressed in SOLO levels, simultaneously coded 620 SOLO responses. The Cohen's Kappa for the SOLO tasks was between .90 and 1.0. For a detailed description of the analysis we refer to Van Buuren, Bijker and Wynants (2006, in press).

RESULTS

The SOLO scores ranged from level 0 till 4. The independent *t*-test revealed statistically significant differences between the means. The mean SOLO-score for the integrated design group (M = 1.65; SD = .485) was significantly higher (t = - 3.687; df = 265; p = .000, two-tailed) than the mean of the non-integrated design group (M = 1.46; SD = .343).

CONCLUSION

Although the comparison is still in its infancy – we could compare only a small part of the curricula – the results of the SOLO analysis are promising. It strengthens our idea that we are on the right track. Nevertheless the present study has its weaknesses. The two samples were not randomly chosen. We did not carry out randomized trials, meaning that possible confounding variables are not controlled for. On the other hand, the samples used in this study are a real-life reflection of innovative versus classic teaching designs, providing the possibility to measure complex processes in different real-life teaching contexts (Reeves, 2005). More studies, especially longitudinal studies, are needed to assess the development of understanding research and statistics within the groups, carried out with the cyclic, whole-task research based design.

This study provides support for the claim that integrated, research-based teaching designs can optimize learning outcomes in statistics. The new teaching context removes from statistics its perception as a discipline with a narrow, mathematically focused and anxiety-provoking perspective and nests it in a research-embedded context, related to the Psychology domain (Bijker *et al*., 2006, in press) that triggers student interest and commitment, and as a consequence, facilitates the study process.

REFERENCES

Barab, S. A., Squire, K. D., and Dueber, W. (2000). A co-evolutionary model for supporting the emergence of authenticity. *The Journal of Education Technology, Research, and Development*, 48(2), 37–62.

Biggs, J. B. and Collis, K. F. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy.* New York: Academic Press.

Bijker, M. M., Wynants, G., and Van Buuren, J. A. (2006). A comparative study of the effects of motivational and affective factors on studying statistics. In A. Rossman and B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. Voorburg: The Netherlands: International Statistical Institute.

Chan, C. C., Tsui, M. S., Chan, M. Y. C., and Hong, J. H. (2002). Applying the structure of the observed learning outcomes (solo) taxonomy on student's learning outcomes: An empirical study. *Assessment and Evaluation in Higher Education*, 27(6), 511-527.

Hattie, J. and Purdie, N. (1998). The SOLO model: Addressing fundamental measurement issues. In B. Dart and G. Boulton-Lewis (Eds.), *Teaching and Learning in Higher Education* (pp. 145-176). Melbourne: ACER.

Johnson, D. W. and Johnson, R. T. (1994). *Learning Together and Alone: Cooperative, Competitive, and Individualistic Learning* (4th edition). Needham Heights, MA: Allyn and Bacon.

Johnson, D. W., Johnson, R. T., and Stanne, M. B. (1985). *Cooperative Learning Methods: A Meta-analysis*. Minneapolis, MN: University of Minnesota.

Lovett, M. C. and Greenhouse, J. B. (2000). Applying cognitive theory to statistics instruction. *The American Statistician,* 54(3), 196-206.

Merriënboer, J. J. G. v. (Ed.). (1997). *Training Complex Cognitive Skills: A Four-Component/Instructional Design Model for Technical Training*. New Jersey: Englewood Cliffs.

Moore, D. S. (1999). Discussion: What shall we teach beginners? *International Statistical Review*, 67, 250–252.

Pappert, S. (1993). *The Children's Machine: Rethinking School in the Age of the Computer*. New York: Basic Books.

Reading, C. (1996). *An Investigation into Students' Understanding of Statistics*. Unpublished Doctoral Thesis, University of New England, Armidale.

Reeves, T. C. (2005). Design-based research for advancing educational technology. Paper presented at the Onderwijs Research Dagen, Gent.

Snee, R. D. (1993). What's missing in statistical education? *The American Statistician*, 47, 149–154.

Van Buuren, J. A. and Bijker, M. M. (2006, in press). Evaluation of the IMTO-pilots 1999-2004. Open Universiteit Nederland.

Wild, C. J. and Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67, 223–265.