University of Wollongong

Research Online

1-1-2013

Innovations of the assessment system in introductory statistics subject

Norhayati Baharun
Universiti Teknologi MARA, norhayat@uow.edu.au

Anne Porter
University of Wollongong, alp@uow.edu.au

Follow this and additional works at: https://ro.uow.edu.au/eispapers

Part of the Engineering Commons, and the Science and Technology Studies Commons

Recommended Citation

Baharun, Norhayati and Porter, Anne, "Innovations of the assessment system in introductory statistics subject" (2013). *Faculty of Engineering and Information Sciences - Papers: Part A*. 2388.
https://ro.uow.edu.au/eispapers/2388

Innovations of the assessment system in introductory statistics subject

Abstract

This paper presents the results of a study investigating the innovations in the assessment system implemented in an introductory statistics subject (STAT131) at the University of Wollongong (UOW). For several years, innovations have been introduced to STAT131 particularly in the assessment system. These included the approaches of assignments, summary and mid-session test in 2003; assignments, laboratory work and midsession test from 2004 to 2005; four tests and three make-up tests in 2006; six tests with three compulsory and two optional tests in 2007; four tests (the best of three test marks were chosen) and opportunity to re-sit the tests in 2008; five tests and retests assignments with a competency requirement of 65% to 70% from 2009 to 2010; and finally the draft and final of the assignment (in the Headstart program) and a group draft and final assignment within session, and three tests and retests assignments with competency 70% required for all assessments in 2011. The findings reveal that there was improvement in the learning outcomes where the failure rate dropped from 18% in 2010 to 13% in 2011 and students in 2011 outperformed than in 2010 in their mean final marks. Furthermore, a dramatic increase in the higher grades of 64% in 2011 was the highest on record since 2000. The paper concludes with a discussion on the issues arise and followed by suggestions for further research.

Keywords

introductory, system, subject, assessment, statistics, innovations

Disciplines

Engineering | Science and Technology Studies

Publication Details

Innovations in assessment system of introductory statistics subject

*Norhayati Baharun and **Anne Porter
*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (Perak), Malaysia
norha603@perak.uitm.edu.my
**School of Mathematics and Applied Statistics, University of Wollongong, Australia
alp@uow.edu.au

Abstract— **This paper presents the results of a study investigating the innovations in the assessment system implemented in an introductory statistics subject (STAT131) at the University of Wollongong (UOW). For several years, innovations have been introduced to STAT131 particularly in the assessment system. These included the approaches of assignments, summary and mid-session test in 2003; assignments, laboratory work and mid-session test from 2004 to 2005; four tests and three make-up tests in 2006; six tests with three compulsory and two optional tests in 2007; four tests (the best of three test marks were chosen) and opportunity to re-sit the tests in 2008; five tests and retests assignments with a competency requirement of 65% to 70% from 2009 to 2010; and finally the draft and final of the assignment (in the Headstart program) and a group draft and final assignment within session, and three tests and retests assignments with competency 70% required for all assessments in 2011. The findings reveal that there was improvement in the learning outcomes where the failure rate dropped from 18% in 2010 to 13% in 2011 and students in 2011 outperformed than in 2010 in their mean final marks. Furthermore, a dramatic increase in the higher grades of 64% in 2011 was the highest on record since 2000. The paper concludes with a discussion on the issues arise and followed by suggestions for further research.**

Keywords - assessments, innovations, statistical learning, students' performances, learning outcomes

I. INTRODUCTION

Students in New South Wales differ in some respects from those in other Australian school systems in that they receive limited exposure to the subject of statistics in secondary schools. Neither has there been any requirement to support the study of statistics with technology... The subject of statistics has consequently not generally been highly regarded by students and they have, more often than not, approached the study of statistics with a negative mindset... (p. 142) [4]

Consequently, there is a need to support the student learning of statistics at the university level. For this reason, many educational innovations in Australia have been developed which aim to assist the teaching and learning of statistics, particularly the innovations in the assessment system. To examine whether such innovations are successful or otherwise requires evaluation within the context in which they are applied. Essentially, in order to understand the context of this study, it has been considered prudent to inspect (i) the reason the subject was selected, (ii) baseline data regarding the pass and failure rates for the subject selected for intervention, and (iii) approaches to assessment.

II. THE CONTEXT

The selection of subject

STAT131 Understanding Variation and Uncertainty is an introductory first year level university statistics subject which has been designed and developed particularly for students in the degree program Information Technology and Computer Science at the University of Wollongong. STAT131 is a six-credit point subject (implying 12 hours of work a week) and is compulsory for most of its students, many of them in the first year of their degree program. This subject has been delivered on-campus throughout the session (although in the years 2004 to 2009 it was also delivered at a remote campus). In short, Table 1 details some components of the subject.

Table 1 Components of STAT131

Component	STAT131
Lectures	3 hours weekly
Laboratory	2 hours weekly
Computer Software	SPSS
Prescribed textbook	Utts, J. M. and Heckard, R. F. (2004), *Mind on Statistics*, Thomson Brookes/Cole packaged with Student Version SPSS
Number of students enrolled	90 to 180 (Wollongong) 8 to 15 (Loftus)
Session offered	Also available pre 2003[a] March/Autumn (2003, 2008 to 2011) Both sessions (2004 to 2007)
Campus offered	Wollongong (on-campus) Loftus (remote campus)[b]

[a]Data were not included in this study
[b]Subject was not offered at remote campus in 2010 and 2011

Based on student-centred and blended learning approaches, the subject has incorporated both face-to-face lectures and

online learning resources through the Blackboard Learning system (the e-learning system used at the UOW). Since the nineteen nineties, all resources have been provided in the e-learning system. Communication with other students/lecturer is by both email and the e-learning forum. Resources in the e-learning sites include lecture notes, laboratory manual (including laboratory tasks), SPSS notes, Edu-Stream (recorded lectures), student forum, video resources, worked solutions, past laboratory tests and exams, assessment, and data sets. In reference to STAT131, reference [3] noted that '[i]n recent years there has been an increasing numbers of students using online resources rather than face-to-face [lectures]'. In 2010 sessions, due to a scarcity of laboratory space, higher year STAT131 students with no previous failures were allowed to complete their laboratory work from the laboratory manual at home or in their own time provided they had the necessary hardware and computer software (i.e. SPSS). The number of students who took this option was small. Assistance is available for students during consultation times (4 hours outside of class time) allocated by the lecturer of the subject in particular session, or by an appointment made through telephone or email at a convenient time. The use of such consultation is very low with most problems or questions addressed through email or the e-learning forum.

Baseline data: failure rates

In STAT131, a review of student grades from 2003 to 2008 revealed failure rates ranging between 9.3% and 24.3% (as shown in Table 2). Interestingly, there was a dramatic reduction of the failure rate in March/Autumn 2004 to 9.3% coinciding with the change in the laboratory classes by producing a laboratory manual including more authentic tasks that engaged students in the data gathering process. The low failure rate was not sustained although the valuing of the resource has remained high. In the later years, worked solutions were available and students had needs to be taught to do tasks not just read solutions. In describing the pattern of student grades according to the sessions, the distribution of the failure rates in March/Autumn was slightly lower (between 9.3% and 20%) compared to August/Spring (between 15.1% and 24.3%). In other words, an increase in the failure rate was evident when offered in August/Spring. However, the lack of a trend in failure rates was possibly due to the variation in the background of cohorts (students enrolled), lecturer(s) or tutor(s) involved, assessment systems used in the subject and other factors. The students, who once completed in August/Spring, now complete in March/Autumn. In addition, a small number of failing students from each session re-enrol in the subsequent session but data on these individuals has not been available within the data collection structure in this study. This subject has had a history of innovations in teaching practice in March/Autumn, with the learning promoted through experiential learning, authentic tasks and the use of technology i.e. online learning [2, 4, 5]. The evaluation practice of this subject has encompassed data from several sources, particularly assessment in this study.

Table 2 Student grades (in percentage) for STAT131

Year[a]	N	Fail	Pass Conceded & Pass	Credit	Distinction	High Distinction
A 2003	165	10.3	48.5	17.6	12.7	10.9
A 2004	205	9.3	49.7	20.0	13.2	7.8
A 2005	162	19.1	45.7	18.5	11.1	5.6
A 2006	136	17.7	27.9	18.4	21.3	14.7
A 2007	125	20.0	29.6	25.6	16.0	8.8
A 2008	108	16.7	25.9	23.2	15.7	18.5
S 2004	172	15.1	34.3	22.1	16.3	12.2
S 2005	156	21.1	43.0	21.1	10.3	4.5
S 2006	84	22.6	33.3	9.5	16.7	17.9
S 2007	107	24.3	37.4	18.7	13.1	6.5

[a]A=March/Autumn, S=August/Spring
(Data source from the University of Wollongong, Performance Indicator Database, 09/12/2010)

Approaches to assessment

Over the years, in the search for the best approach to assessment in terms of enhancing student learning outcomes, a variety of assessment tasks have been employed in STAT131. As can be seen in Table 3, these have included assignments involving the collection and analysis of data, portfolios or laboratory work, summaries, laboratory tests, group presentations (teams of two students), in class and final examinations where different weightings of assessment have been applied across sessions. There was a variation in the continuous assessment approach between March/Autumn 2003 and March/Autumn 2008 with a constant presentation and final examination remained at five per cent and fifty per cent of total assessment marks, respectively. Laboratory work and the mid-session test were applied in four consecutive sessions between March/Autumn 2004 and August/Spring 2005 whereas the mid-session test and summary were applied in March/Autumn 2003 as well as other types of assessment (see Table 3).

Table 3 Assessment weightings for STAT131 across sessions

Session[a]	Assignment	Laboratory work	Summary	Presentation/ Group project	Mid-Session test	Laboratory test	Final exam
A 2003	30	-	5	5	10	-	50
A 2004 S 2004 A 2005 S 2005	30	10	-	5	5	-	50
A 2006 S 2006	20	-	-	5	-	25	50
A 2007 S 2007	20-30[b]	-	-	5	-	15-25[c]	50
A 2008	30	-	-	5	-	15	50

[a]A= March/Autumn, S=August/Spring
[b]Three assignments with one optional assignment (minimum of 20% to maximum of 30%)
[c]Six laboratory tests with three compulsory and two optional tests (minimum of 15% to maximum of 25%)
- Not assessed this session

Assignments have been used to assess student learning and understanding of statistics in each session although there were changes in weightings and the number of submissions between two and three assignments depended on whether they were compulsory or optional.

Laboratory tests have been used in a variety of ways since March/Autumn 2006 until recently. There has been some variation in the number of assessment tasks and their weightings between March/Autumn 2006 and March/Autumn 2008. For example, there were four laboratory tests applied in 2006 for both sessions. In addition, three make-up tests were designed for students who obtained less than seventy per cent in any given test (zero was awarded to students for any test mark less than seventy per cent). In 2007 for both sessions, there were six laboratory tests (three compulsory and two optional tests) in which any test mark that less than sixty per cent was awarded zero and the tests due in laboratory classes. Changes in the number of laboratory tests and opportunity to re-sit the test have been applied in March/Autumn 2008 where the best of three test marks were chosen out of four laboratory tests. As applied in 2007, students who obtained less than seventy per cent were given the opportunity to re-sit the test in the following week with a different data set and completed during laboratory classes. Nevertheless, students who failed the re-test have further been examined by the subject coordinator to clarify any problems they experienced. The reason for having a minimum mark of sixty or seventy per cent in each test was primarily to enhance student competency in the topics examined. In particular, students who were awarded zero were expected to demonstrate their competency through a retest that was offered in the following week. This test retest approach has come to form the basis of the support system to identifying students at risk typically those who needed the encouragement to do a retest. The students who failed to sit retests after having been given feedback were often found to have issues such as anxiety, lack of confidence, depression, obsessive or other difficulties. Tests typically comprise the analyses of data sets in addition to the understanding of theoretical concepts.

The Headstart program

Initiated in February/Autumn 2011, a Headstart program was introduced to STAT131 allowing students to start engaging with statistics prior to the commencement of the session. The students accessed the Headstart program through the e-learning system set up for STAT131 in February/Autumn 2011. The Headstart was originally conceived of as a program that extended the time students would have to learn statistics. The idea of introducing this program in the subject was based on the students' experience, which suggested that the thirteen weeks of session allow insufficient time for them to adequately learn and understand the subject. An alternative approach would involve a curriculum review; it was decided to leave the subject objectives the same and to not reduce the content or processes to be learned. The Headstart was an optional program held in the four weeks before the session formally commenced in 2011. Some 'within session' tasks and resources were included in the Headstart in addition to alternative tasks and resources.

In this program, students were allowed to access the first module of work in the subject and to complete an alternative first assessment. The students who successfully completed the first assessment given in this program were also allowed to skip the first laboratory test assessed in the formal 2011 session. These students were then required to complete the second assessment (i.e. assignment) in a formal session as shown in Figure 1.
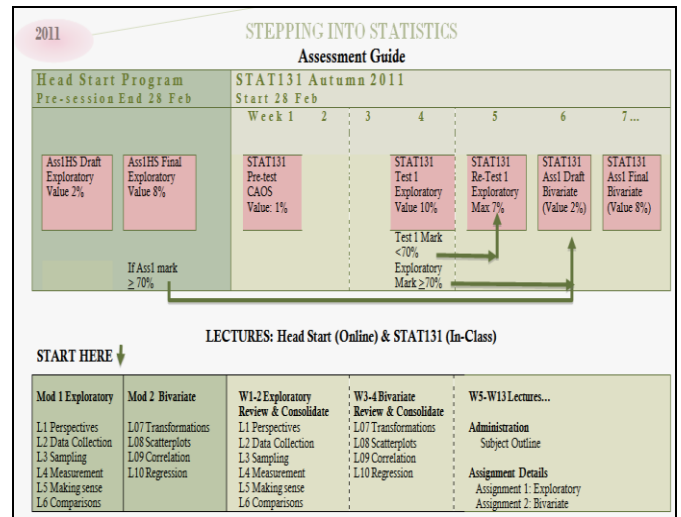
Note: The second module did not materialise due to time involved in addressing HTML coding issues with the e-learning system

Figure 1 Assessment guide designed for the Headstart program

The introduction of the Headstart program has allowed students who engaged with the program to complete a draft, redraft, and submit the first assessment via the e-learning site. Students who attained marks 70% or above for this assessment were not required sitting the first test or retest. The second assessment was of the form draft/final and the remaining three assessments test and retests.

III.   FINDINGS

Over the period of the years 2000 to 2004, the failure rate in STAT131 declined from the highest 19% to the lowest 9%. The proportion of students attaining top grades (High Distinction and Distinction) was also in decline that moved from the top 27% to bottom 21% (see Table 4). From the years 2006 to 2010, it was found that the proportion of students attaining higher grades of High Distinction, Distinction, and Credit was higher but inconsistent. On the other hand, during these years, the failure rate was slightly up on average 20% compared to 2001 to 2004 where the rate remained under 15% (see Figure 2). The results in 2011 showed a dramatic increase

in the top grades rate of High Distinction and Distinction at 51.8% and a drop in the failure rate to 12.8% (25 students) and of these students, 32% (8 students) had effectively ceased to participate and did not sit the final examination.

Table 4 Pass rates for the years 2000 to 2011 in STAT131

Year	N[a]	Fail	PC[b] & Pass	Credit	Distinction	High Distinction
2000	157	18.5	32.5	23.6	14.0	11.4
2001	152	9.9	38.2	25.0	19.7	7.2
2002	141	14.9	32.6	25.5	15.6	11.4
2003	165	10.3	48.5	17.6	12.7	10.9
2004	205	9.3	49.7	20.0	13.2	7.8
2005	162	19.1	45.7	18.5	11.1	5.6
2006	136	17.7	27.9	18.4	21.3	14.7
2007	125	20.0	29.6	25.6	15.2	9.6
2008	108	16.7	25.9	23.2	15.7	18.5
2009	89	22.5	24.7	27.0	22.5	3.3
2010	191	18.3	28.8	25.7	17.8	9.4
2011	195	12.8	23.1	12.3	24.6	27.2

[a]Number of students enrolled
[b]Pass Conceded
(Data extracted from SMP, University of Wollongong. Figures might change slightly at different dates as students were retrospectively withdrawn without penalty from the subject)
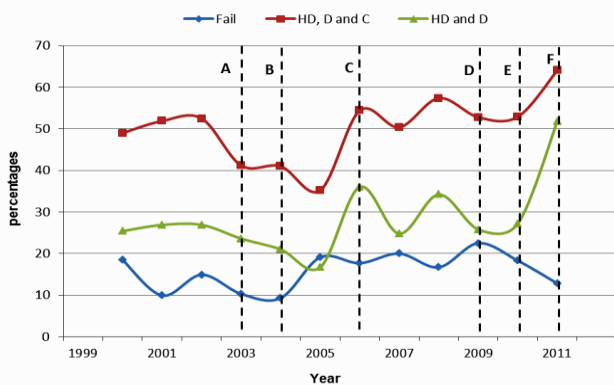
Note: HD = High Distinction ; D = Distinction ; C = Credit
- - - Innovations in the assessment system introduced in the subject
A = Use of real data and working topics of social significance in the laboratory classes
B = Laboratory manual including authentic tasks, and alignment of objectives, tasks, and marking criteria
C = Changes in the assessment system (3 compulsory and 2 optional tests, zero mark for any test less 60%)
D = Test retest approach (retest for any test less 70%)
E = Test retest approach (retest for any test less 70%)
F = Headstart program with draft and redraft of assignment, and test retest approach (retest for any test less 70%)

Figure 2 Percentages of fail versus higher grades and top grades, and innovations introduced in STAT131 for the years 2000 to 2011

Apart from the changes of the assessment system, STAT131 had been taught by the same lecturer all years except in 2009 and that year involved shared lecturing. In 2009, video resources were introduced in the subject as support materials for student learning on the e-learning site. Though the resources were perceived as useful for learning by the majority of the postgraduates in other subject, they seemed less successful in assisting the students in STAT131. This led to a search for a more effective learning design for embedding the resources, tasks, and support materials in the e-learning system.

Consequently, a learning design map within weekly folders was introduced in 2010 incorporating several links to resources. These links included video resources in connection to weekly laboratory tasks on the e-learning site. Accompanied by the test retest approach implemented since 2009, there was a turning point where the failure rate headed down to 18.3% and the higher grades rate (High Distinction, Distinction and Credit) headed up to 52.9% in 2010. In 2011, a further lowering of fails and increases in good grades with a divergence in fails and good grades occurring with the Headstart program, and the draft and redraft of assignments. The improved learning design maps, and the by-type resource folders were also made available in the e-learning site in 2011.

A test for differences in two proportions revealed that the proportion of students failing fell significantly ($Z = 1.99$, $p = 0.023$) in 2011 compared to the overall proportion failing between 2005 and 2010. One-way ANOVA and Scheffe post-hoc tests were used to examine the differences in means of final marks between the three cohorts. The result revealed a strong evidence of differences in mean final marks ($F = 10.004$, $p < 0.001$) between the cohorts. In particular, the 2011 cohort attaining significantly higher mean final marks than the 2010 cohort ($p = 0.004$) with a mean difference of 8.2 marks. Likewise, the 2011 cohorts achieved an average of 12.4 marks better than the 2009 cohort ($p < 0.001$) in the final examination (see Table 5). However, no significant difference was evident in the mean final marks between the 2009 and the 2010 cohorts. This indicates that the new assessment system such as the Headstart program, and the draft and redraft of assignment, along with other innovations in the resources provided in the e-learning site had potentially helped improve student learning outcomes in 2011.

Table 5 Comparison of mean final marks between the three cohorts in STAT131

	N	Average marks	Standard Deviation
March/Autumn 2011	195	67.93[a,b]	22.38
March/Autumn 2010	191	59.73[a]	23.84
March/Autumn 2009	89	55.51[b]	27.66
Total	**475**	**62.30**	**24.48**

[a]The mean difference is significant at $p = 0.004$
[b]The mean difference is significant at $p < 0.001$
Note: Students who did not engaged with the Headstart program in 2011 attained average final marks of 62.50 which were marginally higher than both in March/Autumn 2010 and 2009

An examination of student grades also revealed that the proportion of students with High Distinction and Distinction grades was the highest in record at 51.8% over the years 2000 to 2011. With respect to failures, the rate would have been much higher when having the assessment system not allowed the identification of students at risk and the subsequent work with them to develop their competency. This however was not a controlled study; there may be many factors at play in the overall improved results, and future monitoring will be

necessary in order to see if the impact of the Headstart program remains.

IV.   CONCLUDING REMARKS

Several innovations have been implemented in STAT131 for the past 10 years aimed at improving the student learning and understanding in the subject. These innovations had involved from the collection and use of real data and working topics of social significance. Students continued to work with real data but with a manageable number of outliers and distributions. Assessment tasks involved the collection and analysis of data, portfolios, summaries, tests and presentations to class and final examinations. Currently, the assessment has stabilised on "redeeming approaches", test and retest, and draft and redraft of assignment in addition to a final examination. Besides that, the design in 2009 involved the provision of video supports accessible via by-type resource folders in the e-learning site. In 2010, these resources were accessible via links in the learning design maps provided within weekly folders in the e-learning site. At this point, it seemed momentarily that "everything" had been tried to improve learning outcomes. Finally in 2011, the introduction of the Headstart program, the draft and redraft of assignment, as well as other learning resources provided in the subject resulted in improved student outcomes.

The aim of this study was primarily to investigate the impact of the approaches to assessment system on student learning outcomes in STAT131. This was done by examining changes in performance particularly for three cohorts of students from 2009 to 2011. Students with access to the Headstart program, the draft and redraft of assignment, as well as other learning resources were found to have better performance in their assessment compared to students without access. The mean final marks in 2011 were significantly higher than in 2010 ($p = 0.004$), and in 2009 ($p < 0.001$). The failure rate fell significantly in 2011 compared to the years between 2005 and 2010 ($Z = 1.99$, $p = 0.023$), and most importantly the proportion of failures declined from 23% in 2009, 18% in 2010 to 13% in 2011.

Looking over all the assessment tasks, it would appear that the recent innovations implemented in the subject have been effective in improving student performance in STAT131. This was supported by the students' comments demonstrated in favour of the provision of the Headstart program to support their learning in the subject. The success of these new assessment systems particularly the Headstart program has led to successful funding to implement it in two additional subjects at the UOW [1].

REFERENCES

[1] A. Porter, "Application for Faculty Teaching and Learning Scholar 2011 – 2012", Faculty of Informatics, University of Wollongong, Australia, 2011.

[2] A. Porter, "Improving Statistical Education through the Experience of Reflective Practice", Unpublished Ph.D. Thesis, University of Wollongong, Australia, 2001.

[3] A. Porter, and N. Baharun, "Stepping into Statistics: Providing a Head Start for students", In L. Paditz and A. Rogerson (Eds.), Proceedings of the 11th International Conference, Turning Dreams into Reality: Transformations and Paradigm Shifts in Mathematics Education, Rhodes University, Grahamstown, South Africa, pp. 276 - 281. ISBN: 83-919465-0-9, 2011.

[4] M. M. Morris, "Evaluating university teaching and learning in an outcome-based model: Replanting bloom.", Unpublished Ph.D. Thesis, University of Wollongong, Australia, 2008.

[5] N. Baharun, and A. Porter, "The impact of video-based resources in teaching statistics: A comparative study of undergraduates to postgraduates", In C. Reading (Ed.), Data and context in statistics education: Towards an evidence-based society, Proceedings of the eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia. Voorburg, The Netherlands: International Statistical Institute, 2010.