

Promoting Active Learning when Teaching Introductory Statistics and Probability Using a Portfolio Curriculum Approach

Desmond Adair¹, Martin Jaeger² & Owen M. Price³

¹ School of Engineering, Nazarbayev University, Astana, 010000, Republic of Kazakhstan

² School of Engineering, University of Tasmania, Hobart, TAS7001, Australia

³ Australian Technical & Management College, Melbourne, 3000, Australia

Correspondence: Desmond Adair, Department of Mechanical Engineering & Aerospace Engineering, School of Engineering, Nazarbayev University, 53 Kabanbay batyr Ave., Astana, 010000, Kazakhstan

Received: February 14, 2018

Accepted: March 6, 2018

Online Published: April 2, 2018

doi:10.5430/ijhe.v7n2p175

URL: <https://doi.org/10.5430/ijhe.v7n2p175>

Abstract

The use of a portfolio curriculum approach, when teaching a university introductory statistics and probability course to engineering students, is developed and evaluated. The portfolio curriculum approach, so called, as the students need to keep extensive records both as hard copies and digitally of reading materials, interactions with faculty, interactions with other students and work they have completed on their own, is designed to encourage active learning, mainly in the areas of cooperation and collaboration. In order to investigate the effectiveness of the portfolio curriculum, a controlled experiment applying a pre-test-post-test control group design is conducted. Two tests are conducted, one before the commencement of the course (pre-test) and one after the completion of the course (post-test). The effectiveness is evaluated by comparing within-subject post-test and pre-test scores and by comparing the scores between subjects in the experimental group, i.e., those who learned using the portfolio curriculum approach and subjects in the control group, i.e., those who learned using a traditional method of teaching. In addition to analysis of the controlled experiment, a Survey of Attitudes Toward Statistics (SATS) was completed on the first and last day of the semester by the participants so as to give a measure of student confidence, understanding, liking, and difficulty of the portfolio curriculum approach as opposed to using a traditional method of teaching and learning. The findings of these investigations are reported and discussed, as are the merits and problems encountered regarding the methodology and student attitudes regarding the portfolio curriculum approach.

Keywords: portfolio curriculum, active learning, cooperative learning, collaborative learning, symbolic computation

1. Introduction

A portfolio curriculum approach has been designed for an introductory statistics and probability course at the tertiary level to second-year engineering students in an effort to encourage active learning, including cooperation and collaboration between students. Students were actively encouraged to keep extensive records, summaries, reflections, etc. of a given day's activities in a portfolio, both in hard-copy and digital form.

Active learning has received considerable attention over this past number of years and is often perceived as a radical change from traditional instruction. It is not possible to provide universally accepted definitions for all the vocabulary of active learning, since different authors interpret some terms differently, but it can generally be defined as any instructional method that engages students in the learning process, i.e. it requires students to do meaningful learning activities and think about what they are doing (Felder & Brent, 2016). This definition could be thought of as a tautology, as it is not possible to learn unless the brain or body is active in some way or other (Christie and de Graaff, 2017) and traditional activities such as homework can be thought of as active learning. However in practice, active learning refers to activities that are introduced into the classroom. The core elements of active learning are student activity and engagement in the learning process (Prince, 2004).

The extensive literature on active learning reports that, on balance, the use of active learning gives positive results (Garfield & Ben-Zvi, 2007; Keeler & Steinhorst, 1995; Giraud, 1997; Magel 1998) when the teaching and learning approach is cooperative learning. Meletiou & Lee (2002) organized their curricula along a Project-Activities-Cooperative-Learning-Exercises model emphasizing statistical thinking and reasoning and an orientation towards investigating conjectures and discovery of results using data. Cooperative learning can be defined

as a structural form of group work where students pursue common goals while being assessed individually (Millis & Cottell, 1998; Feden & Vogel, 2003) with the most common model found in engineering literature being that of Johnson *et al.* (1998). This model incorporates five specific tenets, which are individual accountability, mutual interdependence, face-to-face promotive interaction, appropriate practice of interpersonal skills, and regular self-assessment of team functioning (Prince, 2004). Other models for cooperative learning exist but the common element between them is cooperative incentives rather than competition to promote learning. A similar, though different approach to cooperative learning, which also gets improved results is collaborative learning. The latter can refer to any instructional method in which students work together towards a common goal and, as such, collaborative learning can be viewed as encompassing all group-based instructional learning, including cooperative learning (Millis & Cottell, 1998; Smith & MacGregor, 1992; Felder *et al.* 2002). The core element of collaborative learning is the emphasis on student interactions rather than on learning as a solitary activity (Prince, 2004). A common method of instruction called Problem-based learning is always active and commonly is collaborative or cooperative. However, typically problem-based-learning involves significant amounts of self-directed learning on the part of the students with the relevant problems being introduced at the beginning of the instruction cycle and used to provide the context and learning that follows.

Although numerous studies have found active learning to be effective, others have found it to have no effect (Pfaff & Weinberg, 2009) or even hinder student performance (Weltman & Whiteside, 2010). Perhaps one of the reasons for the inconsistent results in the active learning literature is the enormous diversity of approaches that are referred to as active learning (Carlson & Winqvist, 2011).

The present portfolio curriculum approach is intended to move students away from the traditional lecture/tutorial/computer-laboratory approach, where the stress has been on a teacher-centred approach to a more student-centred approach with the inclusion of cooperation and collaboration between students. As the students receive less formal instruction during the portfolio curriculum approach, special care was taken to introduce the teaching material in such a way so as to not overwhelm the students.

A controlled experiment, for the cohort numbering 181 students, applying pre-test-post-test assessments was designed so that the results of the experimental group of students taught using the portfolio curriculum method was compared with a control group taught in a more traditional manner (Pfahl *et al.*, 2004; Adair *et al.*, 2011). Traditional here means an approach to learning statistics and probability using lectures, tutorials, and computer-laboratories.

In addition to the controlled experiment, use was also made of the Survey of Attitudes Towards Statistics (SATS) (Schau *et al.*, 1995). Measuring students' attitudes towards statistics both before and after completing a course is another way to assess a statistics and probability curriculum effectiveness (Harlow *et al.*, 2002; Manning *et al.*, 2006; Sizemore & Lewandowski, 2009). Not only is attitude change an important outcome in its own right, but student attitudes towards statistics before and/or after taking a statistics course are also associated with other outcome variables (Carlson & Winqvist, 2011). That is to say, more positive attitudes are associated with better performance in the course (Elmore *et al.*, 1993; Finney & Schraw, 2003; Chiesi & Primi 2009). However caution must be exercised as reported by (Carlson & Winqvist, 2011) when dealing with measures of the change in students' attitudes toward statistics across a semester.

2. The Portfolio Curriculum Approach

2.1 Overview

The portfolio curriculum approach, can be regarded as one where students read provided content through a learning management system before, and during class, and then work in groups to complete closed and open questions related to the provided material, in addition to later completing projects and assignments. Much effort has been extended to developing more interesting and innovative teaching strategies with the inclusion of real-life projects, a variation of activities and cooperative learning using symbolic computation.

A typical learning and teaching session would start the day before, when students were sent notes and examples through a learning management system on a given statistical problem, (e.g. normal approximation of the binomial distribution, the logic of hypothesis testing). Care was taken to use clear and simple language in these documents. The class period began with an instructor-led discussion where the complete cohort gather together for a question and answer session regarding the information sent and to summarize what was expected to be completed by the end of the session. Following this, the students divided into pre-arranged groups to work through the statistical topic of the day described in worksheets. It was found important that each statistical topic was carefully broken into

subsections to give the students both a structure to work within and to encourage them as they could see progress was being made.

2.2 Forming Groups

This part of the portfolio curriculum approach required careful planning, both for the instructor and the students. It was decided to form group sizes of 4 - 6 students (Ingham *et al.*, 1974). However, determining the composition of each group was more complicated since ideally the groups should be diverse enough to include students with a range of intellectual abilities, academic interests and cognitive styles. Perhaps allowing students to select their own group members can work well with small cohorts but this method always runs the risk of further isolating some students or creating cliques within the cohort as a whole. The cohort here was considered large, and while random selection may have worked, it was decided to use the students' GPA scores coupled with what they put in an index of learning styles questionnaire (Felder & Spurlin, 2005) at the beginning of the semester, to form the student groupings. No allowance was made for gender in the formation of the groups. Regarding learning styles, there does not seem to be much agreement in the literature as to the value of learning styles or that learning styles even exist (Felder & Brent, 2005). There seems to be however some agreement that grouping students with similar learning styles increases the efficiency and success of the group, whereas by forming groups with students who have different learning styles, then students have the opportunity to develop their less dominant styles. Whatever the rights and wrongs of this ongoing argument, it was decided here to form groups, as far as possible, with diverse learning styles, in an effort to homogenize the learning process (Stice, 1987).

Importantly, the method of formation of the groups seemed fair to the students in the cohort.

2.3 Effective Working of the Groups

With the groups formed, it was necessary to spend some time teaching the students how best to work within their group. Students are normally rewarded for individual effort and cooperation and it often does not come naturally for a student, nor have they perhaps thought of what extra skill is needed to best promote group achievement. Some of the better students may not wish to be burdened with students whose success has been moderate in the past, some students may not appreciate having to listen to their equals, and some students are simply shy and not used to working with their peers. Therefore from the start it was important to address the usefulness of group work to motivate students to participate fully. Also, the down side of group work is that a person, i.e. the instructor is not always available to lead a group through a particularly complex intellectual study and students need to build quickly their interpersonal and organizational skills needed for managing the group session, so that students recognize the importance of such things as: listening, clarifying statements and providing good feedback; keeping discussions on task; probing assumptions and evidence; eliciting viewpoints and perspectives; mediating conflicts; and summarizing and presenting findings (Bosworth, 1994).

To help alleviate some of these difficulties, if specific skills were identified in a worksheet, the instructor identified them and provided similar examples of successful use of such skills. Also to help the group operate coherently, quite often, within each group a facilitator was appointed to lead discussions, a note-taker to record and summarize progress, a planner to outline where and how the group was proceeding and an evaluator to elicit critiques. However it was impressed on each student that they must keep their personal portfolio up to date and comprehensive. These group management skills have been identified in the literature as being important, with students encouraged to reflect on their successes and difficulties when exercising these skills, because few participants saw the relationship between completing the group session and achieving the larger goals of the statistics course. The time taken to examine these skills was felt crucial to the success of the portfolio curriculum approach, as already expressed in the literature (Miller *et al.*, 1994).

Occasionally, especially during the teaching of hypothesis testing, use was made of the "jigsaw" teaching strategy (Tewksbury, 1995) where each team became an expert on a topic, and then individuals from that team each taught another team.

2.4 Promoting Learning within Each Group

Some thought was also given to how tasks and work sheets were distributed to each group. Generally for these engineering students the emphasis was on how to apply theoretical knowledge to real-world problems and, where possible, to simulate typical decision-making or problem-solving akin to those made by professionals already in the field. Also thought was given to the inclusion of all members of the group to avoid natural tendencies for some students to dominate while others will become withdrawn. To try to avoid this, each member of the group had to keep a record in their portfolio noting their contribution or, on occasions, there would be a group discussion with the

instructor present. Several times students were also asked to put in their personal portfolio critiques of the other group members' contributions. Also, to promote learning, the students were given problem-based tasks which had a definite outcome. It was observed that this led to many information-rich discussions between students showing that there was a respect for other student's opinions and a realization that the task becomes easier if everyone cooperates. It is recognized that immediate, unambiguous and meaningful feedback was essential. Perhaps this was the weakest aspect of the portfolio method as the number in the cohort was large. It was only possible to give considered feedback on a weekly basis when students handed in their portfolio or sent a digital version by email. This was though not very satisfactory. Of course any verbal feedback was instantaneous, but was also not systematic across the complete cohort.

2.5 Assessment

Individual accountability was thought important, so in addition to peer assessment for each member of the group (10%), each student's portfolio (10%), individual assignments (20%), midterm test (10%) and final examination (50%) also contributed. Individuals were allowed to collaborate, if they wished when doing their assignments. The evaluation form for peer assessment, which was made available early in the course was based on Cramer (1994) and contained such areas as professionalism (attitude, punctuality), initiative (constructiveness, suggestions, achieving common goals), and independence (completion, deadlines, sharing). The final examination for the experimental group A and the control group B was the same examination.

3. The Controlled Experiment

3.1 Description

To investigate the effectiveness of using a portfolio curriculum approach to teaching and learning introductory statistics and probability, a controlled experiment applying a pre-test-post-test control group design was conducted following Pfahl *et al.*, (2004); Adair *et al.*, (2011). The students had to undertake two tests, one before the respective course (pre-test) and one after the respective course (post-test) with the effectiveness of the teaching approaches then being evaluated by comparing within-student post-test to pre-test scores, and by comparing the scores of the students in the experimental group (A), i.e. those taught using the Portfolio Curriculum approach, to those students in the control group (B), i.e. those taught using the traditional approach.

3.2 Hypotheses

To measure the performance of the two groups, five constructs were used, with each construct represented by one dependent variable. Each dependent variable has a hypothesis:

- 1) There is a positive learning effect in both groups (A: experimental group, B: control group). That is post-test scores are significantly higher than the pre-test scores for each dependent variable.
- 2) The learning is more effective for group A than for group B, either with regard to the performance improvement between pre-test and post-test (the relative learning effect), or with regard to post-test performance (absolute learning effect). The absolute learning effect is of interest as it may indicate an upper bound of the possible correct answers depending on the method of teaching.

3.3 Method

The design started with random assignment of students to the experimental group (A) and control group (B) with the members of both groups completing a pre-test and post-test. The pre-test measured the performance of the two groups before the courses and the post-test measured the performance of the two groups after the course. The students did not know that the post-test and pre-test questions were identical and neither were they allowed to retain the pre-test questions with the correct answers only being given to the students on the completion of the experiment. It was made clear to the students that these tests were not included in their course grades.

The students were second year engineering students and a mixture of mechanical, electrical civil and chemical engineering disciplines. The number of students in group A was, $N_A = 91$, and in group B was, $N_B = 90$. The personal characteristics of the students are summarized in Table 1.

Table 1. Personal characteristics

Characteristics	
Average age	19.97 years
Percentage female	51.3% (93 students)
Major	ME 19%, EEE 38%, CE 30%, ChE 13%
Experience with statistics & probability:	
• Never studied probability & statistics	3% (5 students)
• Studied probability & statistics at secondary level	97% (175 students)
• Studied probability & statistics at tertiary level	0%
Has experienced:	
• Traditional & teacher centred learning	98% (177 students)
• Traditional & student centred learning	21% (38 students)
• Group based collaboration learning	5% (9 students)
• Group based cooperative learning	3% (5 students)
• Problem Based Learning	0%

The initial testing was conducted after a short introduction as to the purpose of the experiment and general organizational issues and, the pre-test carried out with data for the independent variables collected. Following the pre-test, the students were placed in either the control group or the experimental group and all students participated in both the pre-test and post-test. After completing their respective courses, both groups performed the post-test using the same questions as given during the pre-test, thus providing data on the dependent variables for the second time.

3.4 Teaching Courses Details

The syllabus used for the experimental and control groups was the same and is outlined in Appendix A. For the experimental group, the course was delivered as described in Section 2 above. The instructing staff consisted of one faculty and three teaching assistants, all of whom were present during each teaching and learning session. The course was run over twelve weeks with two 110 minute sessions per week, and the students had access to resources like the R Project software (R Core Team, 2014) and Mathematica (Wolfram Research, 2016) when appropriate.

For the control group, a more traditional approach was employed with a two-hour lecture followed by a two-hour tutorial per week over the twelve weeks. This group also had access to the R Project software and the Mathematica symbolic computational software when appropriate. Course assessment for this group was by final examination, assignments with and without the aid of computational software and a mid-term test. Any group work the control group used was purely informal and was neither encouraged nor discouraged.

3.5 Data Collection

Data for the dependent variables (J.1...,J.5) was collected with the variables listed in Table 2. The dependent variables are constructs used to capture aspects of learning provided by the courses and each was measured using five questions.

Table 2. Experimental variables

Dependent variables
J.1 Interest in statistics and probability ('Interest')
J.2 General knowledge of statistics and probability ('Understand general')
J.3 Understanding of elementary probability ('Understand elementary probability')
J.4 Understanding of probability distributions ('Understand probability distributions')
J.5 Understanding of hypothesis testing ('Understand hypothesis testing')

The five questions can be characterized as:

J.1 ('Interest'): Questions about personal interest in learning about statistics and probability.

J.2 ('Understand 'general'): Questions to elicit how much students understand the role of statistics and probability in engineering and generally.

J.3 ('Understand 'elementary probability'): Questions on probability set at the elementary knowledge level.

J.4 ('Understand 'probability distributions'): Questions on probability distributions set at the introductory knowledge level.

J. 5('Understand 'hypothesis testing): Questions on hypothesis testing set at the introductory knowledge level.

Selected examples of questions used are shown in Appendix B. The results for the dependent variable J.1 were found using five-point Likert-type scale (Likert, 1932) with each answer mapped to the value range $R = [0, 1]$. The values for variables J.2 - J.5 are average scores derived from five questions and missing answers were marked as incorrect. The descriptive statistics for the experiment are summarized in Table 3 and Table 4 where the columns 'Pre-test scores' and 'Post-test scores' show the calculated values for mean (\bar{x}), median (m) and standard deviation (s) of the raw data collected, and the columns 'Differences' show the difference between the post-test and pre-test scores.

Standard significance testing was used to investigate the effect of the treatments on the independent variables J.1 to J.5. The null hypotheses were:

$H_{0,1}$: There is no difference between pre-test and post-test scores within experimental group (A) and control group (B).

$H_{0,2a}$: There is no difference in relative learning effectiveness between experimental group (A) and control group (B).

$H_{0,2b}$: There is no difference in absolute learning effectiveness between experimental group (A) and control group (B).

Table 3. Pre- and post-test scores for dependent variables

	Pre-test scores					Post-test scores				
	J.1	J.2	J.3	J.4	J.5	J.1	J.2	J.3	J.4	J.5
Group A										
\bar{x}	0.85	0.61	0.43	0.15	0.07	0.88	0.87	0.79	0.59	0.56
m	0.90	0.60	0.48	0.14	0.08	0.90	0.85	0.78	0.57	0.57
s	0.12	0.14	0.17	0.21	0.23	0.13	0.15	0.14	0.16	0.13
Group B										
\bar{x}	0.83	0.62	0.49	0.17	0.09	0.86	0.86	0.71	0.43	0.43
m	0.87	0.63	0.48	0.19	0.08	0.86	0.83	0.73	0.45	0.44
s	0.13	0.12	0.17	0.21	0.24	0.15	0.16	0.12	0.13	0.2

Table 4. Differences scores for dependent variables

	Differences				
	J.1	J.2	J.3	J.4	J.5
Group A					
\bar{x}	0.03	0.26	0.36	0.44	0.49
m	0.00	0.25	0.30	0.43	0.49
s	0.13	0.15	0.16	0.19	0.17
Group B					
\bar{x}	0.03	0.24	0.22	0.26	0.34
m	-0.01	0.20	0.25	0.26	0.36
s	0.14	0.15	0.15	0.18	0.23

Focusing on the experimental group (A), Table 5 shows the results using a two-tailed t-test for dependent samples. Column one specifies the variable, column two represents the Cohen effect size, d (Cohen, 1988), column three the degrees of freedom, column four the t-value of the study, column five the critical value for the significance value

$\alpha = 0.05$ and column six lists the associated p-value. Using the suggestions of Pfahl *et al.* (2004) testing for the normality assumption, analysis to detect outliers and the non-parametric tests of Wilcoxon and the Mann-Whitney U test were carried out for the hypothesis $H_{0,1}$ and for the hypotheses $H_{0,2a}$ and $H_{0,2b}$ respectively. It was found that no normal distribution of the variables could be assumed and that all the data lay within the ± 2 standard deviations around the samples' means. The non-parametric tests did not show any difference from the results of the t-tests. The equivalent results for control group (B) are given in Table 6.

Table 5. Results for 'post-test' versus 'pre-test' for group A

Variable	<i>d</i>	df	t-value	Crit. $t_{0.95}$	p-value
J.1	0.240	90	1.6180	1.662	0.1075
J.2	1.792	90	12.088	1.662	0.0001
J.3	2.312	90	15.594	1.662	0.0001
J.4	2.357	90	15.899	1.662	0.0001
J.5	2.623	90	17.693	1.662	0.0001

Table 6. Results for 'post-test' versus 'pre-test' for group B

Variable	<i>d</i>	df	t-value	Crit. $t_{0.95}$	p-value
J.1	0.214	89	1.4338	1.662	0.1534
J.2	1.697	89	11.384	1.662	0.0001
J.3	1.495	89	10.030	1.662	0.0001
J.4	1.489	89	9.9869	1.662	0.0001
J.5	1.539	89	10.324	1.662	0.0001

The results relating to hypothesis $H_{0,2a}$, which states that the difference between the post-test and pre-test scores of group A is not significantly larger than those of group B are given in Table 7. Table 8 shows for each dependent variable, the results when testing $H_{0,2b}$ using a two-tailed t-test for independent samples.

Table 7. Results for 'performance improvement' (group A versus group B)

Variable	<i>d</i>	df	t-value	Crit. $t_{0.95}$	p-value
J.1	0.000	180	0.000	1.653	1.0000
J.2	0.133	180	0.897	1.653	0.3710
J.3	0.903	180	6.072	1.653	0.0001
J.4	0.973	180	6.542	1.653	0.0001
J.5	0.742	180	4.993	1.653	0.0001

Table 8. Results for 'post-test improvement' (group A versus group B)

Variable	<i>d</i>	df	t-value	Crit. $t_{0.95}$	p-value
J.1	0.142	180	0.959	1.653	0.3390
J.2	0.064	180	0.434	1.653	0.6639
J.3	0.614	180	4.126	1.653	0.0001
J.4	1.098	180	7.379	1.653	0.0001
J.5	0.771	180	5.190	1.653	0.0001

4. Survey of Attitudes towards Statistics

4.1 Description

The SATS-36 (Schau *et al.*, 1995) was used to measure students' attitudes towards statistics, with students still divided into the experimental group (A) and control group (B). Students completed the SATS on the first and last day of the semester, where SATS is a 36 item scale with six subscales. Cronbach's alphas (α) were calculated for each subscale using the pre-test and post-test data. Scales are generally considered as reliable if α is at least 0.7 (Field, 2009). The subscales were:

- *Affect* has six items which assesses the students' feeling towards statistics (e.g., I will like statistics; I am

scared by statistics).

- *Cognitive competence* has six items which assesses students' belief concerning their ability to understand statistics (e.g., I can learn statistics; I will have trouble understanding statistics).
- *Value* has nine items which assesses students' beliefs about the usefulness of statistics in their lives (e.g., Statistics is worthless; Statistics should be a required part of my professional training)
- *Difficulty* has seven items which assesses students' beliefs about the difficulty of statistics (e.g., Statistics formulae are easy to understand; Statistics is complicated).
- *Effort* has four items which assesses students' beliefs about the amount of effort they would or did put in to the class (e.g., I plan to complete all of my assignments; I plan to work hard in my statistics course).
- *Interest* has four items which assesses students' interest in statistics (e.g., I am interested in being able to communicate statistical information to others; I am interested in using statistics).

All thirty-six items use a 7-point Likert (Likert, 1932) and higher SATS scores indicate a more positive attitude toward statistics.

4.2 Statistical Assumptions

The distributions of scores of both the experimental group (A) and control group (B) were analysed for normality assumptions, the Z-scores for skewness and kurtosis showed significant deviation from normality for three of the six SATS subscales, namely *cognitive competence*, *value* and *effort*. These subscales were then analysed using non-parametric statistics and the remaining three subscales, *affect*, *difficulty* and *interest*, were analysed using parametric and non-parametric statistics where it was found that the results did not differ. The non-parametric statistics are reported here with a significance value of 0.01 (two-tailed).

4.3 Data Analysis

Wilcoxon signed ranks were used to test if the students' scores on each of the SATS subscales changed during the semester, with the results summarized in Tables 9 and 10.

Table 9. Changes in students' attitudes from pre-test to post-test

SATS subscale		Group A		Group B	
		Pre-	Post-	Pre-	Post-
Affect	\bar{x}	4.33	4.76	4.51	4.54
	m	4.35	5.01	4.67	4.69
	s	1.11	1.14	0.94	0.98
Cognitive comp.	\bar{x}	4.98	5.54	5.03	5.23
	m	5.04	6.47	4.99	5.38
	s	0.97	1.05	0.93	1.03
Value	\bar{x}	4.67	4.59	4.55	4.59
	m	4.82	4.66	4.83	4.75
	s	0.94	0.92	0.99	1.12
Difficulty	\bar{x}	4.02	5.01	3.97	4.32
	m	4.03	5.00	4.02	4.43
	s	1.16	1.13	0.95	0.97
Effort	\bar{x}	5.96	5.71	5.76	5.54
	m	5.93	5.75	5.80	5.69
	s	1.15	1.04	0.94	1.04
Interest	\bar{x}	4.95	5.33	5.06	5.22
	m	5.04	5.41	5.12	5.32
	s	1.13	1.02	1.12	0.97

Table 10. Analyses of changes in students' attitudes from pre-test to post-test

SATS Subscale	Group A					Group B				
	z-score ^a	p-value ^b	Pre- α^c	Post- α^c	Effect size ($r = z/\sqrt{N}$)	z-score ^a	p-value ^b	Pre- α	Post- α	Effect size ($r = z/\sqrt{N}$)
Affect	-3.45	0.0006	0.92	0.89	0.257	-1.34	0.1802	0.90	0.89	0.0999
Cognitive comp.	-2.64	0.0083	0.95	0.92	0.1968	-2.66	0.0078	0.96	0.91	0.1983
Value	-1.56	0.1188	0.87	0.76	0.1163	-0.14	0.8887	0.85	0.74	0.0104
Difficulty	-3.01	0.0026	0.89	0.84	0.2244	-2.59	0.0096	0.91	0.81	0.1930
Effort	-3.03	0.0030	0.78	0.84	0.2258	-3.96	0.0000	0.83	0.74	0.2952
Interest	-2.17	0.0300	0.92	0.87	0.1617	-1.26	0.2077	0.90	0.86	0.0939

^aThis is the Wilcoxon's signed rank z-scores. They were calculated using $z = (T - \bar{T}) / \sqrt{\frac{n(n+1)(2n+1)}{24}}$ where T is either the sum of the ranked differences that were positive or the sum of the ranked differences that were negative, whichever is smaller, $\bar{T} = (n(n+1))/4$, and $n = (\text{number of paired scores} - \text{the number of difference scores which were zero})$ (Field, 2009).

^bThis is a two-tailed test.

^cThis is Cronbach's alpha (α).

5. Discussion of Results

5.1 Controlled Experiment

It can be seen from Table 5 that the experimental group A achieved statistically and practically significant results for the dependent variables J.2-J.5, whereas J.1 did not, when testing the $H_{0,1}$ hypothesis. Table 6 shows the results for the control group B when testing the $H_{0,1}$ hypothesis and for this group the variables J.2-J.5 were also statistically and practically significant, whereas the variable J.1 was not.

It can be seen from Table 7 that the hypothesis $H_{0,2a}$ can be rejected for the dependent variables J.2-J.5 and for these dependent variables the results support the direction of the expected relative learning effect. For the dependent variables J.1 and J.2 the results were not statistically or practically significant, and, for J.1 the result did not support the direction of the expected learning effect. Table 8 shows for each dependent variable, the results of testing the hypothesis $H_{0,2b}$ using a two-tailed t-test for independent samples with the hypothesis rejected for the variables J.3-J.5 but not for J.1 and J.2. The results for both $H_{0,2a}$ and $H_{0,2b}$ show the relative learning effect to be strongly supported both statistically and practically for the variables J.3 - J.5 but no or little relative impact was made on student's interest in statistics and probability, or, the student's understanding of the role of statistics and probability in engineering and generally when using the Portfolio Approach. On looking at Tables 3, 4, 5 and 6 it can be seen that J.1 and J.2 increased, especially regarding J.2 within each group. Also from Table 3 the dependent variable J.1's mean was already high, so increasing this value was relatively difficult. The disappointing aspect is that the active learning approach did not increase these two variables more when compared to the traditional approach.

5.2 SATS-36 Survey

First the internal consistency of the reliability of the test scores was reported in Table 10 using the Cronbach alpha test. For both the group A and group B results it can be seen that 9 of the 24 results were in the excellent range, 11 were in the good range and the rest were acceptable. Use was made of the Wilcoxon signed ranks test to determine how students' scores for each of the SATS subscales changed during the semester, with the results summarized in Tables 9

and 10. Four of the six subscales produced significant effects for the experimental group A and for three of the six subscales for the control group B. For both the experimental and control groups students had significantly higher cognitive competence, but only group A had significantly higher effect on the last day of the semester when compared to the first day. Both groups significantly overestimated the amount of effort they would need to complete their courses while both groups significantly underestimated the difficulty of their respective courses. The ratings for value and interest were not given as significantly different between the first and last day of the semester by both group A and group B.

Effect size is also shown Table 9, where r is considered as small around a value of 0.1, medium around a value of 0.3 and large around a value of 0.5. It can be seen that most of the six subscales can be considered as medium for group A and for group B all the subscales are tending to the small category except for one.

5.3 Remarks Concerning the Portfolio Curriculum Approach and Active Learning

From the controlled experiment it is clear that use of the portfolio curriculum approach give positive changes to the students' acquisition of statics and probability in the areas of elementary probability, probability distributions and hypothesis testing. The improvement in the learning of the latter is possibility due to its inherent difficulty, and so even small practice and experience of its theory and mechanics will give good returns. However, it is also clear that the introduction of the portfolio curriculum approach, and hence active learning, did not improve interest in nor general knowledge of, statistics and probability when compared to using the traditional approach and obviously this needs to be addressed through further research. It could be that the difference between the two approaches was not pronounced enough to give differential results. For example, if the Portfolio Curriculum approach emphasized the need to find information to solve a problem vs. information being provided by the instructor then perhaps an increase in general knowledge may have resulted.

From the SATS-36 results it may be concluded that the portfolio curriculum approach and the traditional approach used to teach the students both produced significant positive changes in students' attitudes towards statistics and probability. Both groups gave higher difficulty ratings for their course after taking the course compared to the beginning. It was also noticeable that the median values reported in Table 9 mostly increased between the pre- and post- testing for both groups with significantly positive changes for affect, cognitive competence, difficulty and effort reported for group A and for group B, cognitive competence, difficulty and effort.

A common criticism of active learning approaches is that they sacrifice content coverage for direct experience (Carlson & Winqvist, 2011). The portfolio curriculum approach used here did not. What was changed was the delivery method, with in the traditional course students gaining the course content from the instructor and reading and applying during lab sessions. In the portfolio curriculum approach the students acquired the course content from reading and interacting with fellow students and instructors. One of the benefits of the portfolio curriculum approach was that the teaching staff could interact much more with individual students than was possible through the traditional lecture approach. This gave an immediacy to the teaching in that students got answers quickly and directly. There is evidence in the literature that this kind of instructor behavior can increase the student affect (Mottet *et al.*, 2007) as is reported in Table 10.

Although encouraged by the findings of both the controlled experiment and SATS-36, some changes need to be recorded involving instructor attitudes and working. It was found that many more questions than 'usual' had to be answered for the portfolio curriculum approach and students were also found to work at different speeds. The questions from each group became quite similar and students had sometimes difficulty in explaining what exactly their question was about, so instructors had to show patience at times and be versatile in rewording a given student's question. There is a certain loss of control during the group activities although the students were mature enough and eager enough to get on with their tasks.

5.4 Validity and Ethics

The validity of the present work is now discussed. It is first recognized that interest in a topic and evaluation of a teaching session are difficult to measure, and to alleviate this problem the instrument for measuring variable J.1 was derived from measurement instruments that have already been successfully applied in similar kinds of studies, i.e., Pfahl *et al.* (2004) and Vennix (1990).

To alleviate selection threats when dividing the students into two groups, a randomization procedure was used. This together with the students' characteristics of a similar age and coming from a fairly uniform general education system gave reasonable assurance of minimum bias. Also, there were not changes in teaching staff throughout the courses, so reducing any 'selection history effect', and as none of the subjects left their respective group, there was no 'dropout interaction effect'. Students in each group were asked not to discuss their course with members of the other group to try

to reduce the 'diffusion or contamination' effect, but how effective this was is not known. Also it was mentioned to all students that each course was a legitimate method of acquiring knowledge and skills regarding statistics and probability, and one method of delivery was not better than the other. This was an effort to minimize any 'rivalry or resentment' threats.

After selection, any differences in the ability of the groups was captured by collecting pre-test scores. The subjects all came from a fairly traditional national teacher-centred background so it can be expected that they, as a sample, are representative of the general undergraduate population.

Regarding ethics, all students entered the experiment voluntary and were certainly not coerced or recruited by anyone in authority. Students were first assigned to groups and then given the choice, using a 'buddy-buddy' system for exchanging between groups. Less than 5% of the complete cohort changed groups. The data gathered, of course involved the students' results so confidentially and anonymity was used in any subsequent reports or other further dissemination. No teaching and learning time was substituted by research time, i.e. posted times for teaching and learning were completely used for those activities.

6. Conclusions

The empirical studies presented in this paper investigated a more active approach to learning through the use of the portfolio curriculum approach. It was found that this new delivery method helped students towards greater academic performance, though there is some doubt as to whether the current portfolio curriculum approach adds anything more to a student's general knowledge of statistics or to the student's interest in statistics and probability than the traditional approach. These two areas will be looked at in further research.

Acknowledgements

The authors are glad to acknowledge the useful discussions with Mr. Perry Hessenauer concerning the management of groups and Dr. Candace Schau for her permission to use the SATS-36 software.

References

- Adair, D., Jaeger, M. & Stegen J. (2011). Comparison of two-approaches when teaching object-orientated programming to novices. *International Journal of Engineering Education*, 27(5), 1027-1036. ISSN 0949-149X.
- Bosworth, K. (1994). Developing Collaborative skills in college students. In K. Bosworth and S. Hamilton (Eds.), *Collaborative learning: Underlying Processes and Effective Techniques*. San Francisco: Jossey-Bass. <https://doi.org/10.1002/tl.37219945905>
- Carlson, K. A. & Winqvist, J. R. (2011). Evaluating an active learning approach to teaching introductory statistics: A classroom workbook approach. *Journal of Statistics Education*, 19(1). <https://doi.org/10.1080/10691898.2011.11889596>
- Chiesi, F., & Primi, C. (2009). Assessing statistics attitudes among college students: Psychometric properties of the Italian version of the Survey of Attitudes toward Statistics (SATS). *Learning and Individual Differences*, 2, 309-313. <https://doi.org/10.1016/j.lindif.2008.10.008>
- Christie, M., & de Graaff, E. (2017). The philosophical and pedagogical underpinnings of Active Learning in Engineering Education. *European Journal of Engineering Education*, 42(1), 5-16. <https://doi.org/10.1080/03043797.2016.1254160>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences*. 2nd edn., Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Cramer S. (1994). Assessing effectiveness in the collaborative classroom. In K. Bosworth and S. Hamilton (Eds.), *Collaborative learning: Underlying Processes and Effective Techniques*. San Francisco: Jossey-Bass. <https://doi.org/10.1002/tl.37219945909>
- Elmore, P. B., Lewis, E. L., & Bay. M. L. G. (1993). Statistics achievement: A function of attitudes and related experience. *Annual Meeting of the American Educational Research Association*, Atlanta, GA.
- Felder, R. M., & Brent. R. (2016). *Teaching and Learning STEM: A Practical Guide*. Ch. 6 San Francisco: Jossey-Bass.

- Felder, R. M. & Brent, R. (2005). Understanding student differences. *Journal of Engineering Education*, 94(1), 57-72. <https://doi.org/10.1002/j.2168-9830.2005.tb00829.x>
- Felder, R., Brent, R. & Stice, J. (2002). National effective teaching institute workshop materials. 2002 American Society for Engineering Education Annual Conference, Montreal, Quebec, Canada.
- Felder, R. M., & Spurlin, J. (2005). Applications, reliability and validity of the index of learning styles. *International Journal of Engineering Education*, 21(1), 103-112.
- Feden, P., & Vogel, R. (2003). *Methods of Teaching: Applying Cognitive Science to Promote Student Learning*. McGraw Hill Higher Education.
- Field, A. (2009). *Discovering Statistics Using SPSS*. London: Sage Publications Ltd.
- Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, 28, 161-186. [https://doi.org/10.1016/S0361-476X\(02\)00015-2](https://doi.org/10.1016/S0361-476X(02)00015-2)
- Garfield, J. & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, 75(3), 372-396. <https://doi.org/10.1111/j.1751-5823.2007.00029.x>
- Giraud, G. (1997). Cooperative learning and statistics instruction. *Journal of Statistics Education*, 5(3). <https://doi.org/10.1080/10691898.1997.11910598>
- Harlow, L. L, Burkholder, G. J. & Morrow, J. A. (2002). Evaluating attitudes, skill, and performance in a learning-enhanced quantitative methods course: A structural modeling approach. *Structural Equation Modeling*, 9, 413-430. https://doi.org/10.1207/S15328007SEM0903_6
- Ingham, A., Levinger, C., Craves, A., & Peckham, V. (1974). The Ringelmann effect: Studies of group size and group performance. *Journal of Personality and Social Psychology*, 10, 371-384. [https://doi.org/10.1016/0022-1031\(74\)90033-X](https://doi.org/10.1016/0022-1031(74)90033-X)
- Johnson, D., Johnson, R. & Smith, K. (1998). Cooperative learning returns to college: What evidence is there that it works? *Change*, 30(4), 26-35. <https://doi.org/10.1080/00091389809602629>
- Keeler, C. M. & Steinhorst, R. K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education*, 3(2). <https://doi.org/10.1080/10691898.1995.11910485>
- Magel, R. C. (1998). Using cooperative learning in a large introductory statistics class. *Journal of Statistics Education*, 6(3). <https://doi.org/10.1080/10691898.1998.11910621>
- Manning, K., Zachar, P., Ray, G. & LoBello, S. (2006). Research methods courses and the scientist and practitioner interests of psychology majors. *Teaching of Psychology*, 33(3), 194-196.
- Meletioui, M. & Lee, C. (2002). Teaching students the stochastic nature of statistical concepts in an introductory statistics course. *Statistics Education Research Journal*, 1(2), 22-37.
- Miller, J., Trimbur, J. & Wilkes, J. (1994). Group dynamics: understanding group success and failure in collaborative learning. In K. Bosworth and S. Hamilton (Eds.), *Collaborative learning: Underlying Processes and Effective Techniques*. San Francisco: Jossey-Bass. <https://doi.org/10.1002/tl.37219945906>
- Millis, B. & Cottell, Jr., P. (1998). *Cooperative learning for higher education faculty*. American Council on Education, ORYX Press.
- Mottet, T. P., Parker-Raley, J., Beebe, S. A. & Cunningham, C. (2007). Instructors who resist "College Lite": The neutralizing effect of instructor immediacy on students' course-workload violations and perceptions of instructor credibility and affective learning. *Communication Education*, 56(2), 145-167. <https://doi.org/10.1080/03634520601164259>
- Pfaff, T. P. & Wienberg, A. (2009). Do hands-on activities increase student understanding?: A case study. *Journal of Statistics Education*, 17(3). <https://doi.org/10.1080/10691898.2009.11889536>
- Pfahl, D., Laitenberger, O., Ruhe, G., Dorsch, J. & Krivobokova, T. (2004). Evaluating the learning effectiveness of using simulations in software project management education: results from a twice replicated experiment. *Information & Software Technology*, 46, 127-147. [https://doi.org/10.1016/S0950-5849\(03\)00115-0](https://doi.org/10.1016/S0950-5849(03)00115-0)

- Prince, M. (2004). Does Active Learning Work? A Review of the Research. *Journal of Engineering, Education*, 93(3), 223-231. <https://doi.org/10.1002/j.2168-9830.2004.tb00809.x>
- R Core Team. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Schau, C., Stevens, J., Dauphinee, T. L. & Del Vecchio, A. (1995). The development and validation of the Survey of Attitudes Toward Statistics. *Educational and Psychological Measurement*, 55, 868-875. <https://doi.org/10.1177/0013164495055005022>
- Sizemore, O. J. & Lewandowski, G. W. (2009). Learning might not equal liking: Research methods course changes knowledge but not attitudes. *Teaching of Psychology*, 36(2), 90-95. <https://doi.org/10.1080/00986280902739727>
- Smith, B. & Macgregor, J. (1992). What is collaborative learning. in Goodsell, A., M. Mahler, V. Tinto, V., Smith, B. L. & J. MacGregor, (Eds), *Collaborative Learning: A Sourcebook for Higher Education* (pp. 9-22). University Park, PA: National Center on Postsecondary Teaching, Learning and Assessment.
- Stice, J. E. (1987). Using Kolb's learning cycle to improve student learning. *Engineering Education*, 77(5), 291-296.
- Tewksbury, B. J. (1995). Specific strategies for using the "jigsaw" technique for working in groups in non-lecture-based courses. *Journal of Geological Education*, 43(4), 322-326. <https://doi.org/10.5408/0022-1368-43.4.322>
- Vennix, J. A. M. (1990). Mental models and computer models: design and evaluation of a computer-based learning environment for policy making. Doctoral Dissertation, Katholieke Universiteit Nijmegen.
- Weltman, D. & Whiteside, M. (2010). Comparing the effectiveness of traditional and active learning methods in business statistics: Convergence to the mean. *Journal of Statistics Education*, 18(1). <https://doi.org/10.1080/10691898.2010.11889480>
- Wolfram Research, (2016). *Mathematica*, Version 10.3, Wolfram Research, Inc., Champaign, IL.

Appendix A - Teaching Syllabus

Elementary probability:

- Conditional probability
- Independence
- Probability properties
- Continuous random variables
- Probability density functions
- Cumulative density function

Probability distributions:

- Mean and variance of continual random variable
- Gaussian distribution
- Standard normal random variable
- Lognormal distribution
- Binomial distribution
- Poisson process
- Exponential distribution
- Normal approximations to the binomial and Poisson distributions

Hypothesis testing:

- Random samples, statistic, central limit theorem
- Decision making for a single sample
- Point estimation

- Unbiased estimator
- Standard error
- Null hypothesis
- Effect of sample size
- p-values
- Confidence interval on the mean
- Hypothesis testing on the mean
- Inference on the variance of a normal population
- Test for goodness of fit

Appendix B - A Sample of Questions

Dependent variable J.1

I consider it very important for engineering students to know as much as possible about computer programming. (1 = fully agree / 5. = fully disagree). Circle number below.

Agree 1 2 3 4 5 Disagree

Dependent variable J.2

Statistics and probability is important in the following areas of engineering:

Maintenance	True/False
Production	True/False
Research	True/False
Project Management	True/false

Dependent variable J.3

New designs for a wastewater treatment tank have proposed four possible shapes, four possible sizes, three locations for input valves and four locations for output valves. How many different product designs are possible?

Dependent variable J.4

The phone lines to an airline reservation system are occupied 50% of the time. Assume that the events that the lines are occupied on successive calls are independent. Assume that 10 calls are placed to the airline.

- What is the probability that for exactly three calls the lines are occupied/
- What is the probability that for at least one call the lines are not occupied?
- What is the expected number of calls in which the lines are all occupied?

Dependent variable J.5

A semiconductor manufacturer collects data from a new tool and conducts a hypothesis test with the null hypothesis that a critical dimension mean width equals 80 nm. The conclusion is to not reject the null hypothesis. Does this result provide strong evidence that the critical dimension mean equals 80 nm? Give some explanation with your answer.