

The Role of Prior Warning on Test Performance: How Effective Is It to Improve Students' Grades?

Rafiuddin Ahmed

College of Business, James Cook University, Australia

Copyright © 2015 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract Students' use of test information to prepare for a test in a controlled or supervised test environment has been examined in studies outside Australia. This paper reports the findings of the use of test information and its value, in terms of an improvement/decline in marks, in an actual test of an undergraduate subject taught at an Australian university. Using a questionnaire survey of students, the study finds that students overall don't perceive test information useful, there is no statistically significant difference in performance between known and unknown questions, students' scores improve from the use of information and in some instances the improvements are statistically significant between students with different characteristics. The paper contributes to our understanding of students' willingness to use information and the benefits of such information to study and perform for improved test scores. The study has implications for educators making test information available as a preferred practice or universities using it as part of a policy to improve student retention rates or supplement evaluation of students' learning.

Keywords Warning, Test Performance, GLM Method, Anxiety, Study Approach, Management Accounting

1. Introduction

Assessment is one of the most important tools for determining student' achievements from taught materials in university level subjects. A variety of achievement strategies are used in academic institutions and a variety of approaches are used to help students to optimize achievements from assessment items. This paper aims at examining the role of information in students' preparing for an examination and as part of an assessment policy.

One of the most commonly used measures of educational outcomes is evaluation. Evaluation allows measuring students' comprehension of taught materials [1], it tells an educator what and how students learn subject materials [2]. Testing is most commonly observed as an evaluation tool

[see for example, 3, 4, 5]. Kelly, Conant and Smart [6] label testing as an integral part of quality teaching. The literature on assessment in higher education reports two common types of testing: multiple choice and objective/essay type tests [7-9]. In any kind of testing, students are required to demonstrate certain skills and abilities. For example, students studying for professional accounting degrees are required to demonstrate sound analytical and conceptual skills [10]. Accounting academics also endorsed this view [11].

While testing is an educators' tool, tests are formal credentialing methods that students are required to complete to earn their degrees, it is a tool for them to showcase their learned skills in a controlled test environment. Students' responses to educators' testing demands have generated a large body of literature in higher education context. Students' use of strategies to respond to test demands in different subjects, and in different test types are empirically investigated previously [7, 8]. Quite often this test demands (from their instructors) is conveyed through forewarning, cues or formal communication such as sample questions, question types, and level of difficulties. The availability of such information is instrumental in studying for a test [1, 7, 8, 12].

Prior studies examining the effect of information or cues [13, 14] or forewarning [12] on students' extent of preparation for an exam and performance in different types of tests are diverse. Some studies have reported that the release of information (or warning) increases anxiety levels [1, 8, 15] while others report that the availability of such information motivates some students [1, 13]. These studies are mainly confined to the laboratory environment and have examined a number of subjects outside Australia. There are calls for more research into examining students' performance in business subjects in actual tests in a controlled test environment (supervised examination), in different university contexts, for ecological validity of the generalizations in a different context and to gain insights into students' behavior in real test environments [1, 8, 12-15]. This study responds to these calls.

This study is based on the data captured from students'

final examination records of marks in a third year management subject taught at a university in Queensland, Australia. The subject was part of a compulsory degree unit, offered to students enrolled in face to face delivery mode. Three objectives are set for this research (a) to examine students' perception about the value of examination information, (b) the retrospective use of examination information (cues or warning) for the preparation of the examination, and finally, (c) to gauge the efficacy of examination information for students grade achievements.

The study reports a number of findings. The first finding is that students don't perceive examination information useful and there is no difference in perception about the usefulness of exam information across groups profiled by age, gender, schooling background and major. Secondly, there is no statistically significant difference between known and unknown question sets. However, in the known question set, there are statistically significant differences in average marks within different age groups. The final finding is exam information reduced average marks in known questions set in five out of ten groups of students' groups (age, gender, schooling and major). Based on these three findings, the major conclusion of this paper is that examination information is marginally useful to majority of students to prepare for an examination and may be worthwhile releasing to them before an examination.

This paper adds to the growing body of the literature in higher education context by offering insights into performance dimensions across question types, and information availability for an actual examination. The study has also new dimensions to prior literature through the inclusion of within subject (student groups) comparison of performance following the availability examination information. Unlike past studies, this study uses a within-case (a single exam) design to compare students' performances in known and unknown question sets (two questions in each set). The study has enhanced the ecological validity of generalizations reached in the prior works in this area by corroborating the findings in prior studies [1, 7, 8, 13] and examined empirically the concerns raised in Broekkamp & Van Hout-Wolters [14] that examination information may be useful to prepare for an examination.

The remainder of the paper is organized as follows. In the next section the relevant research is reviewed. The research setting, data and instruments used, and the results are described in the following three sections. The summary and conclusions are then elaborated. The limitations and directions for future research conclude the paper.

2. Review of Relevant Literature

Literature on assessment and evaluation reports theory and evidence of different dimensions of test performance following the release of information and students' subsequent processing of the information to prepare for an upcoming examination. Weber and Bizer [12] found that the

availability of examination information triggers anxiety in students. Depending on the level of anxiety, students perform accordingly (e.g. less anxious students perform better than anxious students). Alpert and Haber [16] observed similar findings in their study. However, Ismail and Qayyum [15] observed no role of information on test performance of students in actual classroom setting (in actual tests). They further observed that availability of information is not related to students' examination performance at all, that is, information is proven useless following the completion of assessment grading. Burns [1] took this body of the literature even further. He found that students' test anticipations and engagements are dependent on their comprehension of the information provided to them. Depending on how students form expectations following the availability of information, they will either work hard to achieve or withdraw their efforts if they feel that the upcoming examination would be too challenging for them.

The literature reviewed above is silent on the blanket use of information for test performance. It is unclear as to the efficacy of information and its effect on actual test performance. Our intention here is to explore if information by itself is perceived useful before an examination, the characteristics of the students who expect the information to prepare for the test, and finally if there are differences in perceptions between students grouped by age, gender, schooling background and study major. Thus, our first hypothesis, in alternative form, is:

H_{1a}: Students perceive examination information useful

Following the availability of test information or the lack of it, students prepare for an examination. The literature on the relationships between students' test preparation strategies and preparedness for test is sparse. Hakstian [7] conducted two experiments on students enrolled in an educational research subject. The first experiment used 26 students as subjects who were split into three treatment groups (objective, essay, and essay and objective combined). The students were given cues about the test format, warned about the test format seven days before the test, and the exact type of test questions to be expected. The study explored the relationship between students' study approaches and performance in different types of examinations. The findings are that anticipation of test formats does not affect test performance, item types included, and study approaches used to study for the test. Hakstian's [7] second study found that students stress on factual texts (perhaps problem solving) and essay tests. They observed no difference in performance in test types based on study approach adopted. Foos [8] and Weiner [17] found that students increase efforts when they expect a difficult test than when they expect an easy test, and thus perform better in difficult tests than in an easy test. Ross, Green, Salisbury-Glennon, & Tollefson [18] investigated the linkage between the use of deep study strategies for items requiring deep level study strategies and exam performance. They observed that students performed better when they expected a deep level test item and reported studying at deep

level for that item. The use of deep study strategy resulted in an improved grade performance. They also found that students who used surface level study strategies for a surface level item performed poorer than the students following deep level study strategies. Fattah et al. [13] replicated this study in a psychology subject at an Egyptian university context. They observed the same findings. However, they have examined study strategy as a mediating variable between test expectation and performance.

While the literature reviewed above is conclusive in general that students adopt different study strategies, there is hardly any study that explored within a test item comparison when information is released selectively for some items and not for the others [except 12]. Unlike Abd-El_Fattah [13] and Ross et al. [18], we plan to use information as a mediating variable between students grouped by age, gender, schooling and major and test performance in different groups of examination items. It will be quite interesting to explore if students from different socio economic backgrounds process information differently and performance differently in an examination. Thus we have developed two hypotheses, in alternative forms, to test these premises.

H_{2a}: There are differences in average marks between known and unknown questions

H_{3a}: Students' grades improve from the use of examination information

3. Methods

3.1. Sample

The sample for this study was drawn from students enrolled in a third year management accounting subject at a business school in Queensland. Fifty students were enrolled in this subject, both internally and externally. An ethics approved questionnaire was administered to the students who attended the last lecture of the semester and mailed to students who did not attend the lecture. In total 37 completed questionnaires were returned but five of these could not be used for incompleteness. Thus 32 questionnaires were useable for this study (64% response rate). The collected responses represent a large enough sample in comparison to the population of students (50 in this instance). Small sample studies are quite common in education and evaluation literature but the data in some studies were not screened for normality tests before using the statistical procedures [see for example, 19, 20]. We have addressed this issue to improve the inferential importance (the generalization) of this study.

Statistical tests require assumptions of normality of a population of observations before test statistic and procedures can be used for exact or approximate inferences [21]. Weimer [22] and Howell [21] argue that a sample size greater than 30 is considered normal. Following these arguments, a number of normality tests (skewness, Kurtosis, QQ plot, and Shapiro-Wilk test) were done. We have found

that the dependent variables of the current study are normally distributed as shown in Table 1 below. Due to space limitations, other values are not presented in the table.

Table 1. Normality test values*

Dependent variables	Skewness (standard error)	Kurtosis (standard error)
Test information factor	-.818 (0.414)	-.460 (0.809)
Known questions	.031 (0.414)	-.890 (0.809)
Unknown questions	-.440 (0.414)	.167 (0.809)
Final examination marks	-.249 (0.414)	.417 (0.809)

*Skewness and Kurtosis Z-values between -1.96 to +1.96 represent normally distributed data Cramer and Howitt [23].

3.2. Instruments

The questionnaire had 16 questions in total in four different sections. The research instrument had sections on students' demographic and personal information. The informed consent section of the research instrument also sought students' permission to access their other assessment items. The students were asked to rate sixteen questions on a five point Likert type scale (1= don't agree at all and 5= strongly agree). The questions mainly inquired about the students' study approaches to preparing for the upcoming final examination, their plans of studies for two different types of questions, known (by topic/chapter) and unknown questions (by topic/chapters). The personal information sections mainly asked for students' age, gender, major, and prior accounting knowledge while the demographic information section asked about the students' prior schooling, and location backgrounds.

3.3. Measurement of Variables

We used four questions to measure warnings (or availability of exam information) following the use of the relationship between warning and exam performance by Foos [8], Ismail and Qayyum [15] and Weber and Bizer (Year). As four factors are multi-dimensional we have reduced these four factors into one factor using Principal Component Analysis (PCA). The reduced factors generated 32 regressions co-efficient. The alpha for this new factor is 0.76 (0.81 standardized).

We have used two separate policies on warning, unlike the approach used in Foos [8], Ismail and Qayyum [15], and Weber and Bizer [12]. The final exam comprised of four questions with equal weightings. It was a two hour long supervised examination. Students were not allowed to have formula sheets or notes to the examination. Following the classification schemes Schute [24] [see also 25], we used two sets of questions. Students were warned (cues provided) about two questions (treatment questions) and no cues about the other two questions (control questions). The profiling was expected to help us understand the differences in performance in the treatment (known questions), and control

(unknown questions) compare the efficacy of warning as a policy on exam information disclosure for future years. Other studies in the past used warning for an entire exam [8, 15], however, ours has focused on partial warning on 50% of the exam questions.

The third measurement variable is the contribution of information to students' marks in different questions, and thus overall performance in the test. We measured the contribution of information by a pretest-posttest design by comparing marks before and after including information as a co-variate, and then analyzing the differences in marks in a GLM method (LSD approach). Following Fattah's [13] approach, we have explored the mediating relationship between information, performance and four independent variables (age, gender, schooling and major). While age and gender have been used in past studies [1, 24, 26], we have added schooling background and major as two new variables.

4. Results

Our first hypothesis was aimed to capture students' perceptions about informational value of prior warning about the nature of questions (difficult or easy) in an examination. The null hypothesis that information is not useful is accepted at 5% level of significance ($F=1.42$, $p=0.1556$ (two tailed)). In a GLM model, we have then used the information

co-efficient as a dependent variable of four explanatory variables (age, gender, major and schooling background) to test four summary hypotheses. We did not find any significant relationship between perception, and age ($F = 0.053$, and $p= 0.094$), gender ($F = 0.007$, $p = 0.934$), schooling ($F= 0.0480$, $p= 0.628$), and major ($F= 0.046$, $p = 0.833$). That is, all null hypotheses are accepted, negating the perceived value of information in preparing for the final examination. However, further analyses were required to identify if there were differences in perceptions within student-groups. The pairwise comparison in this GLM model (LSD approach) is summarized in Table 2 below.

The results of the pairwise comparison in Table 2 across all independent variables are insignificant at 5% level of confidence. This implies students in different groups did not perceive information useful for the preparation of the upcoming examination. Even though the results are not significant, some insights into perception differences are observed. Female students perceived information more useful than their male counterparts. Students in other majors perceived information more useful than the students studying for an Accounting major. Students schooled outside Queensland tended to value exam information more than their counterparts schooled within Queensland schooling system. Finally, older students perceived test information more valuable than the other two groups of students (age group 1, 18-21 years).

Table 2. Perceived usefulness of information*

Variable	Grouping	Mean difference (p value)	F statistics (p value)*
Gender	Female vs Male	0.0507 (p= 0.781)	0.921 (p =0.352), Eta (0.058)
Major	Accounting vs Other major	-0.436 (p= 0.439)	0.632 (p= 0.439), Eta (0.040)
Schooling	1 vs 2	-0.204 (p=0.730)	0.297 (p= 0.747), Eta (0.038)
	1 vs 4	-0.545 (p= 0.457)	
	2 vs 4	-0.341 (p = 0.660)	
Age	1 vs 2	0.258 (p = 0.691)	0.389 (p= 0.685), Eta (0.049)
	1 vs 3	-0.358 (p = 0.575)	
	2 vs 3	-0.616 (p = 0.396)	

*Based on GLM contrast table, Eta represents R²

Table 3. Comparison of total marks in question sets

Variables	Unknown (average)	Known (average)	Differences
Age			
- 18-21(Group 1)	21.68	21.79	0.11
- 22-24 (Group 2)	18.00	17.85	-0.15
- 25- above (Group 3)	16.95	17.27	0.32
Schooling			
- Local (Group 1)	18.94	18.83	-0.11
- Within Queensland (Group 2)	20.41	20.60	0.19
-Outside Queensland (Group 3)	18.25	18.62	0.37
Gender			
- Female	18.93	18.14	-0.79
- Male	19.61	19.53	-0.08
Major			
- Accounting (Group 1)	19.05	19.24	0.19
- Other Major (Group 2)	19.51	19.41	-0.10

4.1. Marks in Two Question Sets

We have used a paired sample t-test for known questions set (questions 1 and 3) and unknown questions set (questions 2 and 4). The null hypothesis is accepted at 5% level ($F = 1.20$, $p = 0.20$). We have cross validated the calculations using a one-way ANOVA. In our ANOVA, the null is also accepted at 5% level of significance ($F = 1.05$, $p = 0.309$). We used the GLM method to compare the average marks in unknown and known questions sets. The pairwise comparison in this GLM model (LSD approach) is summarized in Table 3.

In the known questions set (questions 1 and 3), we have found statistically significant differences in mean scores between age groups 1 (18-21 years) and 3 (25 years and above) ($F = 4.763$, $p = 0.017$) and high degree of reliability of the differences (eta squared = 0.268 and observed power = 0.744). The overall schooling background revealed statistically significant mean difference between (5.447, $p = 0.035$). We have also observed significant mean differences between students schooled locally (group 10 and within the university's location state (group 2) (mean difference = 2.325, significant at 0.10 level $p = 0.088$).

Comparison of marks by students' study majors did not reveal any difference in marks within different student groups. The differences in marks between students in group 1 (Accounting major) and group 2 (other major including double major) were statistically insignificant ($F = 0.078$, $p = 0.782$) and had poor reliability measures (Eta squared = 0.003). The pairwise comparison (LSD method) revealed no significant differences in mean scores ($p = 0.782$ and mean difference of -0.386). The finding revealed students doing either a double major or other majors scored slightly higher than the students studying for an accounting major. Comparison of marks by gender revealed no significant difference at 5% level ($F = 0.039$, $p = 0.844$). The pairwise comparison (LSD method) revealed no significant differences in mean scores between male and female students at 5% level of significance ($p = 0.844$, mean difference of 0.271). Female students outperformed male students in the known questions.

In the unknown question set (questions 2 and 4), no significant difference in average marks between age groups were found. Marks of students in age group 1 (18-21 years) and 3 (25 years and above) did not differ significantly ($F = 2.026$, $p = 0.151$, reliability statistics, eta squared = 0.126 and observed power = 0.382). Marks by age groups did not differ either, mean difference between age groups 1 and 2 was 2.822 ($p = 0.109$), groups 2 and 3 was -0.038 ($p = 0.985$) and groups group 1 and 3 was 2.784 ($p = 0.124$). Though the differences are not significant between the groups, group 1 (age 18-21), age group 2 (21-25 years) and 3 (25 years and above) achieved highest marks in descending order.

Comparison of average marks of students in different schooling groups revealed a statistically significant difference in marks at 5% level of significance ($F = 9.475$, $p = 0.001$, reliability measures, Eta squared = 0.404 and the observed power = 0.966). The pairwise comparison (LSD method) revealed significant differences in mean scores between schooling groups 1 and 2 (7.901, $p = 0.000$) and schooling groups 2 and 3 (7.246, $p = 0.000$). The students educated locally performed best within the three groups. We did not find any significant difference in mean scores between students in groups 1 (locally schooled students and 2 (outside locality but within the state schooled students) (0.656, $p = 0.602$).

Comparison of marks by students' major did not explain differences in marks ($F = 0.072$, $p = 0.79$), small values of reliability tests also confirmed this (Eta squared = 0.002 and the observed power = 0.058). Between subject (accounting vs other majors) comparison revealed no significant differences between accounting (group 1) and other majors (group 2) ($F = -0.410$, $p = 0.790$). The finding is quite interesting in that students doing either a double major or other major scored slightly higher than the students studying for an accounting major. Comparison of gender did not reveal any significant differences in marks ($F = 1.183$, $p = 0.286$), small values of reliability statistics confirmed this finding (Eta squared = 0.0039 and the observed power is 0.183). The pairwise comparison revealed no significant differences in mean scores between the students in group 1 (male) and group 2 (female) (difference = -1.612, $p = 0.286$). The finding is quite interesting in that male students outperformed female students (1.612 differences in favor of male students) in the unknown question set.

4.2. The Effect of Information

We determined the impact of information on students' performance in the known questions (treatment) by comparing marks before and after the release of the test information. The table above shows that when information is used as a covariate, students' marks in known questions differed significantly at 5% level by students' age groups ($F = 5.103$, $p = 0.022$, $R^2 = 0.422$). Marks in other three dimension of the students' profile did not differ significantly at all. However, the marks differed between different groups within these student groups (gender, schooling and major). Table 4 below summarizes the results.

As no specific information was released for unknown questions, no comparison was made to determine the effect of information in unknown questions. It is highly unlikely to have any significant influence of information on students' marks.

Table 4. Effect of information on known questions

Variables	Total marks (Before info)	Total marks (After info)	Net change (information effect)	Gains/Losses in marks
<i>Age</i>				
18-21(Group 1)	21.68	21.79	0.11	Gained
22-24(Group 2)	18.00	17.85	-0.15	Lost
25- (Group 3)	16.95	17.27	0.32	Gained
<i>Schooling</i>				
Local (Group 1)	18.94	18.83	-0.11	Lost
Within Queensland (Group 2)	20.41	20.60	0.19	Gained
Outside Queensland (Group 3)	18.25	18.62	0.37	Gained
<i>Gender</i>				
Female	18.93	18.14	-0.79	Lost
Male	19.61	19.53	-0.08	Lost
<i>Major</i>				
Accounting (Group 1)	19.05	19.24	0.19	Gained
Other Major (Group 2)	19.51	19.41	-0.10	Lost

5. Summary and Conclusions

This study was aimed to examine the role of warning in the form of disclosure of test information. We have observed from past studies that students ask for examination information in the belief that such information may help them to prepare for an examination and also help them to choose appropriate study strategies, that is, deep versus surface approaches.

Our first objective was to determine if students perceive warning useful for developing study strategies for an upcoming final examination. The finding was that students did not perceive examination warning useful to prepare for an examination. Even though students demand for examination information, and the institution where this research is conducted has no formal policy on warning, we have conducted this research to test the real efficacy of warning (information about the final exam) on examination performance. The study also finds that there are no differences in perceptions between students grouped by age, gender, major and prior schooling background. Therefore, the conclusion is that even though students demand information about an upcoming test, the findings from the responses contradicted their claims for specific examination guidelines. The students may have possessed the information from the lecturer but may have ignored the information significantly, or it may be that the disclosed test items are difficult to comprehend in a controlled, timed exam setting, and also difficult to solve. Two of the known questions are taken from topics requiring higher order thinking and comprehension skills, and the use of deep level of study strategy. As the survey instruments were filled in before taking the actual examination, students' willingness rather

than the actual use of information was represented in the statistical results.

In order to get further insights, achievement marks in known (treatment) and unknown (control) questions sets were compared and contrasted independently. The key finding was that marks did not differ between the control and treatment groups (null accepted); however, pairwise comparison revealed differences in marks between groups in these two types of question sets. Younger students aged between 18-21 years (group 1) achieved the highest marks, followed by students aged between 21-25 years (group two) and students aged 25 and above (group 3). Thus, the overall conclusion is there is a negative relationship between age and achievement marks, that is, as the students grow older, their marks in an exam can decline. This may be attributed to students' declining abilities to grasp new concepts and ideas with the aging process. We have also explored the relationship between major and marks achieved in control and treatment groups and observed no significant effect of major on marks. However, the pairwise comparisons revealed students doing a double major outperformed mainstream accounting major students. In the subject, students' from five different double major or double degrees were enrolled. It may be the aptitude of these students that contributed to a slight difference in marks in favor of other majors. When schooling and marks were compared, significant differences by schooling background were observed. Students from out of Queensland performed worst followed by students from within the Queensland but outside the location of the University school catchment areas. The students from the local school suburbs (about 100 square Kilometer) achieved the highest marks. Finally, our analysis revealed no statistically significant difference in marks

between male and female students in both types of question sets. However, the pairwise comparison revealed female students outperforming male students in both known and the unknown segments of the exam. The female students usually worked harder in the subject, they attended most of the lectures, attempted most of the tutorial questions and other directed studies. These factors may have contributed to their higher marks.

The final objective of this study was to examine the effect of information on students' marks. Comparison of marks before and after the inclusion of information (as a co-variate) revealed age as the only group to have statistically significant different performance across different age groups. The pairwise comparison revealed five out of ten groups had improvements in marks after the effect of information was considered in the calculations. The average marks of statistically significant different marks before and after the inclusion of information effect was reduced from two to only one group of students (by age groups). However, overall, there were absolute marks improvements in five groups and reduction in marks in the other five groups, and the differences in marks were reduced after the effect of information was considered in the calculations.

Thus, we can infer that information (though very general) helped the students to study for the test and helped the students improve. It may be the students were less anxious after the receipt of information or were able to focus on important text material or that they used the examination guidelines seriously to study for the test. Even though the first hypothesis did not reveal significant differences in willingness to use test information, the students indeed used the information to study for the test, the comparison of marks before and after the inclusion of information effect (co-variate) confirmed this. The marks in the later (control) established that information about a test (cues or warning) made a difference to students' marks. Thus, the finding contradicts earlier works of Foos [8], Ismail and Qayyum [15], Weiner [17] that information does not make a difference to students' performances, and empirically validates the theoretical work of Weber and Bizer [12] that examination information is somehow useful to students.

This paper adds to the current body of the literature in three different ways. First of all unlike prior studies, the current study explores the role of warning or cues as an independent construct. The objective was to gauge the students' willingness to use cues or warning to prepare for a test. The second contribution is unique in that unlike prior studies, we have grouped the questions of a final examination into two groups, by the availability of information, a known question set and an unknown question set, each set comprising of two questions. Comparison of marks of students grouped by major, gender, schooling background and age revealed students' performances and differences in marks in known and unknown questions. Finally, we have shown that information can be quantified and its effect on students' can be determined (in absolute terms) which may be useful for teaching and assessment practices.

6. Limitations and Further Research

This research is based on a study of a single subject taught at an Australian university. Therefore, the generalization applies to one subject area only. Samples can be drawn from classes with larger enrolments (e.g. first year Accounting, Business Statistics etc. where at least 200 students normally enroll in any academic year) to replicate this study and explore the reliability and validity of the conclusions reached in this paper. Financial Accounting and Business Statistics taught at the institution where this research was carried out have similar curriculum and rigor. Only four questions on students' willingness to use test information were used to test the efficacy of information use. More questions can be added to overcome the shortcomings of limited number of questions used here. We have examined the effect of test information on students' anxiety levels and possible responses (motivation to study or withdrawal from studies), and test performance. Other variables may be included to explain the reasons for differing test performance of different student groups used in this study.

Only test of proportions, t-tests and one-way ANOVA tests were performed. The use of other statistical tests such as a regression analysis, ANCOVA or MANCOVA may be used in future studies to improve the reliability and validity of the results reported in this paper. The release of information to students is used as a proxy for students' actual use of examination information which was not followed up by another survey. It may be worthwhile following up the actual use of information after the completion of grading the exam. Finally, the study can be replicated in other grade levels such as first year, second year or post graduate levels for further insights into students' actual and intended use of such information.

Acknowledgements

This paper is an ethics approved research project conducted by the author (Ethics Approval # H4872). An earlier version of this paper was presented at the weekly seminar of James Cook University and benefited from the constructive comments from the participants at the seminar. The paper was also presented at the 29 WBI Conference held in Sydney (24-25 November 2015) and benefited from the constructive comments from Professor Lewlyn Lester Raj Rodrigues.

REFERENCES

- [1] Burns, D.J., Will I Do as Well on the Final Exam as I Expect? An Examination of Students' Expectations. *Journal of the Scholarship of Teaching and Learning*. 8(3): p. 1-19. 2008
- [2] Lewis, R., P. Berghoff, and P. Pheaney, Focusing Students: three approaches for learning through evaluation. *Innovative*

- Higher Education. 23(3): p. 181-196. 1999
- [3] Souchal, C., et al., Assessing does not mean threatening: The purpose of assessment as a key determinant of girls' and boys' performance in a science class. *British Journal of Educational Psychology*. 84(1): p. 125-136. 2014
- [4] Veldhuis, M., et al., Teachers' Use of Classroom Assessment in Primary School Mathematics Education in the Netherlands. *Cadmo*. 2013
- [5] Crooks, T.J., The impact of classroom evaluation practices on students. *Review of Educational Research*. 58(4): p. 438-481. 1988
- [6] Kelley, C., J. Conant, and D. Smart, Examination procedures in marketing education: Profiling marketing's master teachers. *Developments in marketing science*: p. 263-267. 1989
- [7] Hakstian, A.R., The effects of type of examination anticipated on test preparation and performance. *The Journal of Educational Research*: p. 319-324. 1971
- [8] Foos, P.W., Test performance as a function of expected form and difficulty. *The Journal of Experimental Education*. 60(3): p. 205-211. 1992
- [9] Scouller, K.M. and M. Prosser, Students experiences in studying for multiple-choice question examinations. *Studies in Higher Education*. 19(3): p. 267-279. 1994
- [10] Anderson, A., Perspectives on education: Capabilities for success in the accounting profession. 1989
- [11] Sundem, G.L., The accounting education change commission: Its history and impact: Accounting Education Change Commission and American Accounting Association. 1999
- [12] Weber, C.J. and G.Y. Bizer, The effects of immediate forewarning of test difficulty on test performance. *The Journal of general psychology*. 133(3): p. 277-285. 2006
- [13] Abd-El-Fattah, S.M., The effect of test expectations on study strategies and test performance: a metacognitive perspective. *Educational Psychology*. 31(4): p. 497-511. 2011
- [14] Broekkamp, H. and B. Van Hout-Wolters, Students' adaptation of study strategies when preparing for classroom tests. *Educational Psychology Review*. 19(4): p. 401-428. 2007
- [15] Ismail, Z. and R. Qayyum, The effects of immediate forewarning of test difficulty on test performance in applied educational context. *International journal of business and social science*. 2(1): p. 132-139. 2011
- [16] Alpert, R. and R.N. Haber, Anxiety in academic achievement situations. *The Journal of Abnormal and Social Psychology*. 61(2): p. 207. 1960
- [17] Weiner, B., *Human motivation: Metaphors, theories, and research*: Sage. 1992
- [18] Ross, M.E., et al., College students' study strategies as a function of testing: An investigation into metacognitive self-regulation. *Innovative Higher Education*. 30(5): p. 361-375. 2006
- [19] Oxford, R., et al., Factors affecting achievement in a satellite delivered Japanese language program. *American Journal of Distance Education*. 7(1): p. 11-25. 1993
- [20] Howsen, R.M. and S.E. Lile, A comparison of course delivery methods: An exercise in experimental economics. *Journal of Economics and Finance Education*. 7(1): p. 21-28. 2008
- [21] Howell, D., *Statistical methods for psychology*: Cengage Learning. 2013
- [22] Weimer, R.C., *Statistics*. Dubuque, IA: Wm. C, 1993, Brown Publishers.
- [23] Cramer, D. and D.L. Howitt, *The Sage dictionary of statistics: a practical resource for students in the social sciences*: Sage. 2004
- [24] Shute, G.E., *Accounting Students and Abstract Reasoning: An Exploratory Study: a Research Report*: American Accounting Association Sarasota, Florida. 1979
- [25] Davidson, R.A., Relationship of study approach and exam performance. *Journal of Accounting Education*. 20(1): p. 29-44. 2003
- [26] Beyer, S., Gender differences in the accuracy of grade expectancies and evaluations. *Sex Roles*. 41(3-4): p. 279-296. 1999