

Improving feedback in large classes: application of task evaluation and reflection instrument for student self-assessment (TERISSA) in a unit on business statistics

Jennifer Harlim

School of Electrical & Computer Engineering, RMIT University, jennifer.harlim@rmit.edu.au

Ashton de Silva

School of Economics, Finance & Marketing, RMIT University, ashton.desilva@rmit.edu.au

Iouri Belski

School of Electrical & Computer Engineering, RMIT University, iouri.belski@rmit.edu.au

This paper presents findings and lessons learnt from the implementation of Task Evaluation and Reflection Instrument for Student Self-Assessment (TERISSA) in a unit on Business Statistics with over 500 students enrolled, which was conducted in semester one of 2009. Four out of six unit tutors used TERISSA in their tutorials, involving 231 students in the application of TERISSA. It has been found that students who used TERISSA gave an overall higher Good Teaching Scale (GTS) score in the formal RMIT Course Experience Survey (CES), than the students who did not use TERISSA. Students who used TERISSA during three tutorial tests also achieved significantly better marks in their final examination. Results of additional staff and student surveys revealed that by employing TERISSA, the unit coordinator and the tutors were able to gain valuable feedback on students' progress in the unit. Over 40% of the surveyed students agreed that TERISSA has helped them to identify the learning areas that required their immediate consideration. Furthermore, a perceptible shift in student attitude on the issue of whether the lecturers/tutors or the individual students are responsible for the feedback on learning has been identified. Lessons learned as well as ideas on how to use TERISSA in large classes effectively are also discussed in this paper.

Keywords: student feedback, self-assessment, reflection, staff feedback, large classes

Introduction

Providing timely feedback to individual students is a challenge for many educators. Given class sizes are increasing, it is also difficult for a coordinator to gauge whether a course suits the needs of a diverse student body that is changing year after year. Traditional methods, such as individual consultations, are still useful but are no longer practical for large classes. Therefore, alternative forms of feedback that fulfil the needs of both students and educators are necessary to ensure appropriate learning outcomes can be achieved.

Past studies have evaluated different strategies aimed at improving feedback in large classes. Educators reported on successful use of computer technology, web-based tools, blogs and clickers (Lincoln, 2008; Goldman, Cohen, & Sheahan, 2008; Anderson, Anderson, VanDeGrift, Wolfman, & Yasuhara, 2003; Smaill, 2005). Boud (1995) highlighted the nature of learning as socially constructed, where one learns by responding to the opinions of others. Therefore, one of the ways students are engaged in large classes is through peer assessment (Vickerman, 2009; O'Shea & Bigdan, 2008; Ballantyne, Hughes & Mylonas, 2002; Cooper & Robinson, 2000).

O'Shea and Bigdan (2008) redesigned The Biggest Loser competition in order to introduce peer learning to engineering undergraduates. They found that the competition had a statistically significant impact on student learning outcomes. While O'Shea and Bigdan (2008) presented the impact of peer learning on performance outcomes, they did not mention the opinions of the students and staff involved in the scheme.

Vickerman (2009) found that the reactions of students engaged in formative peer assessment were mixed. Only 55% of the 90 students surveyed thought that their knowledge and understanding of the subject was enhanced because of the peer assessment process. The rest of the students found it difficult judging their peers' work and preferred their tutor's intervention (Vickerman, 2009).

Ballantyne et al. (2002) sought student and staff opinions to improve peer feedback implementation. They found also that the experience left students with mixed reactions. In particular, 37% of students felt that they do not need to be involved in assessment and 44% felt that the tutor has the sole responsibility in assessment (Ballantyne et al., 2002). They also found that "student learning improved significantly through the use of peer assessment" (p. 436). However, they have also discovered that peer assessment took significant time and effort from the tutors.

While peer assessment seemed to work in enhancing learning quality, Bonwell and Eison (1991), Boud (1995) and Blayney and Freeman (2008) believed that engaging students to be active learners who are responsible for their own learning is more important. Bonwell and Eison (1991) argued that to ensure that students become active learners they must be engaged in activities that force them to think. The role of reflection in enhancing meta-cognitive skills has been discussed by Dewey (1933). Furthermore, McDonald and Boud (2003) discovered that engaging students in self-assessment can significantly impact the outcomes of their learning. However, this research focused on the impact of self-assessment whereby students are engaged in the assessment process at high school level.

Our study instead focuses on the impact of a meta-cognitive self-assessment tool: Task Evaluation Reflection Instrument for Student Self-Assessment (TERISSA) on feedback provided to students in large classes. TERISSA was developed by Belski (2007 & 2009). Most reflection tools such as journals and portfolios discussed by Beveridge (1997), Woodward (1998), and Stuart (2004), require a student to reflect after the task is completed. TERISSA engages students in evaluation of task complexity before and after they have completed the task. Reflection before a problem is solved makes TERISSA unique. Belski (2009) argued that the reason TERISSA works is that thinking is stimulated when a discrepancy between the initial and final evaluations of the task complexity level occurs. This discrepancy makes the students experience 'a state of doubt, hesitation, perplexity and mental difficulty, in which thinking originates' (Dewey, 1933, as cited in Belski, 2009). This experience engages the student in reflecting to figure out how he/she can resolve this doubt, hesitation, perplexity and mental difficulty. Belski (2009) posits that students gain feedback of their own learning for themselves. Belski (2007) noticed that over 60% of surveyed students from his class in electronic engineering found TERISSA useful in their learning. In a follow-up study that involved non-engineering courses with different applications of TERISSA, over 50% of students found TERISSA useful in providing immediate feedback and helpful in identifying their knowledge gaps (Belski, 2009). In addition, Belski (2009) found that the use of TERISSA can lead to the improvement in Good Teaching Scale (GTS) scores. This is the main reason why TERISSA was utilised in our study.

Past studies have evaluated the use of technology and peer-assisted learning in large classes to enhance learning quality, but not through the use of an individual meta-cognitive self-assessment tool. Some of these studies have only taken into account either the feedback from the students and the staff involved or learning outcomes, but our paper sought to do both. This study attempts to answer the following questions:

- Does the use of TERISSA in a summative assessment context influence students' satisfaction with educational feedback even in a large class?
- Does engaging students in meta-cognitive reflection result in better learning outcomes?
- How can meta-cognitive self-assessment be deployed in large classes more effectively?

Methodology

TERISSA was trialled in the first semester of 2009 in the Business Statistics course with over 500 first year students enrolled. Four out of six unit tutors were utilising TERISSA. Each tutor taught 2 classes, and used TERISSA in only one of his/her classes – the class with larger student numbers. Thus, all 4 TERISSA classes may be considered as the treatment group involving 231 randomly selected students, whereas the four non-TERISSA classes may be considered as the control group.

TERISSA was utilised during three class tests (summative). The first implementation took place in Week 5 of the course. Each test includes two main questions which were broken up into smaller parts. Each student was given two TERISSA sheets (refer to Appendix A), one for each test question. The students were asked to evaluate the complexity of the task before they started the test and after they had done it. While TERISSA is primarily an individual reflection tool, student responses also helped educators with evaluating students' knowledge and progress in the course and with making adjustments to the course (Belski, 2007). Therefore, the responses were collected and analysed. Belski (2007) also recommended sharing the students' responses with the students to enhance student engagement.

Prior to the first class test, the tutors were briefed on how they could introduce TERISSA to their students. It was decided that the treatment group would be briefly introduced to the methodology during the revision session prior to Test 1. This revision session was conducted immediately before the first class test. To accommodate the use of TERISSA, students of the treatment groups were given an extra 10 minutes to fill in the TERISSA template: 5 minutes to carry out the initial evaluation before the test, and another 5 minutes to carry out the final evaluation after the test. In Belski's (2007) study, students' evaluations of the task complexity were shared immediately during tutorials. In our case, to accommodate the time required to compile and analyse both quantitative and qualitative data obtained from the TERISSA responses, we decided to share the responses with the students involved two weeks after the test.

For the second and third tests, the way in which TERISSA was implemented changed as it was found that many students used the extra 10 minutes given for the TERISSA to complete their test. In Week 9, during the second test, the initial evaluation was embedded into the reading time, and an extra 5 minutes was given for the final evaluation after the test paper had been collected. The same procedure was followed during Test 3, which occurred in Week 11. TERISSA responses were not formally shared with students for the second and third tests due to the timing clash with Test 3 and the final examination respectively. However, these analyses were provided to the coordinator and tutors involved.

To investigate the impact of TERISSA on students' satisfaction with educational feedback, formal RMIT CES surveys were used. The CES questions allow students to select from 5 Likert-scale responses ranging from strongly disagree to strongly agree. To calculate the GTS score only responses of agree and strongly agree to 6 particular questions¹ are averaged. These six questions pertain to feedback, student engagement and the quality of the teacher. There were 131 students who used TERISSA and 100 students who did not use TERISSA participating in the formal RMIT CES surveys. The GTS scores for the students involved in the trial were compared.

To explore the impact of meta-cognitive reflection using TERISSA on students' learning, their performance in the final examination was compared. Out of 231 students who used TERISSA, only 171 students sat for all three tests. The latter were considered as TERISSA users for the Independent Samples T-Test. The rationale behind a choice of only 171 students as TERISSA users comes from the work of Boud (1995) and Brockbank, McGill and Beech (2002). They posited that while a single reflection may lead to immediate improvement, values and perceptions are not changed. A lasting change can occur when underlying values and perceptions are changed through continual reflection. Therefore, the number of exposures of students to reflection activities can be a crucial factor in our study. These students' results were compared against the results of 256 students who did not use TERISSA at all.

To evaluate how TERISSA implementation can be improved in large classes, students and staff involved in this study were surveyed. The students were asked 4 questions using 5 Likert-scale responses on whether TERISSA had helped them in areas of their studies and if they would use TERISSA again in the future. Two additional open-ended questions were also included: specifically, which aspect of TERISSA was useful and how implementation can be improved. The staff survey included 13 questions using 5 Likert-scale responses. The staff survey also included a question about whether they would use TERISSA again and in what context. Staff were also given the opportunity to comment more on TERISSA through 4 open-ended questions. For more details on the surveys administered refer to Appendix B. These surveys are similar to those used in Belski (2009). A total of 78 students, the 4 tutors and the course coordinator participated in the surveys.

¹ The GTS statements include: The teaching staff in this course motivate me to do my best work. The staff put a lot of time into commenting on my work. The staff made a real effort to understand difficulties I might be having with my work. The teaching staff normally gave me helpful feedback on how I am going in this course. My lecturers were extremely good at explaining things. The teaching staff work hard to make this course interesting.

Results

GTS Scores

The biggest challenge for the coordinator in Business Statistics was to provide quality feedback to over 500 students who were enrolled in the course. In the past, the feedback components of the GTS had the biggest negative impact on the GTS score of this course. In an aim to increase the GTS score and to ensure feedback quality is improved despite the high enrolments in the course, a variety of changes were implemented by the coordinator of the unit. Amongst many, one additional assistant tutor was trialled in tutorial groups of class size exceeding 60 enrolments. Feedback on how the student is generally coping with course assessments was also given via e-mail in week 9. Students in four tutorials used TERISSA during three tutorial tests. As a result, the overall GTS scores for the unit have increased from 34% (2008) to 46% (2009).

To investigate the extent of TERISSA's influence on the GTS, a comparison of scores of students' responses when TERISSA was used and when TERISSA was not used by the tutors is shown in Table 1. One hundred and thirty one students who used TERISSA and 100 students who did not use TERISSA participated in the CES surveys. The GTS average of the 4 tutors involved when TERISSA was not used was 43% whereas their average GTS score when TERISSA was used was 50%. The differences in the GTS of Tutor 2 and Tutor 3 are negligible but we observed a significant increase in the GTS in the two classes of comparable sizes (Tutor 1 and Tutor 4).

Table 1. Comparisons of GTS Scores of TERISSA Group (T) and Non-TERISSA Group (NT)

	GTS – T (No. of Students)	GTS – NT (No. of students)	GTS Change
Tutor 1	61 (24)	40 (22)	21
Tutor 2	40 (37)	43 (20)	-3
Tutor 3	58 (40)	60 (28)	-2
Tutor 4	39 (30)	30 (30)	9
GTS Averages	50 (131)	43 (100)	7

Therefore, we can conclude that TERISSA positively influenced students' learning experience and resulted in a sizable improvement of student satisfaction with educational feedback.

Impact on students' performance

The final exam results of students who were exposed to TERISSA were compared against those who had never used TERISSA. We wanted to investigate if there was an impact when students were engaged in meta-cognitive reflection on their learning outcome even when TERISSA was used only three times during a twelve week course. Results of 171 students who used TERISSA and sat for all three tests and 256 students who did not use TERISSA at all were compared.

Students who participated in all three tests with TERISSA had a statistically significant better exam performance than those who did not use TERISSA. The average exam mark of the TERISSA group was 50.81 (SE=1.27). This is in contrast to the non-TERISSA group average of 47.08 (SE=1.10). An Independent Sample T-Test showed this difference as statistically significant ($p=0.015$). Students who were in the TERISSA group had an average of above the passing grade, while the control group did not. We conclude in our study there is a positive impact of meta-cognitive reflection on learning outcomes.

Students' opinion

After the use of TERISSA in three class tests, students were surveyed to evaluate the implementation of TERISSA. We were concerned with how TERISSA implementation was viewed and how it can be enhanced. A total of 78 students out of 231 students responded to a TERISSA survey carried out in week 11 of the course. Mixed reactions were visible when students were asked if TERISSA had helped them to develop better judgement of the task given and if TERISSA had provided them with immediate feedback on

their knowledge of the course. Only about 30% responded positively to these questions. When asked if TERISSA had helped them to identify the learning areas which required their immediate consideration, 41% of students surveyed agreed/strongly agreed. Encouragingly, 18% actually agreed/strongly agreed that they will continue to use TERISSA while solving problems.

The comments provided by the students helped us to gain a deeper understanding of some of their responses in the survey. They were specifically asked to comment about the useful aspects of TERISSA and their responses have been grouped thematically below.

The students who responded more positively towards the exercise found TERISSA to be useful in terms of making them think or read the question properly before attempting the task. One student commented,

“Simply to take the objective approach, reading question thoroughly prior to attempting to answer.”

Other students also commented:

‘Being forced to stop and analyse the question.’

‘Thinking about how hard a problem really is and what to do about it rather than panicking.’

‘Makes me read the question carefully.’

‘Made me think about the question before I undertook it.’

The initial evaluation aspect of TERISSA actually helped them in the actual test process itself. The students were also able to focus their attention on areas that they might be having issues with during the test with one student commenting, “*Identifying areas that I will struggle with so I could focus my attention on that even more.*” Therefore, the use of TERISSA during a test process does help the student in tackling the task given.

In addition, students also commented that they were able to see which areas they need to improve on:

‘Identifying things that need improvement.’

‘Showing which areas you need to work on immediately.’

‘Shows me which parts of the course I need to pay more attention to.’

‘Help identifying what techniques I'm short of or what problems I'm facing when revising for exam.’

These corroborate their responses when asked if TERISSA had helped them to identify the learning areas which required their immediate consideration. In addition, TERISSA has helped with their self-assessment. This is evident when students commented that using TERISSA has helped them in:

‘Acknowledging faults and difficulties.’

‘Assessing myself after I've answered the question.’

‘Helped me realise how I evaluate a problem before and after I completed it.’

Most encouragingly, because of using TERISSA, students realised that learning is their own responsibility. Some of the students commented:

‘It does bring you to the realisation that to do better you have to work harder.’

‘It showed me that even if I think I know how to do it, I should still do an example problem because I always found the problem harder after I did it.’

‘Do enough work outside class and try to complete most as much as I can when outside class.’

‘Identifying areas I need help with, knowing I need to study more.’

‘Forces you to think about why you are struggling with any questions.’

We also sought the students’ opinion on how implementation can be improved. Some students who reacted negatively towards the exercise suggested that TERISSA should be removed completely. However, some of the students who found TERISSA helpful suggested that TERISSA should not be implemented during tests with one student commenting,

“Maybe not just after the test because the operational aspect of the TERISSA is making it a bit difficult for students to complete because they are fighting with time to complete the test.”

Other negative comments indicated that some students felt the scheme was a waste of time and distracting when they were trying to deal with timed tests.

Some students also agreed that the presentation of feedback can be further improved. One student commented:

‘Needs to focus more on the individual. Feedback from TERISSA is too long after the test, I’ve already forgotten what it was about.’

Another student commented:

‘Summaries made available online of main problems for future students to keep in mind.’

When TERISSA is used in summative assessment and feedback, there is a need for it to be timely and easily accessed for later reference. One student went on further to request that TERISSA be used as a comparative tool of the class’ performance by commenting,

“Perhaps an average marks graph. As in how many people got high marks and compare this to perceived difficulty.”

Some students also commented:

‘[TERISSA] was simply for the teachers. I didn’t really look at it as a perspective on how it can help me.’

‘Discuss the results with the teacher and discuss written work on TERISSA rather than as an individual self-help strategy.’

‘... put some practice and maybe the feedback can be given individually.’

‘Firstly, there needs to be more consultation times with test and lecturer and more feedback.’

These are indications that students would like the tutors and the lecturer to be more involved in the feedback process.

Academics' opinion

Opinions from the academics were also sought on the use of TERISSA in their classes. The subject coordinator and the four tutors involved were surveyed at the end of the course. Like the students, the academics involved had a mixed reaction. All but one tutor strongly agreed or agreed that TERISSA helped them to gain valuable feedback on the students' progress in their course. Despite this, only one of those surveyed has actually used the feedback to improve course delivery during the semester or changed the way they have conducted the course. Those who indicate that they did not change the course were the tutors surveyed.

Some of those surveyed also indicated that using TERISSA has taken a significant amount of their time. This is not surprising as TERISSA was used during the class tests which were conducted during the tutorials; therefore it was most of the tutors involved who felt this. Three of the academics surveyed agreed that students who have used TERISSA have been able to enhance their learning but only one agreed that the use of TERISSA has improved the students' grades. This is in contrast to what actually happened. Again most of those who strongly disagreed/disagreed or neither disagreed or agreed are the tutors.

Qualitative comments were also sought from the academics involved in the trial for TERISSA. It is clear from some of their comments that there is a mixed reaction to the utilisation of TERISSA in their classes. Their feedback highlighted a need to introduce and promote TERISSA to both staff and students more effectively so that both groups understand and are convinced of the benefits of this learning tool.

A tutor commented:

'It's difficult to get students to see the benefit of using TERISSA ... somehow there needs to be an interesting/fun way to show students how useful it can be.... for staff, I think an introductory session is needed so as to clearly educate them about how to best implement TERISSA and the purpose/usefulness of TERISSA.'

Another tutor also commented:

'I think this method should be explained thoroughly to staff and students as well so that it can be used to its true potential. Addressing it as a feedback process is important. More importantly is the outcome of this – the improvement process of the teaching and learning approach. How far the improvement process goes will play a role in deciding TERISSA in affecting the teaching scores.'

Some of the tutors added:

'Discussing the TERISSA results with students takes up valuable time during the class and generally students are not really interested.'

'The feedback was provided too late after their tests, so students don't remember the questions and problems they had, so the feedback is therefore not as useful.'

'... to benefit the students, I think the presentation of the TERISSA feedback needs to be more immediate.'

These highlight that the tutors perceive the need for timely feedback as crucial in implementation, which indicates that the tutors understand and only recognize TERISSA as a summative tool.

Discussion

GTS scores

We observed a general increase in the GTS scores when the four tutors used TERISSA. The impact of TERISSA on the average GTS scores is consistent with findings of Belski (2009). However, when comparing each individual GTS score of the four tutors, we found a slight variation of results. Those of

comparable respondent groups (Tutor 1 and Tutor 4) had a significant increase in their GTS scores. On the other hand, Tutor 2 and Tutor 3 had a negligible decrease in their GTS when they used TERISSA. We hypothesise that this decrease may have been caused by differences in the number of students who responded to the CES survey in the TERISSA and non-TERISSA group. We observed that when the number of respondents in the non-TERISSA group was nearly double that of the TERISSA group, the GTS scores were lower. It is known from past RMIT survey results that larger class sizes usually result in lower GTS scores.

The non-consistent results between each tutor group can also be explained by the lack of sharing of the TERISSA feedback with students. Since the GTS scores are made up of a mix of feedback, engagement and teacher quality questions, the loss of engagement opportunity caused by the disruptions in sharing of the TERISSA results may have affected our data. In addition, quality of the teaching staff may also have an impact.

Impact on students' performance

Our results show a statistically significant improvement in performance of students when TERISSA was used. The improved learning outcome is consistent with findings from McDonald and Boud (2003) and theories from Boud (1995) and Brockbank et al. (2002) who argued that reflection leads to improvement of learning outcomes. The only difference between the TERISSA and non-TERISSA groups is the use of TERISSA. It can be proposed that students who were exposed to TERISSA, although it was for three times only during the course of twelve weeks, were given three opportunities to assess and reflect on their weaknesses and address them. Some of the students in the group of 171 students took this opportunity and were able to adjust their revision to cover their knowledge gaps. It can be posited that the use of meta-cognitive reflection, where students are encouraged to self-assess, leads to improvement in learning, resulting in better exam performance.

Students' and academics' opinions

The mixed results in students' surveys are consistent with the results of Ballantyne et al. (2002) and Vickerman (2009) who hypothesised that there are always students who feel that feedback is the sole responsibility of their teachers. However, we believe the mixed reaction in our study may have been caused by a lack of explanation of the benefit of TERISSA and engagement of students in the TERISSA process itself. Students were only exposed to TERISSA three times during the course of twelve weeks. In contrast, when TERISSA was used consistently during tutorials, 61% of the 449 students surveyed in Belski's (2007) research actually found TERISSA helped them to develop better judgement of the task given. 65% also felt that TERISSA provided them with immediate feedback on their knowledge of the course (Belski, 2007). When TERISSA was trialled in semester 2, 2007 to include non-engineering courses, the responses were still positive, with 64% strongly agreeing/agreeing that TERISSA helped them to identify learning areas that require their immediate consideration (Belski, 2009). In comparison, 42% of these students strongly agreed or agreed that they will continue to use TERISSA (Belski, 2009). These findings confirmed Boud's (1995) argument that self-assessment can be rejected by students if not implemented properly.

The tutors' and coordinator's survey responses were also mixed. Again, we believed the lack of explanation to some of the tutors may have led to the non-consistent survey responses. This may have caused a lack of understanding on how to utilise the TERISSA responses from the students. In section 3.4, the tutors indicated that they did not adjust the course content; it can be argued that to be able to use TERISSA feedback to improve course delivery, a degree of autonomy is required by the tutors. We believe the feedback that the academics received through the TERISSA process has enabled at least the course coordinator to deliver materials that were appropriate for the students. This in itself may have had a positive impact on the GTS scores.

Like Ballantyne et al. (2002) who reported that the staff found the process of implementing peer learning time consuming, we also found tutors felt that TERISSA took too much of their time. Since TERISSA feedbacks in our study were used in summative assessment, the tutors involved need to spend some time preparing and discussing the results with the students. On the other hand, Belski's approach in his studies (2007; 2009) was to use TERISSA more in formative assessments.

In addition, some tutors were not convinced of the benefit of using TERISSA in improving the students' grades, although it is clear from our findings, meta-cognitive reflection impacts students' learning and in this

case, student exam marks were significantly higher in the TERISSA group. There are further discrepancies between the opinions of the students who found the process useful and the tutors who used TERISSA. The tutors were convinced that for TERISSA to work, the sharing of feedback on their part is crucial. In our study, TERISSA does indeed provide immediate feedback even without the tutors' intervention. This is consistent with findings by Belski (2007 & 2009).

As seen in some of the previous comments by students, TERISSA does help students during test conditions by at least making them re-read or even focus their attention on certain parts of the problem. In our trial, in Test 1, it was found that students were using the time allocated for TERISSA to answer the test. In Test 2 and 3, students were only allowed to use TERISSA during the reading time and at the end of the test after the paper had been collected. Perhaps, the use of TERISSA in a test assessment should further be defined such that students understand that the time for TERISSA is separate from the test time.

Most encouragingly, consistent with Boud's (1995) argument that self-assessment can lead to autonomy and self-direction, we found students commenting that TERISSA made them more aware of their learning responsibility. We posit that when the students were using TERISSA, because they are forced to think about what they are doing, they internalise the feedback more.

Again, similar to Ballantyne et al's (2002) and Vickerman's (2009) studies and Boud's (1995) argument, there are always students who expect that learning and assessment is the sole responsibility of the academics. Some of the students' comments highlight the need for students to understand more about how TERISSA can help them in their studies as an individual formative feedback tool rather than a mere summative tool that requires the academics' intervention.

While comments from both the students and academics surveyed indicated that the implementation still requires much improvement, we believe ensuring that tutors are briefed clearly on the value of TERISSA will lead to better utilisation.

Conclusion

In this paper we evaluated the implementation of TERISSA in Business Statistics, a large class, with over 500 students enrolled. Of these 500 students, 231 were exposed to the TERISSA procedure in three class tests. We investigated the impact of TERISSA by analysing the GTS score. We found that the average GTS scores were higher for tutorial groups that used TERISSA. We found classes of comparable respondent numbers have significant increases due to TERISSA. Interestingly, classes with higher numbers of respondents have negligible decreases when TERISSA was utilised. These results may have been influenced by the disruptions in the sharing of TERISSA feedback and the quality of the teaching. In addition, it is known from past RMIT survey results that larger class sizes usually result in lower GTS scores.

We also investigated the impact of meta-cognitive reflection on learning outcome and found that the use of TERISSA, led to a statistically significant improvement in learning outcomes. Results suggest that by engaging students in meta-cognitive reflection, students who used TERISSA were given the opportunity to address their knowledge gaps. This in turn gave them an advantage over those who did not use TERISSA.

In order to improve implementation, students and the academics involved were surveyed. The reaction to the implementation itself is mixed from both students and staffs involved. This highlighted that the implementation process needs to be improved, particularly in the area of explaining TERISSA to the tutors and students.

This study has highlighted that there is a discrepancy between the academics' and students' opinion on what self-assessment can do in ensuring the quality of teaching and learning in large classes. The academics were seeing TERISSA as a summative assessment and feedback tool, while it is clear from the students' opinions the students are actually gaining value from the self-assessment that they were exposed to. Despite the lack of timeliness and difficulty in discussing the TERISSA feedbacks, students were finding it useful for themselves.

Therefore in implementation of TERISSA in large classes, it is recommended that:

- Before the semester the tutors must be briefed clearly and properly to ensure that there is a congruency between the tutors' and students' expectations of how self-assessment can help in learning. By having clear understanding of the scheme tutors are more likely to convey the benefits of using TERISSA to their students. This will lead to better co-operation.
- In summative assessment, tutors and lecturer should choose students' responses and share them with the students. By doing so, feedback quality will be further enhanced.
- There is more value in using TERISSA as a formative assessment and feedback tool rather than a mere summative assessment and feedback tool.

While TERISSA had the desired effect on the overall GTS score, the individual impact of other changes in the course have not been taken into account. There is an impact of TERISSA on learning outcome; however, this study has not taken into account of the students' performance in other course subjects. It is not clear whether the students continued to utilise meta-cognitive reflection in other subjects and were able to be better learners, despite some students indicating that they will continue to use TERISSA. Since participation in the surveys was optional, not all the students who used TERISSA were involved in the final survey.

Despite these limitations, we believe our study has contributed to the knowledge of how learning quality can be maintained in large classes through the use of meta-cognitive self-assessment.

References

- Anderson, R.J., Anderson, R., VanDeGrift, T., Wolfman, S.A., & Yasuhara, K. (2003). Promoting interaction in large classes with computer-mediated feedback. In *Computer Support for Collaborative Learning 2003 Conference*. Amsterdam, Netherlands: Kluwer.
- Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an Action Research process. *Assessment & Education in Higher Education*, 27 (5), 427-441.
- Belski, I. (2007). Using Task Evaluation and Reflection Instrument for Student Self-Assessment (TERISSA) to improve educational assessment and feedback. *Australasian Association for Engineering Education Conference*. Melbourne.
- Belski, I. (2009). Improving student satisfaction with feedback by engaging them in self-assessment and reflection. *RMIT ATN Assessment Conference 2009*. Melbourne.
- Beveridge, I. (1997). Teaching your students to think reflectively: the case for reflective journals. *Teaching in Higher Education*, 2 (1), 33-43.
- Blayney, P. & Freeman, M. (2008). Individualised interactive formative assessments to promote independent learning. *Journal of Accounting Education*, 26 (2008), 155-165.
- Bonwell, C.C., & Eison, J.A. (1991). *Active learning: Creating excitement in the classroom*. ERIC Digest. Washington, DC: School of Education, George Washington University.
- Boud, D. (1995). *Enhancing learning through self-assessment*. London: Kogan Page.
- Brockbank, A., McGill, I., & Beech, N. (Ed.) (2002). *Reflective learning in practice*. England: Gower Publishing Limited.
- Cooper, J.L., & Robinson, P. (2000). The argument for making large classes seem small. *New Directions for Teaching and Learning*, 81 (Spring), 5-16.
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Lexington, MA: D C Health and Company.
- Goldman, R.H., Cohen, A.P., & Sheahan, F. (2008). Using seminar blogs to enhance student participation and learning in public health school classes. *American Journal of Public Health*, 98 (9), 1658-1663.
- Lincoln, D.J. (2008). Teaching with clickers in the large-size principles of marketing class. *Marketing Education Review*, 18 (1), 39-45.

- McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education*, 10 (2), 209-220.
- O'Shea, P., & Bigdan, V. (2008). The Biggest Loser Competition. *IEEE Transactions on Education*, 51 (1), 123-130.
- Smail, C.R. (2005). The Implementation and Evaluation of OASIS: A web-based learning and assessment tool for large classes. *IEEE Transactions on Education*, 48 (4), 658-663.
- Stuart, P. (2004). Evaluation of an on-line reflective journal in engineering education. *Computer Applications in Engineering Education*, 12 (4), 209-214.
- Vickerman, P. (2009). Student perspectives on formative peer assessment: An attempt to deepen learning? *Assessment & Evaluation in Higher Education*, 34 (2), 221-230.
- Woodward, H. (1998). Reflective journals and portfolios: Learning through assessment. *Assessment & Evaluation in Higher Education*, 23 (4), 415-423.

Appendix A: TERISSA pro-forma template

Before you attempt the question:

Initial evaluation (Step 1)

Please evaluate the complexity of the problem in a scale of 1-5;
1 – very easy, 2 – easy, 3 – so-so, 4 – difficult, 5 – very difficult.

Your evaluation: 1 2 3 4 5

Please write the reasons why you have not evaluated the question as one level **less** difficult.

After you have completed the question:

Final evaluation (Step 2)

Please evaluate the complexity of the problem in a scale of 1-5;
1 – very easy, 2 – easy, 3 – so-so, 4 – difficult, 5 – very difficult.

Your evaluation: 1 2 3 4 5

Please write the reasons why you have not evaluated the question as one level **more** difficult.

Please write the reasons for any discrepancy between the original (Step 1) and the final (Step 2) evaluation.

Please write the actions you need to undertake to become more confident with a similar task next time.

Staff Survey

1. TERISSA has helped me gain valuable feedback on students' progress in my course.

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

2. TERISSA helped me discover areas of student knowledge, which required my urgent attention.

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

3. I have used the TERISSA feedback to improve course delivery during the semester.

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

4. Using TERISSA in my course has taken significant amount of my time.

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

5. TERISSA has changed the way I conducted the course.

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

6. It was difficult to convince students to use TERISSA.

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

7. I am going to make changes to other courses I coordinate to incorporate TERISSA.

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

8. Students enrolled in my course who used TERISSA have been able to enhance their learning.

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

9. TERISSA has made NO impact on my teaching practice.

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

10. Improvements in CES of the course from the previous year are entirely caused by application of TERISSA.

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

11. I would recommend TERISSA to other teaching staff

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

12. I would recommend TERISSA to all my students

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

13. Students who used TERISSA have been able to get better marks

Strongly Agree (5) 4 3 2 Strongly Disagree (1)

14. I plan to use TERISSA in the future in the following course components (you may tick more than one box and even add your own items if required):

Tutorials

Class assignments

Home assignments

Laboratory reports

Add more

Add more

Add more

Add more

15. Please comment on your experience in using TERISSA:

16. What are your recommendations to other staff and students on TERISSA usage?

17. What is the best way to promote TERISSA among staff and students?

18. Please comment on how has TERISSA impacted the CES results for your course?