Siew Fun Tang
Loshinikarasi Logonnathan *Editors*

Assessment for Learning Within and Beyond the Classroom

Taylor's 8th Teaching and Learning Conference 2015 Proceedings

Springer

Assessment for Learning Within and Beyond the Classroom

Siew Fun Tang · Loshinikarasi Logonnathan
Editors

Assessment for Learning Within and Beyond the Classroom

Taylor's 8th Teaching and Learning Conference 2015 Proceedings

Springer

Editors
Siew Fun Tang
Taylor's University
Subang Jaya, Selangor Darul Ehsan
Malaysia

Loshinikarasi Logonnathan
Taylor's University
Subang Jaya, Selangor Darul Ehsan
Malaysia

ISBN 978-981-10-0906-8 ISBN 978-981-10-0908-2 (eBook)
DOI 10.1007/978-981-10-0908-2

Library of Congress Control Number: 2016939376

© Springer Science+Business Media Singapore 2016
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Science+Business Media Singapore Pte Ltd.

Preface

Assessment is widely recognized as an important process that can either optimize or inhibit learning, depending on how it is applied. Within this space, there has been much discussion on the concept of **Assessment for Learning**. While assessment has traditionally been thought of as occurring at the end of a learning cycle, many educational institutions have re-examined the way they think about assessment to better emphasize its role in supporting the next stages of learning. In principle, Assessment for Learning promotes the idea that students will improve most if they understand their learning target, where they are in relation to this target and how they may close their gap in their knowledge. Educators, in the process, must then use assessment information to not only provide students valuable feedback for their learning, but also adjust their teaching strategies accordingly and consider more carefully how they use learning activities in their classrooms.

This year, the Taylor's 8th Teaching and Learning 2015 (TTLC2015) has chosen to focus on "Assessment for Learning: Within and Beyond the Classroom" in recognition of the power of assessment for learning as a way of raising student achievements. By placing this topic at the heart of the conference, TTLC2015 aims to explore in breadth, depth, and quality the best models and practices, strategies, lessons learnt, and success stories from implementation of assessment within the classroom to beyond its borders and up to the virtual space. Additionally, the conference aims to provide fertile ground to stimulate and compare responsive assessment approaches and practices on relatively new areas of assessment such as graduate capability assessment in view of the need of educational institutions to evidence graduate employability. The manuscripts focus on the subthemes related to accessing extracurricular learning practices, innovative assessment approaches, e-learning assessment, assessment related to graduate capabilities and professional competencies, and other elements related to assessment for learning.

Contents

Part V Other Issues in Assessment for Learning

Part I
Assessing Extra-Curricular Activities for Graduate Employment Outcomes

Students' Perceptions About Assessment Modes for Service-Learning Modules

Sarah Abedi Abdullah, Pauline Hwa Ling Teo and Kelly Pei Leng Tee

Abstract Most service-learning modules or community service initiatives (CSI) assess students' engagement in designated projects and activities which are carried out during a semester and are graded according to a set of approved rubrics. In this study, the diploma and degree CSI students who were engaged in two projects were given the space to share their perceptions and preferences for how assessments can be conducted in relation to their CSI projects. The objective of this paper was to discover if similar trends of students' perceptions and preferences about CSI assessments exist between the two student cohorts. Therefore, an instrument was designed to capture students' perceptions and feedback in relation to the assessment modes of the two CSI modules. A pre-course questionnaire and a post-course questionnaire were given out to the students to capture their responses. The data obtained were analysed quantitatively using SPSS 20 with the method of descriptive analysis. The key findings of this paper indicate that the diploma and degree students favour the current assessment modes which include reflection, presentation and fieldwork. Hence, these findings are in line with Millican's (Student community engagement—A model for the twenty-first century? The Community University Partnership Programme, The University of Brighton, 2007) views on what a typical service-learning course should encompass.

Keywords Service-learning · Students' perceptions · Assessment modes

S.A. Abdullah (✉) · P.H.L. Teo · K.P.L. Tee
Taylor's University, Subang Jaya, Malaysia
e-mail: sarahabedi.abdullah@taylors.edu.my

P.H.L. Teo
e-mail: paulinehwaling.teo@taylors.edu.my

K.P.L. Tee
e-mail: peileng.tee@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_1

1 Introduction

1.1 Background of Community Service

Community service has been widely accepted as a commendable and effective way to help those in more deprived sectors of our society. Many opportunities and avenues to serve the community are waiting to be taken up by interested parties. However, this is not just about serving. To get the most out of the experience, participants have to be able to take something away from the engagement and be able to use their newly discovered life lessons or deeper sense of self-awareness to be more engaged and committed to improve the quality of their future servitude for the betterment of communities-in-need around them. Thus, to gain the most out of community service, there should be an element of learning, balanced with the serving aspect (Edgerton 1995). In progressive steps, students will be equipped and enriched with the learning, serving and understanding to make an impact on their world.

Generally, institutions of higher learning acknowledge that community engagements do benefit students, faculty and the served community, in terms of collaborative character building, exposure to the real world, focused awareness, knowledge application and research opportunities. As a result, the last decade has seen an upsurge of interest among many educational establishments globally, in these so-called third stream activities (Millican 2007). In Millican's terminology, the "third stream" is a priority ranked alongside the teaching and research functions of a university. It bridges the gaps between higher education and broader society, business, industry and community. According to Driscoll et al. (1996), parties who benefit from service-learning include students, faculty staff, the community and the institution. Since the benefits are so widespread, it is not easy to assess the impact service-learning has on the institution and the degree and quality of students' engagement in the various projects.

Service-learning courses vary widely in duration, quality and in the balance of "service" and "learning" stressed in the course. A typical service-learning course, as explained by Millican (2007), should have these three common factors:

- A "service" component where students spend time serving in the community to meet actual needs.
- A "learning" component where students seek out or are taught information that they will integrate into their service. Learning is often both interpersonal and academic.
- A "reflection" component that ties the service and learning together. Reflection is sometimes symbolized by the hyphen in the term "service-learning" to indicate that it has a central role in learning by serving. Reflection is simply a scheduled consideration of one's own experiences and thoughts. This can take many forms, including journals, blogs and discussions.

1.2 The Role of CSI in the Malaysian Context

In the wake of growing interest in community service initiatives (CSI), its implementation as a part of the school curriculum has not been widely welcomed. The Malaysian school system and parents have traditionally regarded academic excellence as a success measure of student performance. This will consequently restrict students to a learning diet of syllabi content and examinations for their development.

Nevertheless, there are fragmented efforts initiated by private schools and individuals for students to engage with and reflect on real-world projects away from their classrooms during the longer school holidays. The Project Malaysia Kuala Selangor (Bui and Naga 2014) was a programme which called for the participation of high school and university students from around the world to engage with orphans in Kuala Selangor. Exact details of how the programme would be conducted, who would sponsor travel costs and what participants can hope to get out of the course were not provided on its web page. Above all, assessment instruments to gauge students' involvement and contributions were not specified. In the context of tertiary education, Ismail (2007) detailed the important role of service learning as an enhancer of students' diversity and academic performance. However, the tools and methods for assessing students' engagement in the CSI module were not mentioned.

1.3 Assessment Tools for CSI

The development of assessment approaches or evaluation measures for CSI are still in its infancy and difficult to develop as stated in Driscoll et al. (1996). Colleges and universities in the USA have not reached a mutual understanding about its intent and how best to incorporate these community-based projects into their curriculum. Other issues which challenge CSI assessment involve the degree of commitment from the various parties and evaluating the impact the engagement has on students, faculty, the parent institution and the community served.

While it is not uncommon for some CSI modules to require short written tests and oral presentations as assessment procedures, others may stipulate collaborative projects and fieldwork with a target community and keeping reflective journals to enhance learning with understanding as well.

1.4 Research Questions

This study seeks to answer these research questions:

1. In relation to CSI assessments, are there similarities in students' perceptions and preferences which can be identified from the responses obtained from two student cohorts?
2. What are the areas where students' perceptions showed the most marked differences?

2 Literature Review

2.1 The Need for Assessment

It is common knowledge that assessment has always been an area of scrutiny with respect to education, specifically so for CSI. Efforts are constantly taken to improve and create suitable assessment to gauge students' learning. This particular study seeks to look at students' perception on assessment for CSI. However, before looking at assessment and students' perception, it is imperative to understand why there is a need to assess students in the first place. Assessment is defined as

> the process of gathering and discussing information from multiple and diverse sources in order to develop a deep understanding of what students know, understand, and can do with their knowledge as a result of their educational experiences; the process culminates when assessment results are used to improve subsequent learning (Huba and Freed 2000, p. 10).

In other words, the purpose of implementing assessment in a curricular is, in actual fact, to make learning better for students. It is through assessment that educators are able to understand what students have learnt in their progression of a certain course or module, and to find ways to fill in the missing gaps. This not only helps students to be aware of what they know and do not know but to also give educators the opportunity to improve students' learning approach.

In terms of CSI per se, it is extremely essential to have well-developed assessment tools available to evaluate students' performance and cognitive achievements in their service-learning module. According to Steinke and Fitch (2007), assessment for service-learning efforts must ultimately indicate attainment of locally specified students' learning outcomes through the CSI. In addition, by having organized assessments, it is possible to establish the great influence that service-learning can create on student learning (Steinke and Fitch 2007). Assessments in CSI can indirectly demonstrate the effectiveness of service-learning in helping students to achieve better learning by providing service to the community.

2.2 Views on Assessment

In terms of grading types of assessment, evaluation of students' learning can be done through formative or summative assessment. Formative assessment which is usually used to check students' understanding is a part of the instructional process while summative assessment is generally carried out periodically to evaluate students' knowledge at that particular point in time (Garrison and Ehringhaus 2006). Formative assessment helps lecturers gauge how students are learning in order to improve understanding. In contrast, summative assessment such as final examinations enables lecturers to determine students' grades. According to Diaz et al. (2010), an effective practice is to take both types of assessment in a collaborative effort to

produce better outcomes. As both grading types of assessment are somewhat different in nature, integrating them would result in having the best of both worlds.

It is also undeniable that current learning theory stresses the significance of learning with understanding as compared to rote learning. Students must be assessed in a way that allows them to establish and demonstrate what they understand rather than an assessment that only requires superficial memorization (Bransford et al. 2000). This has many years been an issue of discussion, whereby students at every level of their education tend to recall facts without properly comprehending certain theories or specific contents of the subject.

Another related aspect of assessment concerns the choice of having individual or group work evaluation. Assessments that usually incorporate group work reflect real-world practices more closely when compared to individual assessments (Lombardi 2008). It is, however, extra challenging to evaluate in terms of distinguishing the individual contribution of students. Yet, according to Magin (1993), "averaged group assessment" may give a more consistent evaluation than individual assessment while Lin et al. (2001) reported in their study that Chinese speaking students seem to have a preference for group harmony and hence are in favour of group work assessment.

Lecturers, therefore, bear the heavy responsibility in determining the suitable type of assessment that would evaluate students' understanding of the CSI while helping them to maximize their learning and, at the same time, fulfilling their preference for assessment. The question is "*what is/are students' preferred assessment*?"

2.3 Negative Perceptions Towards Assessment

Regrettably, many students fail to understand that assessment is not meant to penalize them but rather to help them improve their learning process. Most students find assessments have an unfavourable effect on their learning process and are put off by it. It was reported that in a study carried out by Mezzaroba (2000), students actually displayed negative perceptions towards assessment which included fear, submission, anxiety and passiveness. In the case study which involved 26 pharmacy teachers and the same number of pharmacy students, a survey as well as semi-structured interviews for four of the teachers involved was administered. These negative perceptions of assessment clearly interfered with students' learning.

In another study, 75 students and two teachers were asked about their perceptions on assessment used for a course on English as a Foreign Language. Again, surveys, classroom observations and semi-structured interviews found students associated assessment with "nervousness, bureaucracy, control, punishment, fear, power and pressure" (Pellisson 2007). Similar responses on students' negativity towards assessment were reported by Sambell et al. (1997) in their study. In Sambell's study, students commented that exams did not actually help them to understand their subject better. These findings somewhat indicate the reverse to the

essence of assessment, which is to benefit and develop students' learning. When assessment is applied, minus the advantage of increasing students' learning, it is time for educators to deliberate ways to ensure that assessment can be more meaningful to students.

This goes on to show that the types of assessment set for a certain module are crucial as they affect the way students' react to the module. What can be done is to look at ways to help students realize and reap the benefits of assessment, and to do that, one method is to actually match students' expectations of how assessment should be carried out. Many times the way students expect to be assessed is rather inconsistent with *how* they are truly assessed.

2.4 Students' Preferences for Assessment

Assessment has indeed evolved over the years and traditionally, assessment typically referred to multiple-choice examination and the conventional evaluation by essay (Sambell et al. 1997). Presently, assessment can include other forms of evaluation such as portfolios, presentations, self- and peer assessments, simulations, practical and other innovative methods in its carte du jour (Race and Brown 1998; Sambell et al. 1997).

In a study on portfolio assessment, Vieira (2006) applied the social representation theory to draw responses from 178 students. Three main findings were reported from the study which saw most of the students approving the use of portfolio as a mode of assessment as they felt portfolio reduced the punishment, selection and exclusion effects which could be found with other types of assessment. A small group of students from the study did not prefer portfolio while a less significant number had reservations about the validity and the effectiveness of portfolios.

Correspondingly, Janssens et al. (2001) in their study found that student teachers felt portfolios helped them to reflect and demonstrate their professional development as future teachers. They perceived portfolios as an instrument to help them develop and improve in terms of career personal growth. In other related studies, Slater (1996) also claimed in his research, that students like portfolio assessment. Students reportedly said that they felt that they would be able to retain what they have learnt longer and better as they had internalized the knowledge through the process of preparing the portfolio. The study also stated that "students enjoyed the time they spent on creating portfolios and believed it helped them learn".

Reflections, especially unstructured ones, are also reportedly preferred by students as a form of assessment. White (2012) in her study on the students' perceptions on the role of reflection in leadership learning discovered that students appreciated reflection as a practice to improve their leadership skills, but found both the structured and the unstructured reflections to be challenging. This mirrors Boswell's (2010) findings in her study on students' perceptions of reflection on an immersion service trip which revealed that although students had positive feedback about the activities, they preferred unguided reflections.

Many studies have also been done to investigate students' perceptions of self- and peer assessments. De Grez et al. (2012) investigated 57 university students on their perception of peer assessment in oral presentation. Students responded positively to peer assessment and appreciated the high degree of learning from the feedback they received. Langan et al. (2008) also reported optimistic conclusions from structured interviews conducted with 12 students who participated in a residential field course. The students agreed that self- and peer assessments of oral presentation were very beneficial practices that stimulated reflection. In another study on self- and peer assessments in a problem-based learning environment, Segers and Dochy (2001) found a similar trend whereby students perceived self- and peer assessment procedures to encourage deep-level learning and critical thinking. Yet, at the same time they found that students were uncomfortable and unsure of their capability of assessing each other fairly. This was despite the fact that the scores given by the peer and the tutor were significantly correlated (Mires et al. 2001).

Based on these literatures, it is not rare that portfolios, reflections and self- and peer assessments are frequently used as assessment method for CSI. As Driscoll et al. (1998) state, students in service-learning settings are required to "learn and reflect on the community context in which service is provided" and to correlate the learning objectives of their course with the service activity which they have carried out. As seen in the reviews, these evaluation methods are not only preferred by students but are also deemed valuable and appropriate assessment methods for CSI. Through portfolios, reflections and self- and peer assessments, students are able to demonstrate what they have learnt and achieved through service-learning modules. This is in line with Entwistle and Tait (1990) who suggested that students who adopt "a deep approach to learning are more inclined towards assessment procedures that allow them to demonstrate their understanding". With all the above in mind, this study investigates the students' responses on their choice of assessment for CSI with the hope of understanding their assessment preferences.

3 Methodology

The present study involved the community service module at a private university in Selangor, Malaysia. This was a compulsory subject offered to Diploma in Business and Business Degree (Finance & Economics and Banking & Finance) students. The purpose of the module is to provide opportunities for students to use their skills in real-life situations and recognize the importance of contributing to the community. There were a total of 150 students (61-Diploma and 89-Degree) who participated in this study. They were in their second year of their diploma and degree studies, respectively. The students were of mixed-gender between the ages of 19 and 21 and were made up of local and international students.

The diploma students provided two-to-one coaching to refugee students (mentees) from a refugee school in Gombak on two projects: (1) shoe painting and (2) IT support. The students acted as mentors to help their mentees develop the skills

needed for work, life and academic success. The degree students, on the other hand, worked alongside a government agency to provide financial education to a group of small business individuals from different areas within the Klang Valley. They worked in pairs or groups of three to mentor their mentees in the marketing and budgeting aspects of their business.

The method of data collection used was a pre- and post-questionnaire survey to find out the students' perceptions of the existing service-learning assessment modes. This research also collected the qualitative data from students' responses in the questionnaire. The qualitative responses were used to support the quantitative data in order to better understand their choices of assessment. The quantitative data were analysed using SPSS 20 with a simple descriptive analysis which looked at frequencies and multiple-response frequencies.

4 Results

The results are compiled from the analysis of both the pre- and the post-questionnaire and are discussed below in detail. It touches on the hours allocated for the course, preferred mode of assessment, preferred types of assessment, the most beneficial session in the module, the most enjoyable type of assessment, usefulness of reflection and time allocation for group presentation.

Firstly, it is found that the students were agreeable with the hours allocated for community service which is 2–3 h per week. 95.1 % of diploma students and 93.7 % degree students strongly agree and agree with the allocated hours (see Table 1) as the students mentioned, "*We have other subjects to concentrate on as community service is not the only module we are taking*". Therefore, the time allocated for the community service module is sufficient for the students to comprehend the essence of community service and have a meaningful community engagement.

Secondly, it is interesting to note the results of the students' perceptions on the best assessment method for a community service module. In the pre-course questionnaire, both groups of students (diploma and degree) indicated that they favoured group work as their preferred assessment method over pair and individual effort. It is found that the majority (65.4 %) of the diploma students indicated that it was best

Table 1 Hours spent for community service

Statement	Percent (diploma)	Percent (degree)
Strongly agree	24.6	25.4
Agree	70.5	68.3
Disagree	3.3	3.2
Strongly disagree	0	1.6
Neutral	1.6	1.6
Total	100.0	100.0

Table 2 Students' perceptions on the community service assessment methods

Assessment methods	Percent (diploma: pre-course)	Percent (diploma: post-course)	Percent (degree: pre-course)	Percent (degree: post-course)
Individual effort	3.8	20.0	9.1	19.5
Group effort less than 5	28.8	33.8	59.1	39.1
Group effort 5–10 members	65.4	30.8	19.3	5.7
Pair effort	1.9	15.4	12.5	35.6
Total	100.0	100.0	100.0	100.0

to work in groups which consisted of 5–10 members. Similarly, the majority (59.1 %) of the degree students preferred to be assessed in smaller groups of fewer than 5 members (see Table 2).

Surprisingly, it is found in the post-course questionnaire that both the diploma and the degree students have a change of thought in their choices on the best way to be assessed for their community service performance (see Table 2). There is almost a 30 % drop in students' preference for group effort (less than 5 or 5–10 members) in both the diploma and the degree students. There is an increase of 16.2 % for diploma students and 10.4 % for degree students in the individual effort and 13.5 % for diploma students and 23.1 % for degree students in pair effort as compared to the students' initial preference for group work. Hence, this confirms that the current practice of the subject is in line with the students' preference when working with the community which is assigning students to work in pairs with their respective mentees.

Thirdly, students were asked about the best way to assess their community service experience (see Table 3). In the pre-course questionnaire, it was found that the majority of diploma students favoured peer feedback (40.4 %) and lecturer/tutor

Table 3 Students' preference on the types of assessment

Assessment type	Percent (diploma: pre-course)	Percent (diploma: post-course)	Percent (degree: pre-course)	Percent (degree: post-course)
Peer feedback	40.4	23.7	18.2	22.1
Lecturer/Tutor observation	36.5	23.7	34.1	23.3
Individual reflection	7.7	18.4	14.9	19.8
Written examination	5.8	2.6	5.7	1.2
Evaluation by mentees	9.6	19.7	26.1	18.6
Others (essay writing, video and case study)	0	11.8	0	15.2
Total	100.0	100.0	100.0	100.0

observation (36.5 %). The assessment which they preferred least was written examination (5.8 %). On the other hand, the degree students preferred lecturer/tutor observation (34.1 %) and evaluation by mentees (26.1 %). The least preferred type of assessment was written examination (5.7 %) which was similar to the diploma students.

Upon completion of the course, both the diploma and the degree students had changed their assessment choices compared to their indicated choices in the pre-course questionnaire. For the diploma students, there was an increase of 10.7 % for individual reflection, 10.1 % for evaluation by mentees and 11.8 % for others which consisted of essay writing, video and case study (see Table 3). Similarly, the degree students also indicated changes in their assessment choice as there was an increase for peer feedback (3.9 %), individual reflection (4.9 %) and others (15.2 %) as compared to lecturer/tutor evaluation, evaluation by mentees and written examination which had a decrease in percentage (see Table 3).

Next, both the diploma and the degree students indicated fieldwork (Week 3–7) as the most useful session in their community service module (see Table 4). This clearly indicates that the design of the course meets the course module objectives which are to inculcate a sense of social responsibility in the students, to promote the culture of caring and respect towards all members of society, to make them responsible citizens of this country, to understand the community around them and to understand themselves in relation to their community.

Additionally, the students were asked on the types of assessment which they enjoyed most. Both the diploma and the degree students agreed that they enjoyed the fieldwork component most (see Table 5) due to several reasons. Students mentioned that fieldwork benefited them most as they "*learnt new things and skills*

Table 4 The most beneficial session in the community service module

Weekly activity	Percent (diploma)	Percent (degree)
Week 1: Intro lecture	4.9	1.6
Week 2: Project briefing	14.8	11.1
Week 3–7: Field work	62.3	63.5
Week 8: Project closure ceremony	6.6	7.9
Week 9: Coursework finalization	8.2	4.8
Week 10: Presentation	3.3	11.1
Total	100.0	100.0

Table 5 The most enjoyable assessment types

Assessment types	Percent (diploma)	Percent (degree)
E-portfolio	18.0	19.0
Group presentation	23.0	19.0
Fieldwork	59.0	61.9
Total	100.0	100.0

Table 6 Usefulness of writing a reflection

Statement	Percent (diploma)	Percent (degree)
Strongly agree	34.4	20.6
Agree	60.7	69.8
Disagree	3.3	6.3
Strongly disagree	1.6	3.2
Total	100.0	100.0

such as seeing how a business is being run and also improving my interpersonal skills". It also allowed students to experience "…*hands-on activity which is more effective*", they "*get to experience learning outside of class which is something that is rarely done nowadays*" and "*every week we had different tasks to complete which was more challenging and a good experience*". The findings above are in line with the results from Table 4 above that show the students found fieldwork to be the most beneficial activity throughout the 10-week semester.

Subsequently, the students were asked on the usefulness of writing a reflection which is related to the e-portfolio. Students were asked to write an average of 5 entries of reflection in their e-portfolio submissions. Both the diploma (95.1 %) and degree (90.4 %) students agreed that the reflection was useful (see Table 6). The students mentioned that writing a reflection helped them keep track of their progress as mentioned by one student, "*I can keep track of what I have learned and apply to daily life*", "*I can determine my weaknesses every week and improve myself for the future*". It also made them aware of the changes they had experienced as stated by another student, "*Because as the weeks went by, the experiences were different, situations grew more challenging and our style of writing began to change - it is then that I realized that I was changing too. Growing!*" Furthermore, students were able to reflect and dwell on their inner thoughts and emotions as a student stated that "*The e-portfolio was a great way to keep up with what I have been doing for the entire module. Also due to the e-portfolio being private and only to be read by the module leaders, there were more intimate and realistic details that could be written out*".

Lastly, the students were asked if sufficient time had been allocated for the group presentation (see Table 7). 72.1 % of the diploma students and 87.3 % of the degree students agreed that the time allocated was sufficient. Each student was given 5 min to share his/her community service experience in their group presentations. Therefore, the results confirm that the current allocation of time for presentation is adequate.

Table 7 Time allocation for group presentation

Statement	Percent (diploma)	Percent (degree)
Strongly agree	13.1	15.9
Agree	59.0	71.4
Disagree	24.6	9.5
Strongly disagree	3.3	1.6
Other	0	1.6
Total	100.0	100.0

5 Conclusions

In brief, the findings of this paper confirm that both cohorts were satisfied with the modes of assessments of the CSI module which offered them opportunities to learn and serve at the same time. This is clearly shown in the students' overwhelming support for fieldwork as a mode of assessment which is in line with the views of Edgerton (1995) and Millican (2007) who advocate a balance of the learning and service aspects for a service-learning module to be meaningful. In addition, in terms of types of assessment, students' preferences seem to point to peer feedback and lecturer's observation.

With reference to the second research question, it is obvious that there are marked differences in certain areas in the pre- and post-course responses in areas like the number of members for group work, willingness to invest individual effort and acceptance of individual reflection as a type of assessment. Nevertheless, a similar trend of perception is evident in the responses given by both the diploma and the degree students in the pre- and post-course questionnaire.

Overall, the results of this study indicate that students managed to achieve the learning outcomes of the module in terms of collaborative learning through the assessments assigned. Their favourable responses in the pre- and post-course questionnaire in relation to how they are assessed reflect that the course has been generally well received by both cohorts. In essence, an understanding of students' CSI assessment preferences can ensure future courses can be conducted more effectively to cater for the needs of future student cohorts.

References

Boswell, L. (2010). *The structure trap: Students' perception of reflection on a co-curricular immersion Service_Learning Trip (unpublished Master's Thesis)*. Arcata, CA: Humboldt State University.

Bransford, J., Brown, A. L., Cocking, R. R., Donovan, M. S., & Pellegrino, J. W. (Eds.). (2000). *How people learn, brain, mind, experience and school. Expanded edition, National Research Council*. Washington: National Academy Press.

Bui, K., & Naga, R. (2014). Project Malaysia Kuala Selangor 2014. Available at: http://www.sealnetonline.org/past-projects/summer-projects-014/project-malaysia-kuala-selangor-2014/. Accessed on July 13, 2015.

De Grez, L., Valcke, M., & Roozen, I. (2012). How effective are self and peer-assessment of oral presentation skills compared with teacher's assessment? *Active Learning in Higher Education, 13*(2), 129–142.

Diaz, V., Brown, M., & Salmons, J. (2010). Assessment of collaborative learning project outcomes. *Educause Learning Initiatives*.

Driscoll, A., Gelmon, S., Holland, B., Kerrigan, S., Spring, A., Grosvold, K., et al. (1998). *Assessing the impact of service learning: A workbook of strategies and methods* (2nd ed), Centre for Academic Excellence, Portland State University.

Driscoll, A., Holland, B., Gelmon, S., & Kerrigan, S. (1996). An assessment model for service-learning: Comprehensive case studies of impact on faculty, students, community, and institution. *Michigan Journal of Community Service Learning, 3*(1), 66–71.

Edgerton, R. (1995). Crossing boundaries: Pathways to productive learning and community renewal. *AAHE Bulletin, 48*(1), 7–10.

Entwistle, N. J., & Tait, H. (1990). Approaches to learning, evaluation of teaching and preferences for contrasting environments. *Higher Education, 19*, 169–194.

Garrison, C., & Ehringhaus, M. (2006). Formative and summative assessment in the classroom, National Middle School Association (Reprinted in Schools Connection). *A Journal of the College of Education*, *18*(2). Kean University, Union, NJ.

Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. Needham Heights, MA: Allyn & Bacon.

Ismail, N. (2007). Enhancing tertiary students' diversity and academic performance through service learning. *URDC Colloquium*, *4*, 137–143 (MARA University of Technology (Johor)).

Janssens, S., Boes, W., & Wante, D. (2001). Portfolio: een instrument voor toetsing en begeleiding (Portfolio: An instrument for evaluation and coaching). In F. Dochy, L. Heylen, & H. Van de Mosselaer (Eds.), *Assessment in onderwijs (Assessment in education)* (pp. 203–224). Utrecht, The Netherlands: Lemma.

Langan, A. M., Shuker, D. M., Cullen, W. R., Penny, D., Preziosi, R. F., & Wheater, C. P. (2008). Relationships between student characteristics and self, peer and tutor evaluations of oral presentations. *Assessment and Evaluation in Higher Education, 33*(2), 179–190.

Lin, S. S.-J., Liu, E. Z.-F., & Yuan, S.-M. (2001). Web based peer assessment: Attitude and achievement. *IEEE Transactions on Education*, *44*, 2. Assessed July 27, 2015.

Lombardi, M. M. (2008). Making the grade: The role of assessment in authentic learning. *Educause Learning Initiative.*

Magin, D. (1993). Should student peer rating be used as part of summative assessment? *Research and Development in Higher Education, 16*, 537–542.

Mezzaroba, L. (2000). Concepts of Evaluation among Pharmacy and Biochemistry Faculty and Students at Universidade Estadual de Londrina, Paraná, Brazil. Concepções de Avaliação de Professores e Alunos de Farmácia e Bioquímica da Universidade Estadual de Londrina, Paraná. *Revista Brasileira de Educação Médica*, *24*(3), 53–61.

Millican, J. (2007). Student community engagement—A model for the 21st century? Workshop on Models of Student Community Engagement for the Global Citizens Conference at the University of Bournemouth, September 2007. The Community University Partnership Programme, The University of Brighton.

Mires, G. J., Friedman Ben-David, M., Preece, P. E., & Smith, B. (2001). Educational benefits of student self-marking of short-answer questions. *Medical Teacher, 23*(5), 462–466.

Pellisson, J. A. (2007). Perceptions of two teachers of foreign language (English) and of their students about evaluation: Implications for teacher's pre-service education. Unpublished Master's thesis, Universidade Estadual de Campinas, Campinas, São Paulo, Brasil. In D. A. S. Matos, S. D. Cirino, G. T. L. Brown, & W. L. Leite. (2013). Assessment in higher education: Multiple conceptions of Brazilian students, Estudos em avaliaçãoeducacional, *24*(54), 172–193.

Race, P., & Brown, S. (1998). The lecturer's toolkit (2nd ed). London: Kogan Page.

Sambell, K., McDowell, L., & Brown, S. (1997). 'But is it fair?': An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation, 23*(4), 349–371.

Segers, M., & Dochy, F. (2001). New assessment forms in problem-based learning: The value-added of the students' perspective. *Studies in Higher Education, 26*(3), 327–343.

Slater, T. F. (1996). Portfolio assessment strategies for grading first-year university physics students in the USA. *Physics Education, 31*(5), 329–333.

Steinke, P., & Fitch, P. (2007). Assessing service learning. *Research and Practice in Assessment, 2*, 24–29.

Vieira, V. M. O. (2006). Social representations and educational assessment: What the portfolio reveals. Unpublished doctoral dissertation, São Paulo, Brasil: Pontifícia Universidade Católica de São Paulo.

White, J. V. (2012). Students' perception of the role of reflection in leadership learning. *Journal of Leadership Education, 11*(2), 140–157.

SHINE Point Templates: The Taylor's Journey to Capture the Essence of Extra-Curricular Assessment

Su Ming Cheah and Mayriana Arman Michael

Abstract In September 2014, Taylor's University launched the 2nd Transcript which provides formal recognition and rewards students' extra-curricular achievement both on- and off campus. The extra-curricular activities recognised under the 2nd Transcript are weighed against approved SHINE point templates (SPoTs). The action research study explored whether the SPoTs created for competitions were *fit-for-purpose* and could distinguish competition activity submissions from other activity types. Secondly, the study aimed to evaluate whether these competition point templates had been distinctly mapped to the 6 SHINE point principles with *no overlaps*. Thirdly, the study aimed to formulate actions to modify the templates as appropriate. Theoretical sampling was used to select events related to competitions from September to December 2014. Data were collected from activity submissions, student reflections and one semi-structured interview. The findings surmised that the definition of "competitions" needed to be amended to better distinguish competitions from other activity types such as celebrations, to ensure the competition point templates were fit-for-purpose. Another key realisation underscored the importance of the SHINE point principle of "prestige" as a way to recognise and acknowledge students who were selected for events. The findings also suggested combining the two competition point templates to remove the need for students to make a choice. Lastly, the study also suggested that the SHINE point principle of "role" be broadened to enable the selection of dual roles. In conclusion, with the Ministry of Education's intention to implement the integrated CGPA (iCGPA) in local universities, it is hoped that this study will shed some light on how to objectively and creatively assess this relatively new but exciting dimension of out-of-class engagement in a way that motivates and celebrates student engagement.

Keywords Extra-curricular activities · Competitions · Assessment

S.M. Cheah (✉) · M.A. Michael
Taylor's University, Selangor Darul Ehsan, Malaysia
e-mail: suming.cheah@taylors.edu.my

M.A. Michael
e-mail: mayriana.armanmichaelanniah@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_2

1 Introduction

1.1 Graduate Employability in the Malaysian Context

According to Malaysia's Human Resources Minister Datuk Seri Richard Riot Jaem, the number of unemployed fresh graduates have steadily grown from 42,954 in 2010 to 52,282 in 2013 (The Star, June 2015). The Corporate Recruiters Survey 2015 Report (Graduate Management Admission Council (GMAC) reported in the Star, June 2015) showed that 92 % of the companies surveyed based recruitment on the graduate's proven ability to perform. It is no longer the lack of qualifications that stops fresh graduates from getting jobs but what else they have to offer aside from academic qualifications. The report also highlighted that 89 % of employers placed strong emphasis on strong oral communication skills. The Malaysia Education Blueprint 2015-2025 (Ministry of Education, April 2015) also stated that that unemployed graduates lacked 3 key characteristics such as a strong command of English, the right attitude and problem-solving abilities. In anticipation of growth over the next 5 years, employers projected that they would need new employees who are skilful in communication, leadership and strategy.

According to the 2015 U21 report, an acknowledged reference on higher education benchmarking that ranked higher education systems in 50 countries, Malaysia's high expenditure on higher education had not truly translated into improved efficiency. Malaysia currently allocates 7.7 % of its budget for higher education and training, far more than countries such as Japan (1 %), Korea (3.9 %), Indonesia (3.1 %), Thailand (3 %) and Singapore (6.4 %). The report indicates that the quality of Malaysian graduates as measured by employability remains a challenge primarily due to the skills mismatch with industry requirements.

1.2 Educational Reforms in Malaysian Higher Education

Educational reforms over the last 2 years have seen greater commitment to fundamentally transform the higher education system and how higher learning institutions (HLIs) operate in order to stay ahead of global trends. The respective launches of the Malaysia Education Blueprint 2015-2025 (Higher Education) in April 2015 and the integrated Cumulative Grade Point Average (iCGPA) system in August 2015 signify an overt shift to create an education system that produces holistic, balanced, lifelong, entrepreneurial graduates with a drive to create jobs, rather than to only seek jobs, i.e. nurturing "job creators" rather than just "job seekers".

The Malaysia Education Blueprint 2015-2025 (Higher Education) or MEB-HE outlines ten shifts that will spur continued excellence. iCGPA is one of the initiatives under Shift 1 of the Malaysia Education Blueprint 2015-2025 (Higher Education). Under iCGPA, assessment will be reported in the students' score cards

using a "Spider Web" methodology which enumerates all the achieved, expected or desired performance of the student, detailing not just their subjects and performance, but the skills that they have picked up along the way.

1.3 Explaining the 2nd Transcript

In September 2014, Taylor's University launched the 2nd Transcript and the SHINE Award Programme in order to equip its graduates with life, personal, social and emotional skills to enhance their graduate capabilities. The 2nd Transcript was designed to complement and supplement the academic transcript by providing opportunities to develop graduate capabilities through co-curricular and extra-curricular experiences. The 2nd Transcript is an attestation given to students to recognise and reward their involvement and achievements in out-of-class activities during their studies at Taylor's University.

The SHINE Award Programme covers four "learning packages", namely personal development, people and leadership, professional development and global engagement. Each learning package is made up of empowerment workshops and experiential opportunities. Extra-curricular activities in the SHINE Award Programme are managed and tracked via an online portal called the SHINE Portal. Students upload videos, pictures or testimonials onto their online portfolio as proof of participation.

The SHINE Award Programme, officially launched at the end of last year, is now being used by the Education Ministry as a model under the Malaysia Education Blueprint 2015-2025's (Higher Education) university transformation programme. The 2nd Transcript initiative is in line with the government's objective of

> producing graduates who not only excel in their fields of study (academically), but are also equipped with the necessary soft skills (such as English proficiency), knowledge (of the world at large, the sciences and arts), values (ethics, patriotism, and spirituality), leadership abilities (including the love of volunteerism), and the ability to think critically (accepting diverse views, innovation and problem solving).

The 2nd Transcript is a direct university strategy to enable employers to easily gauge the true potential of our graduates and address the skills and graduate–employer expectations mismatch and enable our graduates to find meaningful employment and in the long run, truly address the unemployment crisis amongst graduates.

1.4 Explaining the 2nd Transcript's SHINE Point System

The SHINE point system was developed in August 2014 and is based on a verification and validation system based on "points", given based on different levels of

participation, victories and positions which the students hold. The point system calculates the points accrued for activities undertaken. The extra-curricular activities recognised under the 2nd Transcript are weighed against the SHINE point templates (SPoTs). The SPoTs are underpinned by 6 SHINE point principles, of which 4 are compulsory principles whilst the other 2 are achievement-based.

1.5 Problem Statement: Measuring Competencies and the Need for an Acceptable Instrument

As the 2nd Transcript is the first of its kind in the region and wholly developed by the SHINE Award Centre, more data-driven research is highly necessary to test, refine and enhance the SHINE point system in order to produce a fair and equitable system to assess a range of activities, roles and levels of achievement.

2 Literature Review

Employability awards such as the 2nd Transcript are a relatively new concept in Malaysia and the Asian region. However, in the United Kingdom and Australia, the emphasis on cultivating graduate attributes through award schemes is nothing new. Graduate attribute projects have been at the forefront of changing higher education outcomes since 2005 as key strategies to enhance the employability of university graduates. In Australia, much of the graduate attribute research was spearheaded by the National Graduate Attributes Project (National GAP), funded by the Australian Learning and Teaching Council. Since the GAP Project, most Australian university institutional policies now state the university-endorsed graduate attributes, which are a key part of their curriculum renewal strategies. In the United Kingdom, the Quality Assurance Agency for Higher Education (QAA), an independent body entrusted with monitoring, and advising on standards and quality in UK higher education, provides the required guidance and leadership. In fact, initial findings by the AGCAS Award Task Group, which conducted a "skills and employability awards" survey (2011), highlighted 67 UK universities with their own award schemes or "skills awards", particularly those from the 1994, Russell, Santander, Million+ and Alliance Groups of Universities.

Generally, award schemes surveyed were structured around (i) hours such as Bath Award (University of Bath), Red Award (University of Reading); (ii) points—Southampton Graduate Plus (University of Southampton), Phoenix Award (University of Southern Queensland), UTS Build (University of Technology Sydney); and (iii) modular—Nottingham Advantage Award (University of Nottingham). Generally, 26.9 % of the award schemes in the United Kingdom were points based (Jeffries-Watts 2011). Generally, the point systems and the principles of

many of the award schemes, upon which the points are accorded, are not readily available or transparent. Fortunately, an analysis of the point systems for two award schemes—Southampton Graduate Plus and UTS Build—indicated that the key criteria for according points revolved around the length of involvement, role undertaken (degree of responsibility), types of activities and location of the activity. Furthermore, as illustrated by the UTS Build Elective Activities Reference Table (UTS Build 2015), it was also observed that points were generally set first by activity type (i.e. voluntary activity) and then, differentiated by duration of the activity (25 h—10 CLiCs points, 50 h—CLiCs 20 points) and/or role (participant—10 CLiCs points, facilitator—20 CLiCs points, etc.).

In the Malaysian education scene, there is a dearth of research literature regarding measuring graduate capabilities and competencies derived from employability awards. The emphasis on graduate employability also indicates that this research area is becoming increasingly important as indicated in the Malaysian Education Blueprint 2015-2025 (Education Shift 1: Producing Entrepreneurial, Holistic and Balanced Graduates) (MOE 2015). The search for assessment models appropriate for the Malaysian HEIs is becoming fundamental due to the pressures faced by universities to produce employable graduates. In fact, in a move to address this national gap, the Ministry of Education (MOE) has piloted a tool called iCGPA to measure the university students' competencies across 5 public universities in September 2015. An immediate concern highlighted by the Malaysian Qualifications Agency in relation to this new tool was the transparency, validity and reliability of the rubrics and criteria (The Star, August 2015). MQA further reiterated that constant evaluation is necessary in order to iron out the issues in order for the system to be accepted by all the public universities.

3 Research Methodology

The 2nd Transcript includes a wide spectrum of extra-curricular activities that can be classified under 17 different point templates. Each extra-curricular activity has its own attendant complexities. This paper explores the process undertaken to test, refine and enhance the SPoTs for the 2nd Transcript, with reference to "competitions".

The specific aims of the study are as follows:

1. "competition point templates" are fit-for-purpose
2. the descriptors in the 6 SHINE point principles do not overlap
3. refine the existing point templates if required

To explain, the action research study explored whether the competition point templates were fit-for-purpose and could distinguish activity submissions related to competitions from other activity types. Secondly, the study aimed to evaluate whether the competition point templates had been distinctly mapped to the 6 SHINE point principles with no overlaps. Thirdly, the study aimed to formulate actions to modify the templates as appropriate.

The competition point templates created in August 2014 were used in this study:

- Single Competitions (One-off)
- Competition Series

It must be stated that information on any of the SPoTs cannot be disclosed and is confidential.

The mode of this action research study was practical or technical in nature and had as its goal "to test a particular intervention based on a pre-specified theoretical framework" (Berg in Newton and Burgess 2008). This study was conducted based on McNiff (2002)'s action research model where improvement of practice was the primary purpose. The key elements of this model which were particularly valuable were the "investigation" and "reflection and evaluation" phases.

4 Data Collection

Theoretical sampling was used to select events related to competitions during the "September to December 2014" *time period* to understand how the competition point templates (an emerging theoretical concept) worked on real-life competitions.

Qualitative data were collected from three key sources as follows:

- Activity submissions related to competitions on the SHINE Portal
- Reflective submissions (essays) of approx. 150–300 words
- A semi-structured interview with an activity organiser: Sports Recreation Centre staff

Firstly, the 27 competition activity submissions were categorised into 9 distinct events due to duplication of the events. It was important at this stage to analyse and determine whether the 9 distinct events submitted could be defined as "competitions" and not another activity altogether, i.e. fit-for-purpose.

It was also considered that the competitions in the study were a general representation of the types of competitions undertaken by students at Taylor's University *but* may not be representative of the full range of competitions available, given the short period of the study.

The definition of competitions in Cambridge Dictionaries Online (2015) was taken as a baseline as follows:

> an organised event in which people try to win a prize by being the best.

Event descriptions and event resources of the 9 distinct events were mapped to the following scopes:

- Aims and/or objectives of the event
- The competition point template that was used

After that, the mapping matrix further correlated these scopes to the 6 SHINE point principles in the two competition point templates—Single Competitions and Competition Series:

- Any of the Taylor's Graduate Capabilities
- Location
- Length of activity
- Role
- Prestige
- Achievement

Secondly, the 27 reflection submissions for these 9 distinct events were analysed according to the 3 aims of the study. These reflective submissions (essays) were approximately 150–300 words at length.

Thirdly, a semi-structured interview with 2 Sports Recreation Centre (SRC) staff with 5 years of handling sports events for the university was conducted to round-out the observations gathered from the analysis of the events and the reflection submissions.

5 Findings of the Study

An analysis of the aims and objectives of the 9 distinct events to assess whether the competition point templates were fit-for-purpose was based on these dimensions: nature of the competitions, the degree the competitions was institutionalised (official rules, governing bodies), competition levels or stages and types of achievement, i.e. prizes.

Nature of the competitions: the analysis showed that 7 of the 9 distinct events could be categorised into any of these 3 competition orientations, i.e. platforms to demonstrate skills related to the participants' *discipline-specific knowledge*, *interest* or *sport*. The study showed that Hack'em All was appropriately categorised as a skill-based competition, with descriptions such as "non-stop coding and app development" and "simple app prototype". Next, the TGC Video Competition exemplified a competition which related to one's interests—in this case, video-making. The third orientation was the most ubiquitous—sports competitions which related to one's physical skills/prowess in relation to particular sports such as ping-pong or table tennis, netball, marathon and extreme sports (obstacle-racing). The names of the events also clearly indicated this competition type—Kejohanan Ping Pong Terbuka MASISWA 2014, MAPCU Table-tennis Competition 2014, Nestle Fitnesse Netball Carnival, Standard Chartered KL Marathon and Viper Challenge 2014.

In contrast to the events above, the study found that there were two (2) events which did not seem fit into the competition mould. It was found that these events were of a more celebratory nature, i.e. celebrations using sports as a medium. The

description for Scholarsseum seemed to indicate that it was more of a "celebration" of student-bonding and more focused on using sports as a medium to cultivate and strengthen social bonding amongst students. Descriptions taken from the event included "lots of fun", "awesome", "fun games" and "free food", and this indicated a festive tone for the event. The Pharmacy Sports Day further supported this finding with its objectives clearly stating that the event was aimed at "improving relations between all pharmacy students from years 1 to 4" and "to build teamwork amongst students". It was apparent from this description that building rapport was the key objective of this event. It was suggested that these 2 events were a better fit for another category entitled "Festivals and Celebrations".

It was important to note that although Scholarsseum and Pharmacy Sports Day may not be indicative of how the general population think, the inclusion of these 2 events seemed to suggest a misconception at play, i.e. that students may similarly albeit mistakenly connect all sports-related events with "competitions". This raised an interesting concern that should be addressed.

Degree of institutionalisation: the study also found that official sports competitions were governed by official sporting bodies/councils and guidelines. For example, marathons were long-distance running events with official distances. Additionally, in the case of "Kejohanan Ping-pong Terbuka MASISWA 2014" and "MAPCU Table-tennis Competition 2014", the degree of institutionalisation was apparent for both the events as MASISWA (Majlis Sukan IPTS Malaysia) and MAPCU (Malaysian Association of Private Colleges and University) sports competitions were officially recognised varsity games for private higher education institutions which were supervised strictly by these official sporting bodies. Thus, due to the institutionalisation inherent in these sports, achievement in these organised events may translate into concrete recognition, honour or acknowledgement of the respective competitor's skills amongst other genuine competitors and could even contribute to the competitors' ranking in the sports concerned.

Indeed, the 2 more "celebratory" events, i.e. Scholarsseum and Pharmacy Sports Day, tended to skew towards being fun-oriented rather than a serious test of skills; that is, the games were more related to luck rather than skill or knowledge, for example, the inclusion of station games, lucky draws and game-show games/mission races such as in the "Running Man" variety show on television. For example, Pharmacy Sports Day had 2 mission races entitled "To the Battlefield (Running Man: Part 1)" and "Do or Die (Running Man: Part 2)".

Through these findings, it was surmised that the definition of "competitions" on the SHINE Portal needed to be tightened to better distinguish submissions related to competitions from other activity types such as celebrations and thus, ensure the competition point templates were fit-for-purpose. This signalled an area for refinement which will be addressed under *Implications of the study*.

The analysis of the dimensions of competition levels or stages and achievement did not yield anything that was out of the ordinary, i.e. competitions either happen at single or multiple stages. Rewards and recognition also came in a variety of formats, i.e. medals, certificates and trophies.

Next, the study aimed to evaluate whether the competition point templates had been appropriately structured and mapped to the 6 principles without any overlaps amongst the SHINE point principles.

Whilst most event submissions were straightforward and correctly submitted into the appropriate competition categories in the SHINE Portal, there were two exceptions. "Kejohanan Ping-pong Terbuka MASISWA 2014" and "MAPCU Table-tennis Competition 2014" were mistakenly submitted and matched to Single Competitions instead of the Competition Series point template. Reflection submissions for these events also depicted the students' progress through the tournament levels, indicative that it should have been submitted under Competition Series. The excerpt from Student X's reflection on the MAPCU Table-tennis Competition supported this as follows:

> … we still managed to move to the quarter final. Our next opponent was Sunway University in the quarter final but we lose to them. This was because we made the wrong order of player listing and we also never considered that their men's doubles team was so strong. We thought our men's doubles players could beat them but we predicted this wrongly, so ended up losing 1-3 to Sunway University. Thus, our team was stopped at quarter final (Student X).

In addition, comparing the reflections of different activity submissions also yielded an important discovery. Some students were selected to participate in competitions whilst there were many others who joined the events voluntarily on their own accord. The competition point template was used to *distinguish* competitors who were especially selected or chosen by the faculty, school or university department to represent the university or campus community at different competition levels. This key realisation underscored the point principle of "prestige" as an important way to recognise the selection of students for the events as an important differentiator. To explain, the study found that "prestige" incorporated the distinctive element of distinguishing whether a competitor was "selected" to participate at multiple tournament levels such as university, state, national, international-level competitions, for example, "Kejohanan Ping-pong Terbuka MASISWA 2014" and "MAPCU Table-tennis Competition 2014". In direct contrast, events in which students voluntarily joined and for which there was no pre-selection were considered as "open entry" competitions, for example, Standard Chartered KL Marathon and Viper Challenge 2014. It was also found that points were pegged to commensurate ascendingly in correlation with the increasing degree of prestige at higher tournament levels.

All in all, with reference to the second aim of the study, there were no overlaps in the competition point templates with reference to the 6 point principles. In fact, certain point principles like "prestige" was an area in which the competition templates may have worked really well.

Interestingly, a comparison of the two (2) competition point templates found little difference between the two. However, the one key difference worth highlighting was the extra roles found under the "role" point principle where there were *additional roles available for selection* due to the ascending tournament stages such

as "competitor @ group stage" and "competitor @ quarter-finals" in the Competition Series template vis-à-vis only "competitor" in the Single Competition template. This signalled an area for refinement which will be addressed under *Implications of the study*.

Finally, the semi-structured interview with the Sports Recreation Centre staff also yielded an interesting perspective which highlighted an area for refinement for the competition point templates. The interview highlighted that some competitors played more than one (1) role in competitions, especially in sports competitions where team captains or team leaders were important roles with significant responsibilities related to strategy and teamwork that set them above normal competitors. However, both the Competition Series and Single Competition templates only had one level of roles and did not manage to capture any additional roles the competitors may have undertaken in their respective tournaments.

6 Implications of the Study

The study highlighted some areas of refinement for the competition point templates.

Definition of competitions: Firstly, the research team strongly felt that "competitions" could be clearly identified based on the two (2) fundamental dimensions "nature of the competition" and "degree of institutionalisation". The description of "competitions" could take these two dimensions into consideration in its "redefinition" on the SHINE Portal to provide greater clarity on what was generally accepted as competitions under the 2nd Transcript. The following revised description was proposed:

> Competitions must be based on merit or achievement. Competitions include sports and non-sports competitions and also include single and multiple-stage competitions.
> Competitions based on number of likes gained in social media, or games played at fairs/celebrations/festivals are not acceptable. If the competition is held at school level - e.g. School Sports Day, Scholarsseum, friendly matches in campus of which the nature is more of a social gathering as compared to a formal sports competition, please select the Fair/Festival/Celebration/Performance activity option.

The study also highlighted that it was better to emphasise what *could not be included* rather than the opposite, keeping in mind the breadth of events that could fall under this category.

Combining "Competition Series" and "Single Competitions" templates: the findings also suggested that these two highly similar point templates be combined to remove the need for students to make a choice between two competition categories and thus, reduce the probability of incorrect selections.

Extra role selections: the study also suggested that the point principle of "role" be broadened to enable the selection of more than 1 more role or dual roles. It was suggested that the point template be revised to enable students to *first or initially* select the "main" or "fundamental" role played in the competition and then, to select any "other additional" role or responsibility undertaken, if *applicable*. It was

further suggested that the additional roles should include approved discrete selections such as "team captain" or "coach".

7 Conclusion

The study highlighted that each SHINE point template could take a long time to create and refine as exemplified by the process to refine the competition point templates. This is essential so that the competition point templates derived can be applied to a comprehensive range of real-life competitions. Thus, it cannot be understated how challenging it is to create a SHINE point template which makes clear what is valued according to the 2nd Transcript and the SHINE Award Programme.

This study will have a major impact on the Taylor's student population who have voluntarily enrolled in the SHINE Award Programme as they are actively participating in activities and collecting points across the 4 learning packages in order to attain one of the award tiers available, i.e. SHINE Silver, Gold and Platinum Awards, respectively. Currently, the SPoTs play a defining role in appropriately capturing and assessing the range of activities submitted by the SHINE participants. Therefore, point templates that are regarded as comprehensive, fair and equitable would raise less resistance from our student stakeholders and make the evaluation tasks simple and clear-cut especially for complex, multilayered activities.

On a national scale, with the Ministry of Education's intention to implement the iCGPA in universities, it is hoped that this study will shed some light on how to objectively and creatively assess this relatively new but exciting dimension of out-of-class engagement in a way that motivates and celebrates student engagement.

References

Abas, A. (2015, August 10). 300 students to take part in iCGPA pilot project. *New Straits Times Online*. Retrieved September 6, 2015, from http://www.nst.com.my/news/2015/09/300-students-take-part-icgpa-pilot-project

Action Research Projects. (2015). LLAS Centre for Languages, Linguistics and Area Studies. Retrieved September 1, 2015, from https://www.llas.ac.uk/projects/2837

Aeria, S., & Khor, A. (2015, June 21). Angling for jobs. *The Star online*. Retrieved September 6, 2015, from http://www.thestar.com.my/News/Education/2015/06/21/Angling-for-jobs/

BUiLD Elective Activities. (2015). UTS: Build. Retrieved September 7, 2015, from https://www.uts.edu.au/sites/default/files/CLiCs%202014%20reference%20table.pdf

Cambridge Dictionaries Online. (2015). Retrieved September 1, 2015, from http://dictionary.cambridge.org/dictionary/english/competition

iCGPA ready to roll out. (2015, August 11). *The Star online*. Retrieved September 6, 2015, from http://www.thestar.com.my/News/Nation/2015/08/11/iCGPA-ready-to-roll-out-Grading-system-six-years-in-the-making-says-minister/

Jeffries-Watts, S. (2011). *Initial findings from the skills and employability awards survey.* AGCAS Award Task Group. Retrieved September 7, 2015 from file:///C:/Users/tia0072/Downloads/skills_awards_survey_report.pdf

Ministry of Education. (2010). *Eleventh Malaysia Plan: Strategy paper 10: Transforming education system.* Retrieved September 6, 2015, from http://rmk11.epu.gov.my/pdf/strategy-paper/Strategy%20Paper%2010.pdf

Ministry of Education. (2015). *Executive summary: Malaysian Education Blue Print 2015-2015.* Retrieved September 6, 2015, from http://jpt.mohe.gov.my/corporate/PPPM%20(PT)/4.%20Executive%20Summary%20PPPM%202015-2025.pdf

Newton, P., & Burgess, D. (2008). Exploring types of educational action research: Implications for research validity. *International Journal of Qualitative Methods.* Retrieved September 7, 2015 from http://creativecommons.org/licenses/by/2.0

Williams, R., Leahy, A., de Rassenfosse, G., & Jensen, P. (2015). *Executive summary and full 2015 report.* U21 Ranking of National Higher Education Systems 2015. Retrieved September 6, 2015, from http://www.universitas21.com/article/projects/details/153/executive-summary-and-full-2015-report

Part II
Innovative Assessment Approaches

Assessing Writing Skills of Postgraduate Students: Perspectives of Supervisors and Supervisees

Peck Choo Lim, Gurnam Kaur Sidhu, Yuen Fook Chan, Lai Fong Lee and Leele Susana Jamian

Abstract There has been a significant increase in the number of students who have enrolled for postgraduate degrees in Malaysia for the past couple of years. However, the number of postgraduates (PG) who have successfully completed their degrees remains low due to high attrition rate. To address this problem, it is therefore important to investigate the critical factors that facilitate successful and efficient completion of the PG degree. One key factor is the writing skills of PG students as it is an essential skill for academic success. This paper aims to assess students' readiness for PG study in terms of writing skills from the perspectives of supervisors and supervisees. The study involved 209 supervisees and 121 supervisors from two local public universities in Malaysia. Data were collected using questionnaires and semi-structured interviews. The results show that respondents perceived that supervisees were moderately ready for their PG study with regard to their writing skills. This suggests that PG students have adequate writing skills to cope with PG study. Additionally, the independent t test results also show that there was a significant difference between the perceptions of supervisees and supervisors regarding PG students' writing skills. It shows that supervisees think that they have a higher moderate readiness in their writing skills as compared to the supervisors. The results also show that students were less ready to write content with clarity and needed improvement in the quality of their argument. This has implications for enhancement of supervisory practices in terms of advancing approaches to further

P.C. Lim (✉) · G.K. Sidhu · Y.F. Chan · L.F. Lee · L.S. Jamian
Faculty of Language Studies, UiTM, Shah Alam, Malaysia
e-mail: pclim@salam.uitm.edu.my

G.K. Sidhu
e-mail: gurnamsidhu@salam.uitm.edu.my

Y.F. Chan
e-mail: yuenfook@salam.uitm.edu.my

L.F. Lee
e-mail: leela679@salam.uitm.edu.my

L.S. Jamian
e-mail: leelesusana@salam.uitm.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_3

develop postgraduate students' writing skills that are essential for successful completion of their study.

Keywords Writing skills · Postgraduate supervision · Supervisors · Supervisees

1 Introduction

There has been an upward trend in the number of postgraduate (PG) students registered for PG study globally (Aranda-Mena and Gameson 2012; Denicolo et al. 2010). This comes as no surprise given the types and modes of PG study available. Whilst the push for knowledge and innovation is strong motivating factor for PG study, Malaysia also capitalizes on PG study to develop and empower its human capital as a means to transform Malaysia into a knowledge economy. To realize this goal, the National Higher Education Strategic Plan (NHESP) beyond 2020 has implemented programmes that provide better access to higher education, particularly for PG study. The Ministry of Higher Education (MOHE) hopes to produce 60,000 quality graduates with doctoral degrees by 2023 with initiatives such as MyBrain 15 which finances PG studies. Such initiatives have resulted in an upward trend in PG enrolment in local universities (Sidhu et al. 2014).

Despite these rigorous initiatives in Malaysia, the number of PG completion is not proportionate to the increase in the PG enrolment. Sidhu et al. (2014) found high attrition and low completion rates of local PG students. They found that about 60 % of doctoral graduates dropped out. This poses a serious problem.

One of the reasons for low completion rate is students' inadequate reading and writing skills which are crucial for academic success (Watson and Reissner 2014). McNaught and Hoyne (2012) also confirmed this by stressing that writing skills underpin academic success in any tertiary course, irrespective of the discipline. According to Goldfinch and Hughes (2007) and Tuan (2010), the ability to write in English is considered fundamental for students to be successful in college and their future careers. This is particularly critical for ESL learners who are enrolled in tertiary studies. The importance of English language skills is underscored by Hyland (2013) who postulates that English language skills have become less a language than a basic academic requirement for many users around the world. This is further reinforced by Knoch et al. (2015) who stated that university students are expected to achieve high levels of language skills.

Although English language skills are important, Goldfinch and Hughes (2007) note that few university students are adequately equipped with writing skills necessary for academic success at tertiary levels. They further add that many of these tertiary students were under the impression that their writing skills are adequate. Zakri (2006) also confirmed that students were unable to clearly present their ideas due to their limited language proficiency and writing skills. Moreover, Lau (2003) established that the inadequate writing skills often result in students' failure to cope with normal coursework, thus leading to their eventual withdrawal from tertiary

education. This underlines the importance of writing skill in tertiary education, in particular for PG students who are ESL learners.

Another problem that results in attrition rates according to Bierman and Jordaan (2007) is students' inability to define the focus of their study. They found that students often included too many objectives in a single project and in the end they were unable to relate to their objectives and this has negatively affected their ability to complete writing their report.

Hence, the purpose of the study is to examine PG students' readiness for PG study in terms of their writing skills from the perspectives of supervisees and supervisors. It is important to look into the perspectives of supervisees as McAlpine and Norton (2006) note that student voice is rarely heard in research on PG studies. In addition, the perspectives of supervisors need to be explored as they would be able to provide insights into PG students' writing skills as they are important stakeholders in this supervisory practice. Therefore, this study hopes to answer the following questions which are as follows: 'What are supervisors' perceptions of PG students' writing skills?' 'What are supervisees' perceptions of PG students' writing skills?' and 'Is there a difference between supervisors and supervisees' perceptions of PG students' writing skills?'

2 Methods

This is a descriptive study carried out in two local public universities in Malaysia. Bickman and Rog (2008) posit that descriptive research is considered a better research design compared to other designs because it provides an opportunity to answer the 'what is' or 'what was' of a topic. Additionally, this design allows for a multifaceted approach to data collection used in this study which employed survey and semi-structured interviews. Therefore, this design is useful in answering the many aspects of the 'what is' in this study which explores the assessment of supervisors and supervisees regarding PG students' writing skills.

A mixed-methods approach was employed. Data were collected through questionnaires and semi-structured interviews. Two sets of questionnaires were used one for supervisees and another for supervisors to assess their perceptions regarding PG students' writing skills. Out of these, 12 PG students (6 from each university) and 6 supervisors (3 from each university) were interviewed, respectively. Respondents were coded as follows. For example, L1AFSS refers to Lecturer 1 (L1) from University A (A) who is a female (F) from the Social Science (SS) discipline, whilst S2BMSc refers to Student 2 (S2) from University B (B) who is a male (M) from the Pure Science (Sc) discipline.

For the purpose of this study, an adapted version of the Postgraduate Supervision Questionnaire (PGSQ) developed by Sidhu et al. (2014) was used. It comprised two sections: sections A and B. The former explored the demographic variables of supervisors (e.g. years of supervision, number of supervisees) and PG students (e.g. gender, mode of study). Section B investigated the writing skills of PG students as

perceived by supervisors and supervisees, respectively. Questions regarding writing content, organization and structure, and language were included. Data from the questionnaires were first analysed using descriptive statistics such as means and standard deviation to assess the writing skills of PG students. Then, inferential statistics such as independent sample *t* test was also conducted to compare the perceptions of supervisors and supervisees regarding PG students' writing skills. Interviews with the respondents were recorded and transcribed for qualitative analysis to help triangulate data collected from the questionnaires.

Since this is a small study involving only two public universities in Malaysia, the results cannot be used to make generalizations about the writing skills of other PG students in other universities in Malaysia.

3 Results

This section presents the results of the assessment of PG students' writing skills based on the perceptions of supervisors and supervisees. Finally, a comparison between PG students and supervisors' perceptions was carried out to investigate whether there was a significant difference between the two.

Table 1 shows the demographic profile of respondents based on gender and course discipline. There were 209 PG students comprising 67 male and 142 female. Out of these, 95 were Social Science and 114 were Science students. Of the 121 supervisors, 53 were male and 68 were female. Seventy-one supervisors were from the Social Sciences, whereas 50 were from Science.

Table 2 shows the overall mean scores of PG students' readiness in terms of their writing skills as perceived by both supervisors and supervisees. The results show that PG students perceived that they were moderately ready ($M = 3.53$, SD = 0.57). Similarly, supervisors perceived PG students to be moderately ready ($M = 2.94$, SD = 0.74). This suggests that both supervisors and supervisees perceived that PG students have adequate writing skills to undertake PG studies.

However, an independent sample *t* test was carried out to confirm whether there was a significant difference between the perceptions of supervisees and supervisors. The comparison in Table 3 shows the PG students group ($M = 3.53$, SD = 0.57) had a higher mean score than the supervisor group ($M = 2.94$, SD = 0.74). The *t* test confirms that the difference between the two groups was significant at $p < 0.05$ ($t[320] = 7.909$). The Levene's test further confirmed that homogeneity of variance had been assumed ($F = 2.425$, $p > 0.05$); i.e., the different number of PG students and supervisors in the groups did not interfere with the t test result. It is thus concluded that PG students perceive that they are more ready for their PG studies in terms of their writing skills as compared to their supervisors who perceive them to be less ready to handle writing for PG studies.

Table 1 Demographic profile of respondents

Variables	Supervisors		PG students	
	Frequency	Percent (%)	Frequency	Percent (%)
Discipline				
Social Science	71	58.7	95	45.5
Science	50	41.3	114	54.5
Total	121	100.0	209	100.0
Gender				
Male	53	43.8	67	32.1
Female	68	56.2	142	67.9
Total	121	100.0	209	100.0

Table 2 Mean score of PG students' readiness in terms of writing skills

	Supervisors			PG students		
	N	Mean	SD	*N*	Mean	SD
Writing skills	121	2.94	0.74	209	3.53	0.57

Scale: *1* very limited readiness, *2* limited readiness, *3* moderate readiness, *4* high readiness, *5* very high readiness

3.1 Supervisors and Supervisees' Perceptions of PG Students' Writing Skills

This section presents supervisees and supervisors' assessment of PG students' writing skills. These writing skills were investigated based on the writing content, organization and structure, and language skills. Table 4 below shows that supervisees perceived that they were moderately ready with regard to their overall writing skills ($M = 3.53$, SD = 0.57). Among all the writing skills listed, they perceived that they were most ready to 'use appropriate word choice' ($M = 3.20$, SD = 0.911) followed by their readiness to 'deliver clear transition and flow between sentences, paragraphs or sections' ($M = 3.18$, SD = 0.910); 'use accurate grammar and sentence structure' and 'adhere to academic writing conventions' ($M = 3.17$, SD = 0.909). In contrast, supervisors perceived PG students as most ready to 'adhere to the proposal/thesis structure' ($M = 3.18$, SD = 0.860), followed by 'adhere to academic writing conventions' ($M = 3.10$, SD = 0.978) and 'provide logical order and organization of information and ideas' ($M = 3.08$, SD = 0.845). It can be inferred that supervisees perceived that they were most ready for PG study in terms of accuracy of language followed by organization and structure of their writing. On the other hand, supervisors perceived PG students to be most ready in terms of organization and structure. This shows that there is a difference between the perceptions of supervisors and supervisees regarding PG students' readiness for PG study in terms of writing skills.

Table 3 Independent sample *t* test of supervisees and supervisors' perceptions on PG students' writing skills

		Levene's test for equality of variances		*t* test for equality of means						
		F	Sig.	*t*	df	Sig. (2-tailed)	Mean difference	Std. error difference	95 % confidence interval of the difference	
									Lower	Upper
Writing skills	Equal variances assumed	2.425	0.120	7.909	320	0.000	0.58401	0.07384	0.43874	0.72929
	Equal variances not assumed			7.372	195.006	0.000	0.58401	0.07922	0.42778	0.74025

Table 4 Supervisors and supervisees' perceptions of PG students' readiness in terms of their writing skills

Items/dimensions	Supervisors (n = 121)		Supervisees (n = 209)	
	Mean	SD	Mean	SD
1. Content				
a. Deliver clear and understandable statements	2.87	0.833	2.94	0.724
b. Provide pertinent research issues/problems	2.95	0.757	2.91	0.712
c. Link own studies to appropriate theories/previous research	2.92	0.875	3.14	0.907
d. Show gaps in literature	2.83	0.870	3.05	0.822
e. Align between the nature of research and research design	2.96	0.827	3.05	0.822
f. Provide arguments and justifications in proposal/thesis	2.86	0.836	3.04	0.821
2. Organization and structure				
a. Adhere to the proposal/thesis structure	3.18	0.860	3.16	0.908
b. Deliver clear transition and flow between sentences, paragraphs or sections	2.92	0.926	3.18	0.910
c. Provide logical order and organization of information and ideas	3.08	0.845	3.03	0.831
d. Adhere to academic writing conventions (format and references)	3.10	0.978	3.17	0.909
3. Accuracy of language				
a. Use appropriate word choice	2.80	0.944	3.20	0.911
b. Use accurate grammar and sentence structure	2.76	0.991	3.17	0.909

Scale: *1* very limited readiness, *2* limited readiness, *3* Moderate readiness, *4* high readiness, *5* very high readiness

Meanwhile, supervisees perceived that they were weakest in terms of writing content with regard to their readiness to 'provide and discuss pertinent research issues/problems' (M = 2.91, SD = 0.712), followed by 'deliver clear and understandable statements' (M = 2.94, SD = 0.724) and 'provide logical order and organization of information and ideas' (M = 3.03, SD = 0.831). The supervisors, however, perceived that PG students were weakest in their language skills particularly 'use accurate grammar and sentence structure' (M = 2.76, SD = 0.991) followed by 'use appropriate word choice' (M = 2.80, SD = 0.944). This suggests that supervisees perceived their weakest readiness was in the area of writing content, whereas supervisors perceived it to be language. Despite the difference in their perceptions regarding PG students' weakest writing skills, the findings show that PG students have adequate writing skills to cope with PG study.

In terms of the perceptions of PG students' overall writing skills, interviews conducted with supervisees and supervisors revealed similar findings. For example, both supervisees (M = 3.53, SD = 0.57) and supervisors (M = 2.94, SD = 0.74) perceived PG students to be moderately ready in terms of their overall writing

skills. Most of the PG students (11 out of 12) rated themselves as quite ready for PG study. S6AMSS rated himself '*7 out of 10*' for writing skills and S5AMSS rated his writing competency to be '*good to satisfactory*'. Only S3AFSS felt that she was poor in her writing skills and added that '*I need to improve my writing skills*'. However, most of them felt that their writing skills have improved over time. Likewise, almost all supervisors from both universities perceived PG students' writing skills as modest. L3BMSc said, '*Writing skills is a problem…. Because of the education system in Malaysia, with Bahasa Malaysia as the focus*'. This was further corroborated by L2BFSc when asked what she thought about PG students' readiness in writing, '*I think it is 50/50. Most students are not really ready in terms of writing. It has been a problem*'. L3AMSS also confirmed this when she said, '*Most students cannot write. Even PG students!*'

With reference to perceptions regarding PG students' readiness in terms of the individual writing skill components, interview data similarly confirmed that most of the supervisees perceived that they had no problems with their use of appropriate words (M = 3.20, SD = 0.911). This was supported by S1BMSc who said that '*I am quite good in English. I have good vocabulary and I can write well. My grammar is good*'. Moreover, supervisees also rated their ability to connect ideas well (M = 3.18, SD = 0.910). S4BFSc said that she was able to write coherently, '*No problems. I don't have problems connecting ideas…*'. S2BFSc reiterated that she had no problem adhering to academic writing conventions because '*the format of the thesis is given on the university website. Also, there are many theses that I can get from the library to help me…*'.

Unlike supervisees who perceived that they were most ready in their language skills, almost all supervisors on the other hand perceived PG students' ability to organize and structure their writing like 'adhere to proposal/thesis structure' (M = 3.18, SD = 0.860) to be their main strength followed by 'adhere to academic writing conventions' (M = 3.10, SD = 0.978). For example, L3BFSc believe that '*…the university website provides templates on how to write the proposal and thesis. This helps PG students. No problem with that*'. Further corroboration was given by L2AFSS who said, '*I always refer students to APA. They have no problem referring to the online version. So… generally they have no problems with formatting*'.

Most of the PG students perceived 'provide pertinent research issues/problems' (M = 2.91, SD = 0.712) as their weakest writing skill followed by 'deliver clear and understandable statements' (M = 2.94, SD = 0.724). S3AFSS admitted that '*I have difficulty defining research problems. My supervisor asked me to do literature review to help me understand the issues because it will impact my method*'. Likewise, S6BMSS also felt that he lacked reading to help him describe the gap in his study. In contrast, supervisors perceived PG students to be weakest in their language skills particularly grammar (M = 2.76, SD = 0.991). L1AMSc revealed that '*PG students do not know how to write. There are always so many grammatical errors and they do not make sense. Sometime it looks like they copy and paste materials from various sources…*'. This was further corroborated by L2AFSc when asked what she thought about PG students' language skills, '*I think it is quite*

bad. They cannot write. Most times I have to ask them to get someone to proofread their work …. Language is a problem'. The second weakest skill as perceived by supervisors was PG students' weak vocabulary skills (M = 2.80, SD = 0.944). L3AMSc confirmed this when she noted '*Very poor vocab. They don't read enough. … this is a serious problem because they hardly use English*'.

All these are testament to the fact that PG students have misconceived ideas about their language ability and their ability to structure and organize their ideas. Further, it indicates that there is a gap between the perceptions of the supervisors and supervisees.

4 Conclusion and Implications

This study will without doubt provide some insights into how to enhance PG supervisory practices, in particular PG students' writing skills seen from the perspectives of both supervisees and supervisors. The findings in this study clearly show that there is a significant difference between the perceptions of supervisees and supervisors regarding PG students' writing skills. Whilst both groups perceive PG students to be moderately ready for PG study in terms of their writing skills, they are not in agreement as to which types of writing skills PG students are good at or weak in.

Supervisees perceived their language skills as their greatest strength in writing, but supervisors viewed PG students' language skills (vocabulary and grammar) as their weakest. This is not surprising because researchers such as Goldfinch and Hughes (2007) and Tuan (2010) reveal that many students are under the illusion that they possess adequate writing skills to complete their study. One reason could be because most PG students in Malaysia are ESL learners. Despite being exposed to the English language in primary, secondary as well as tertiary education, most still have problems with the language. It is understandable because Richard and Renandya (2002) acknowledge that of all the English language skills, writing is the most difficult skill to acquire.

In contrast to their positive assessment of their language skills, supervisees rated their ability to write content as their weakest. They acknowledged that they were not so adept in writing problem statements and writing clear statements. This inability to skilfully convey content effectively shows that they are unable to reflect their clarity of thought in their writing. The ability to transmit content clearly is important as attested to by Baik (2008) who found the ability to demonstrate understanding of content/concepts is one of five important features of academic writing. Another important feature is quality of argument. Both supervisees and supervisors are in agreement that the quality of argument in PG students' writing needs improvement. Adding to this discourse, Watson and Reissner (2014) note that one of the biggest let-downs that they encounter in student work 'is a lack of argument: it is not clear what the student wants to say (lack of claim) and/or why this is important (lack of justification)—all this show students' weak critical writing

skills' (p. 62). They highlight that critical writing displays careful crafting of an argument by determining the claim, justification and any supporting evidence in the light of the intended audience (p. 62). Therefore students need to ask critical questions about their work so that they can develop 'academic assertiveness'. By understanding PG students' limitations in writing, supervisors will be better able to improve their supervisory practices by referring PG students for appropriate help and support.

What all these imply is that PG centres in local universities should set up student writing centres which offer writing courses to help them cope with writing demands. This is even more critical because most local PG students are also ESL learners. It would be helpful if these English writing centres offer a variety of writing courses such as writing courses for proposal writing, literature review and final research report. This would allow PG students to take appropriate courses as they progress in their PG study. As attested by Watson and Reissner (2014), writing plays an important role in aiding students to deconstruct their ideas on paper clearly. They added that without good writing skills, students will not be able to express and share their ideas effectively. In short, PG students will not be able to complete their studies if they are inadequately equipped with writing skills necessary for academic success and would consequently face the high risk of withdrawing from tertiary education (Goldfinch and Hughes 2007; Lau 2003).

References

Aranda-Mena, G., & Gameson, R. (2012). An alignment model for the research higher degree supervision process using repertory grids—Reflections on application in practice in built environment research. *Australasian Journal of Construction Economics and Building, 12*(3), 66–81.

Baik, C. (2008). *Issues in assessing the academic writing of students from CIPD*. USA: Chartered Institute of Personnel and Development.

Bickman, L., & Rog, D. J. (Eds.). (2008). *The SAGE handbook of applied social research methods* (2nd ed.). Newbury Park, CA: Sage.

Bierman, E., & Jordaan, M. C. E. (2007). *Developing applied research skills in 4th year students using e-learning: A case study*. Paper presented at the WWW Applications Conference held from 5–7 September 2007 at the University of Johannesburg, South Africa.

Denicolo, P., Fuller, M., Berry, D., & Raven, C. (2010). *A review of graduate schools in the UK*. Staffordshire: UK Council for Graduate Education.

Goldfinch, J., & Hughes, M. (2007). Skills, learning styles and success of first-year undergraduates. *Active Learning in Higher Education, 8*(3), 259–273.

Hyland, K. (2013). Writing in the university: Education, knowledge and reputation. *Language Teaching, 46*, 53–70.

Knoch, U., Rouhshad, A., Oon, S. P., & Storch, N. (2015). What happens to ESL students' writing after three years of study at an English medium university? *Journal of Second Language Writing, 28*(2015), 39–52.

Lau, L. K. (2003). Institutional factors affecting student retention. *Education, 124*(1), 126–136.

McAlpine, L., & Norton, J. (2006). Reframing our approach to doctoral programs: An integrative framework for action and research. *Higher Education Research and Development, 25*(1), 3–17.

McNaught, K., & Hoyne, G. (2012). *Predicting a student's success in health sciences based on their academic writing skills.* Retrieved from http://www.fyhe.com.au/past_papers/papers12/Papers/3A.pdf

Richard, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge: Cambridge University Press.

Sidhu, G. K., Kaur, S., Chan, Y. F., & Lee, L. F. (2014). Establishing a holistic approach for postgraduate supervision. In *Taylor's 7th teaching and learning conference 2014 proceedings*, 2015 (pp. 529–545).

Tuan, L. T. (2010). Enhancing EFL learners' writing skill via journal writing. *English Language Teaching, 3*(3), 81–88.

Watson, G., & Reissner, S. C. (2014). *Developing Skills for Business Leadership. CIPD*. Chartered Institute of Personnel and Development: USA.

Zakri, A. H. (2006). Research universities in the 21st century: Global challenges and local implications. Global keynote scenario at the UNESCO forum on higher education, research and knowledge: *Colloquium on research and higher education policy*, Nov 29–Dec 1, 2006, Paris.

Assessing the Critical Reading Skills of Postgraduate Students: Perspectives of Supervisors and Supervisees

Gurnam Kaur Sidhu, Sarjit Kaur, Peck Choo Lim and Yuen Fook Chan

Abstract The advent of the Information Age witnessed the democratization and massification of higher education in the twenty-first century all around the globe. Despite the significant increase in postgraduate study, attrition rates have been rather high, whilst graduate completion rates continue to fall. One of the reasons cited has been students' incompetency to handle postgraduate study due to their limited academic literacy and research skills. Critical reading skills are often viewed as a fundamental pillar for postgraduate study. Therefore, the main aim of this paper was to assess the postgraduate students' critical reading skills from the perspectives of both students and their supervisors. This descriptive study involved 209 postgraduate students and 121 supervisors from two local public universities in Malaysia. Data were collected using questionnaires and semi-structured interviews. The findings revealed that there was a significant difference in the assessment of critical reading skills of postgraduate students from the perspectives of supervisors and postgraduate students. Students perceived that they had a higher moderate level of readiness compared to supervisors. Supervisors also felt that students possessed limited readiness in terms of critical reading skills such as questioning and evaluating texts. These results have implications for the need for innovative approaches to advance the critical reading skills of postgraduate students.

Keywords Critical reading skills · Postgraduate study · Supervisors · Postgraduate students

G.K. Sidhu (✉) · Y.F. Chan
Faculty of Education, UiTM, Shah Alam, Malaysia
e-mail: gurnamsidhu@salam.uitm.edu.my

Y.F. Chan
e-mail: yuenfook@salam.uitm.edu.my

S. Kaur
School of Humanities, Universiti Sains Malaysia, Penang, Malaysia
e-mail: sarjit@usm.my

P.C. Lim
Faculty of Language Studies, UiTM, Shah Alam, Malaysia
e-mail: pclim@salam.uitm.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_4

1 Introduction

The democratization of education has witnessed a global increase in tertiary education all around the globe including Malaysia. According to the Malaysian Education Blueprint 2015–2025 (Higher Education) (Kementerian Pendidikan Malaysia 2015), both public and private institutions of higher learning (IHL, hereafter) experienced a 70 % increase in enrolment rates over the last decade to reach 1.2 million students. Despite significant enrolment gains, challenges and issues of concern such as high attrition and low completion rates continue to plague universities around the globe, including IHL in Malaysia.

Researchers such as Sidhu et al. (2013) and Lee et al. (2014) reveal that among the reasons for this scenario include poor supervisory practices and students' limitations in terms of their twenty-first-century academic skills such as communication, critical thinking, and entrepreneurial skills. Reports from potential employers further cited graduates' lack of critical thinking, communication skills, and low English language proficiency (Kementerian Pendidikan Malaysia 2015). Sustainability and success in postgraduate study require students to be well equipped with twenty-first-century skills such as critical thinking and effective reading and writing skills. Baik (2008) notes that even in Australia, there is increased concern regarding students' English language skills as many students struggle to meet the demands of mainstream university courses. According to Watson and Reissner (2014), reading and writing skills are viewed as pillars of postgraduate study. They further reiterated that the ability to read and write critically is the basis of independent thinking and the creation of knowledge necessary for the continuous improvement of professional and managerial practice. Therefore, this paper aims to analyse the perspectives of both supervisors and supervisees in assessing postgraduate students' academic readiness for postgraduate study in terms of their critical reading skills.

Reading is an integral and fundamental academic skill required for postgraduate study so that students can develop a deeper understanding of the research topic. According to Sengupta (2002, p. 2), academic reading is purposeful and students need to undertake "critical reading of a range of lengthy academic texts for completing the study of specific major subject areas". Therefore, a fluent reader must be capable of "recognizing the word forms, the graphic form and phonological information, activate appropriate semantic and syntactic resources, recognize morphological affixation in more complex word forms, and access her or his mental lexicon" (Grabe 2009, p. 27).

Dreyer and Nel (2003) note that many underprepared students underestimate the challenge in meeting the academic reading demands when pursuing postgraduate study. Instead, they rely heavily on summaries, and with limited critical reading ability, they are bound to face challenges in their academic journey. Sidhu et al. (2013) further reiterated that supervisors noted that students experience problems in postgraduate study due to their limited reading, writing, and research skills. Spack (1993) points out that in academic reading, students need to engage actively with

the text and apply critical reading strategies in order to comprehend and interpret deeper meaning within the text and evaluate relevancy of texts to their own study.

Postgraduate study often demands a variety of academic reading skills. Students need to have the ability to select relevant materials and read and evaluate them critically for use in their postgraduate projects. Critical reading is often viewed as a three-stage process that requires students to understand, question, and evaluate reading materials (Kress 2010; Kaur 2013). Carrell and Carson (1997) stated that reading, understanding, and synthesizing the material from different references are necessary in academic reading. In understanding texts, readers need not only be able to identify key points but also be able to understand the relationship between points and the relevance of the texts to their study. Huckin and Flower (1990) support this claim and added that examining the author's purposes and having the ability to critique are essential in developing students' critical reading performance. Academic reading involves questioning and evaluating texts in particular statements and claims put forward by authors alongside identifying the strengths or weaknesses of the arguments. Finally, critical readers must be able to evaluate and conclude the value of a text by looking into the robustness of the arguments and evaluating the assimilation of supportive evidence provided. This aspect of critical reading was expressed succinctly by Hughes (2000, cited in Watson and Reissner 2014) who stated that critical thinking and reading involve the application of the following three skills:

- interpretive skills to identify the meaning of a statement,
- verification skills to determine the veracity of a statement, and
- reasoning skills to analyse the inferences made in an argument.

A study conducted by Flowerdew and Peacock (2001) highlighted that besides macro reading skills, students also need to be equipped with micro skills such as the capability to create logical relationships, definitions, generalizations, examples, explanations, and predictions and distinguish fact from opinion. Hence, critical reading is an activity that requires critical and active thinking. With critical reading being viewed as a pillar to successful study at postgraduate level, it is important to assess postgraduate students' readiness in this academic skill. Therefore, this paper examines supervisors' and supervisees' assessment of postgraduate students' readiness in critical reading skills for postgraduate study. This is a report on work that is still in progress based on a larger study that investigates postgraduate supervision practices in Malaysia.

2 Research Methodology

This descriptive study employed a mixed-methods research design where data were collected using surveys and semi-structured interviews. The population sample consisted of two public universities located in Peninsular Malaysia. For reasons of

anonymity, the universities will not be named. The study involved a total of 121 supervisors and 209 supervisees (referred to as "PG students" in this study).

For the data collection process, two sets of questionnaires were employed: one for the PG students and one for the PG supervisors. These questionnaires were adapted from the Postgraduate Supervision Questionnaire (PGSQ) developed by Sidhu et al. (2013). The questionnaires for this study is comprised of two sections: Sections A and B. Section A in both questionnaires explored the demographic variables of the respondents. Section B in both questionnaires assessed PG students' readiness for postgraduate study in terms of their critical reading skills. Questions examined aspects such as students' ability to understand texts, make inferences, question claims made by authors, and evaluate reading texts.

For the interview sessions, a total of 6 supervisors, three from each university, and 12 PG students (six from each university) were interviewed. The respondents were coded based according to university location, gender, and academic discipline. For example, the coding R1-LAMSc, indicated Respondent (R) number 1 a lecturer (L) from University A who is a male (M) teaching the pure science discipline (Sc) whilst the coding R3-SBFSs refers to Respondent 3 who is a Student from University B a female from the social science discipline. Data from the questionnaires were analysed using descriptive statistics such as frequency, per cent, mean, and standard deviation to gauge supervisors' and supervisees' assessment of students' critical reading skills. However, inferential statistic technique such as independent-samples *t* tests was used to identify whether there was any significant difference between the perceptions of supervisors and supervisees on postgraduate students' critical reading skills in preparing thesis. Interviews with the respondents were recorded and transcribed for qualitative analysis to help triangulate data collected from the questionnaires.

3 Results

The following section presents the findings of the study which investigated supervisors' and PG students' assessment of students' critical reading skills. Table 1 shows the demographic profile of the respondents involved in the study. Out of the total 209 postgraduate students, there were 67 male and 142 female students, whilst the 121 supervisors comprised 53 male and 68 female PG lecturers. The demographic data also revealed that 58.7 % of the supervisors and 45.5 % of the PG students were from the social sciences discipline, whilst the remaining 41.3 % of the lecturers and 54.5 % of the PG students were from the pure science disciplines.

Table 2 shows the overall mean scores of the supervisors' and PG students' perceptions of students' readiness for postgraduate study in terms of their critical reading skills. The supervisors' assessment revealed that their students were moderately ready in terms of reading ($M = 2.96$, SD = 0.751). A rather similar view was also held by the PG students ($M = 3.51$, SD = 0.696). This implies that

Table 1 Demographic profile of respondents

Variables	Supervisors		PG students	
	Frequency	Percent (%)	Frequency	Percent (%)
Discipline				
Social science	71	58.7	95	45.5
Science	50	41.3	114	54.5
Total	121	100.0	209	100.0
Gender				
Male	53	43.8	67	32.1
Female	68	56.2	142	67.9
Total	121	100.0	209	100.0

Table 2 Mean score of students' readiness in terms of their reading skills

Dimensions	Supervisors (n = 121)			PG students (n = 209)		
	N	Mean	SD	N	Mean	SD
Critical reading skills	121	2.96	0.751	209	3.51	0.696

Scale 1—very limited readiness, 2—limited readiness, 3—moderate readiness, 4—high readiness, 5—very high readiness

both groups felt that postgraduate students possessed a moderate readiness for postgraduate study in terms of their reading skills.

Table 3 shows a comparison between the two groups' assessment of postgraduate students' critical reading kills. The descriptive data reveal that both PG students (M = 3.51, SD = 0.696) and the supervisors (M = 2.96, SD = 0.751) felt that students possessed a moderate level of readiness in terms of their critical reading skills. The independent-samples t test further confirmed that there was a significant difference between the two groups' assessment of students' critical reading skills at $p < 0.01$ (t[328] = 7208). The Levene test indicated that there were no variances due to the difference in the number of respondents in the supervisors' (n = 121) and the PG students' (n = 209) groups. The independent-samples t test result indicated that the PG students perceived a higher readiness in reading skills when compared to their supervisors' perceptions of students' readiness in handling academic reading demands.

3.1 Comparison Between Supervisors' and PG Students' Assessment of Students' Critical Reading Skills

This study further explored postgraduate students' critical reading skills based on three aspects—i.e. understanding, questioning, and evaluating texts. The questionnaire administered to both supervisors and students required the respondents to

Table 3 Independent-samples *t* test of supervisors' and PG students' perceptions on student readiness in terms of reading skills

Independent-samples *t* test										
		Levene' test for equality of variances		*t* test for equality of means						
		F	Sig.	*t*	d*f*	Sig. (2-tailed)	Mean diff.	Std. error diff.	95 % confidence interval of the difference	
									Lower	Upper
Critical reading skills	Equal variances assumed	2.217	0.137	7.208	322	0.000	0.54748	0.07595	0.39805	0.69691
	Equal variances not assumed			6.770	199.442	0.000	0.54748	0.08087	0.38800	0.70696

assess students' ability to understand the purpose of reading and their academic skills in questioning and evaluating reading materials.

In terms of understanding and comprehension of reading texts, both groups expressed a moderate level of student readiness. This is illustrated in Table 4. The findings also indicate that both supervisors and students agreed that students were able to understand texts read and identify main points and paragraph topics and had the ability to evaluate the relevancy of text to their own study. Among all the critical reading skills, students perceived that they were most ready to "ensure that the source of information is up-to-date, valid and free from bias" (M = 3.66, SD = 0.920). This was followed by their perceived readiness to "identify research topics" (M = 3.29, SD = 1.09) and "locate information related to their own studies" (M = 3.12, SD = 1.13). Nevertheless, supervisors felt that students had limited readiness in terms of inferential skills (M = 2.63, SD = 0.062), whilst students felt they had a high moderate ability (M = 3.72, SD = 0.823) in their inferencing skills and were able to read between the lines to obtain deeper meaning of texts.

Interviews conducted with both postgraduate students and supervisors revealed similar findings. Both groups highlighted that students had little problems with literal comprehension of texts. A large majority of the students highlighted that they had no problems understanding academic texts and could identify main and

Table 4 Supervisors' and postgraduate students' perceptions of students' critical reading skills

	Supervisors (n = 121)		PG students (n = 209)	
Items/dimension	Mean	SD	Mean	SD
1. *Understand the purpose of reading*				
(a) Read actively, precisely, and purposefully	3.15	0.986	3.68	0.858
(b) Identify main points in texts	3.43	0.874	3.49	0.986
(c) Identify paragraph topics/component subtopics	3.85	0.862	3.31	0.080
d) Read between the lines for deeper meaning	2.63	0.062	3.72	0.823
(e) Relate relevancy of texts to own study	3.44	0.936	3.01	0.181
	3.30		3.44	
2. *Question texts (L = 2.27) *s = 2.82*)				
(a) Check claims from primary and secondary sources	2.27	0.911	2.98	0.118
(b) Question reasons and debates in texts	2.30	0.855	2.81	0.168
(c) Question claims made by authors	2.24	0.765	2.67	0.342
3. *Evaluate texts*				
(a) Evaluate strength of arguments in texts	2.35	0.919	2.48	0.718
(b) Judge texts for facts/opinions/appropriateness	2.45	762	2.67	0.832
(c) Evaluate and link to own ideas/context/study	2.58	0.836	3.08	0.094
(d) Evaluate and link to other articles/author's argument	2.48	0.852	2.92	0.719
(e) Evaluate the usefulness of texts to own study	2.93	0.896	3.05	0.821
	2.96		3.51	

Scale 1 very limited readiness, *2* limited readiness, *3* moderate readiness, *4* high readiness, *5* very high readiness

supporting ideas. Nevertheless, three out of the 12 students interviewed said that they had problems understanding the purpose in reading materials due to their limited English language proficiency. All three students were from the social science disciplines. For example, R5-SBFSs, a management student, highlighted that she is weak in English and realized she had "*problems understanding what I read. I have to read many times before I can understand, this makes my study quite slow*".

Meanwhile, student R8-SAMSs admitted that he could understand texts written in simple language but found it difficult to understand "*some journal articles as the language is too complex*" and it required him to often "*refer to the dictionary to look up meanings of difficult words. For me this makes reading* difficult".

Likewise, supervisors interviewed also agreed that the PG students had no problem understanding basic reading materials and could identify main ideas and synthesize simple information. Nevertheless, supervisor R5-LBFSs pointed out that "*even though students can read, understand and summarise texts, I feel some of my postgraduates cannot restate accurately what the author is trying to say…these students usually end up plagiarising texts because they fail to have a good understanding of the purpose of the journal article or text they are reading*".

Besides comprehension, supervisors felt that their students had limited readiness in questioning claims made by authors ($M = 2.24$, SD = 0.765) and found it difficult to check claims from primary and secondary sources ($M = 2.27$, SD = 0.911) and question reasons and debates found in reading materials ($M = 2.30$, SD = 0.855). Students, however, expressed the view that they possessed moderate readiness in all these aspects.

Critical reading also requires PG students to have the ability to judge and evaluate reading materials. Here again, supervisors felt students possessed limited readiness in terms of evaluating the strength of arguments in texts ($M = 2.35$, SD = 0.919) and judging texts in terms of differentiating it from fact and opinion ($M = 2.45$, SD = 762). In contrast, students expressed more confidence in their abilities to evaluate the relevancy of reading texts to their study ($M = 3.05$, SD = 0.821), but admitted to their limitation in evaluating the strength of arguments in text ($M = 2.48$, SD = 0.718) and their ability to read and judge texts for facts/opinions/appropriateness ($M = 2.67$, SD = 0.832). Nevertheless, these items still fell within the scale of moderate readiness.

Interview sessions revealed that supervisors were rather vocal with regard to students' critical reading skills. Almost all of them felt that postgraduate students need to enhance their critical reading skills. For example, supervisor R2-LAFSs pointed out that though her students can interpret reading texts accurately, their critical reading skills left much to be desired. She said:

> I think a majority of my postgraduate students lack analytical skills in reading, they have difficulty in identifying patterns between texts and interpreting texts such as what are facts, what are assumptions, and what is the underlying meaning in a reading text.

This lack of critical reading skills was also articulated by supervisors R5-LBFSs and R1-LAFSc. For instance, supervisor R5-LBFSs stressed that some students had difficulty understanding various viewpoints and making their own judgments and

"*so these students often end up stating what each author is saying, I feel they cannot synthesize information read in a cohesive and critical manner*".

This lack of critical reading skills was also admitted by a majority of the students interviewed. In fact, students R1-SAFSs, R4-SAMSc, R1-SBFSs, and R10-SBFSs all agreed that they could understand texts well but had difficulty with inferential skills. For instance, student R4-SAMSc pointed out that he has "*no problem reading basic science papers but I have some problems understanding arguments by different writers and sometimes all this makes understanding difficult*". Meanwhile, student R10-SBFSs felt that she understood what she read but sometimes experienced difficulty "*to understand what different writers are saying about the same topic and is gets very confusing when I have to compare and contrast what writers are saying or arguing about*".

Finally, the need for effective critical reading skills was succinctly articulated by supervisor R2-LAFSs who stressed that postgraduate study demands students to "*read large quantities of reading materials such as journal articles and books quickly, accurately and critically. So in my opinion, postgraduate students must be critical and discerning readers who are able to evaluate the strengths and limitations of the materials read. From my experience, I have seen that students with limited critical reading skills often end up plagiarizing texts read*".

4 Conclusion

The findings of this study show that there is a significant difference between supervisors' and students' assessment of PG students' critical reading skills. Both groups, however, agreed that students display a moderate level of readiness for postgraduate study in terms of their critical reading skills. A similar finding was also recorded by Moraru and LeBoutillier (2009) who stressed that tertiary students have limited critical reading skills as they often experience difficulty in analysing and interpreting materials read. It also concurs with findings from other local studies carried out among tertiary students in Malaysia (Koo et al. 2012; Kaur 2013).

The findings of this exploratory study have wide-reaching implications for the workforce readiness of these students once they complete their postgraduate studies. It is evident that more can be done by both supervisors and supervisees towards instilling effective critical reading practices among students. It is essential for postgraduate supervisors to engage their supervisees in meaningful learning experiences towards further development of their students' critical thinking repertoires in various academic literacies. Indeed, supervisors can make their supervisees aware of the importance of becoming more critically engaged when reading different genres of academic texts that are relevant to their research study.

Whilst such recommendations are made, it is also incumbent on postgraduate supervisors to realize that altogether supervisors, supervisees, and texts play important roles in creating a context that fosters critical literacy. The role of the postgraduate supervisor can include raising awareness in terms of designing

supervision activities and intellectual discussions that foster critical thinking. On the part of the postgraduate students, they need to realize that there are indeed many benefits for them to work independently to hone their critical reading skills by adopting a critical stance when reading academic texts. Among others, one clear benefit is that postgraduate students who are critical readers become more open-minded, active, and strategic when they read various types of texts. It is also reported that critical readers understand that information presented in a range of academic texts has been authored from a particular perspective for a particular purpose (Gee 2004; Kress 2010). In realizing these benefits, students can become more active participants in their postgraduate academic journey.

References

Baik, C. (2008). Issues in assessing the academic writing of students from diverse linguistic and cultural backgrounds. Retrieved from http://www.iaea.info/documents/paper_2b711b9e6.pdf

Carrell, P. L., & Carson, J. G. (1997). Extensive and intensive reading in an EAP setting. *ESP Journal, 16*(1), 47–60.

Dreyer, C., & Nel, C. (2003). Teaching reading strategies and reading comprehension within a technology enhanced learning environment. *System, 31*, 349–365.

Flowerdew, J., & Peacock, M. (2001). *Research perspectives on english for academic purposes*. Cambridge: Cambridge University Press.

Gee, J. P. (2004). *Situated language and learning: A critique of traditional schooling*. New York: Routledge.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.

Huckin, T., & Flower, L. (1990). Reading for points and purposes. *Journal of Advanced Composition, 10*, 347–361.

Kaur, S. (2013). Critical literacy practices of english majors in a tertiary institution. *GEMA Online Journal of Language Studies, 13*(2), 21–39.

Kementerian Pendidikan Malaysia. (2015). *Malaysian education blueprint 2015–2025 (higher education)*. Ministry of Education: Malaysia.

Koo, Y. L., Wong, F. F., & Ismail, Kemboja. (2012). Students' critical consciousness through critical literacy awareness. *Gema Online Journal of Language Studies, 12*(1), 143–727.

Kress, G. (2010). *Multimodality: A socio semiotic approach to contemporary communication*. London, UK: Routledge.

Lee, L. L., Sidhu, G. K., & Chan, Y. F. (2014). Exploring 21st century skills among postgraduate students in Malaysia. *Procedia—Social and Behavioral Sciences, 123*, 130–138.

Moraru, A., & LeBoutillier, J. C. (2009). Enhancing undergraduate critical reading skills in neuroscience using instructor-developed study guides. *CELT* (Vol. 2). Retrieved on 23 July 205 from: http://celt.uwindsor.ca/ojs/leddy/index.php/CELT/article/view/3215

Sidhu, G. K., Kaur, S., Chan, Y. F., & Yunus, F. (2013). Postgraduate supervision: Exploring Malaysian students' experiences. *Procedia—Social and Behavioral Sciences, 90*, 133–144.

Sengupta, S. (2002). Developing academic reading at tertiary level: A longitudinal study tracing conceptual change. *The reading matrix*, *2*(1).

Spack, R. (1993). Student meets text, text meets student: Finding a way into academic discourse. *Reading in the composition classroom: Second language perspectives*, 183–196.

Watson, G., & Reissner, S. C. (2014). *Developing skills for business leadership*. USA: CIPD. Chartered Institute of Personnel and Development.

OSCE Standard Setting by Borderline Regression Method in Taylor's Clinical School

Keng-Yin Loh, M.I. Nurjahan, K.K. Ong, G.S. Roland and A.R. Noor

Abstract Objective structured clinical examination (OSCE) is the main assessment tool for the end of semester (EOS) summative continuous assessment for the clinical training (Phase 2) at the Taylor's Clinical School. OSCE is a clinical competency assessment, whereby a candidate who passes the OSCE is deemed as clinically competent. The traditional way of deciding the pass score as 50 % in any examination is arbitrary; this may lead to problem in OSCE where the score of 50 % (pass score) may not represent the actual competency required. In recent years, standard setting methods have been applied in OSCE by many medical schools so that a defensible, fair and absolute pass score is determined. The aim of this study is to describe the OSCE cut-off score by the borderline regression method (BRM) in Taylor's Clinical School (TCS) compared to the conventional arbitrary pass mark of 50 %. This study focused on the following two research questions: (1) What is the difference in cut-off (pass mark) of OSCE if BRM standard setting is applied? (2) What is the difference in OSCE pass rate if BRM is applied? The results of EOS 5 and EOS 7 were tabulated, and the BRM standard setting was applied to these two OSCEs. The results showed that the mean score of both EOS OSCEs was significantly lower after BRM standard setting ($P = 0.001$). With BRM standard setting, it was able to identify more poor performers in OSCE who may have passed if the conventional arbitrary pass mark of 50 % was applied. We concluded that BRM standard setting is feasible and is a reasonable as well as defensible method of standard setting for OSCE. We recommend BRM for all OSCEs in Taylor's Clinical School.

Keywords Borderline · Regression · Standard setting · OSCE

K.-Y. Loh (✉) · M.I. Nurjahan · K.K. Ong · G.S. Roland · A.R. Noor
Taylor's University School of Medicine, Hospital Sungai Buloh,
Jalan Hospital, 47000 Sungai Buloh, Selangor, Malaysia
e-mail: KengYin.Loh@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_5

1 Introduction

Objective structured clinical examination (OSCE) is used in many undergraduate medical schools as the main clinical assessment tool for the final professional examination. In the OSCE, candidate is required to perform a task within a specific time frame, and the examiner will score the candidate's performance by using a standardised predetermined checklist. The checklist comprises the entire items (tasks) that the candidate is expected to perform, and for each task, a specific score is allocated. The main objective of OSCEs is to assess the clinical competency of the candidates especially in high-stakes examinations. Therefore, it is crucial to have a robust method and a justifiable cut-off score (pass mark) in order to maintain good reliability of the OSCE (Newble and Dawson 1994).

In the conventional way, the pass mark is usually set arbitrary at 50 %, i.e. those who have scored above 50 % are considered to be passed. However, there are some concerns regarding clinical assessment: whether getting an arbitrary 50 % as cut-off score really predicts the actual competency of the candidate and certifies him to have passed the examination. This arbitrary setting of the pass mark as 50 % may not provide robust and valid evidence for pass–fail decisions as each examination is unique in several ways (Ben-David 2000). In order to obtain a fairer cut-off score (pass mark), applying standard setting in OSCEs is the current practice in many high-stakes examinations (Mash 2007). Borderline regression method is one of the standard setting methods for OSCE which provides a reliable, credible and feasible standard for an OSCE (Kramer et al. 2003). In this method, the cut-off score (pass mark) is determined by a regression correlation between the actual score in the OSCE checklist and the global rating by the examiner (Fig. 1).

The aim of this paper is to describe the OSCE cut-off score by borderline regression method (BRM) in Taylor's Clinical School (TCS) compared to conventional arbitrary pass mark of 50 %. This study focused on the following two research questions: (1) What is the difference in cut-off (pass mark) of OSCE if BRM standard setting is applied? (2) What is the difference in OSCE pass rate if BRM is applied?

Study approval was granted by the dean and the clinical school programme director. The results of this study will not affect the candidates' performance in any way.

2 Methods

In Taylor's Clinical School (TCS), OSCEs are conducted as the main clinical assessment in the end of semester (EOS) examinations. A total of 12 OSCE stations were examined in both the end of semester (EOS) 5 and EOS 7. Four out of 12 stations were observed stations where the candidate is expected to complete a clinical task or procedure within 6 min. Forty-one candidates sat for the EOS 5, and

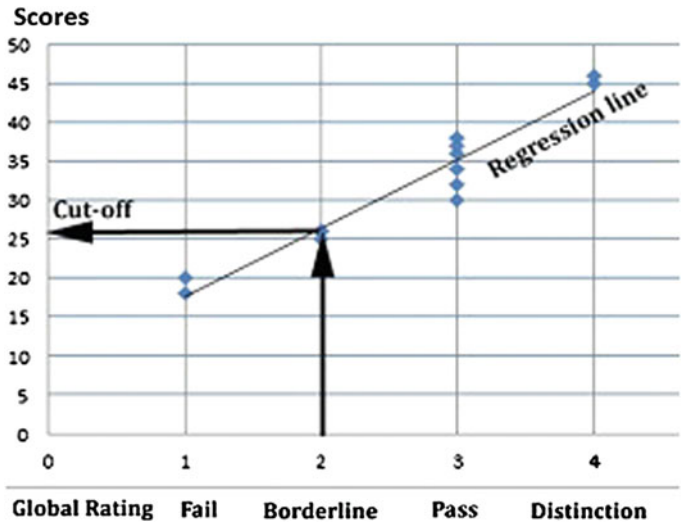


Fig. 1 Cut-off score obtained by borderline regression standard setting method

58 candidates sat for the EOS 7 OSCEs. In each OSCE station, the candidate is given specific instruction to complete a task within 6 min, and they were observed by an examiner. Marks were given according to a predetermined OSCE checklist. The examiners were also required to assess the candidate using a global rating scale which comprises of four categories to capture the overall perception of each candidate's performance. These categories were as follows: clear fail, borderline, pass and distinction. At the end of the OSCE examination, the candidate's actual scores were plotted against their global rating category. A regression line that is drawn across the scores for these 4 categories which intercept the borderline group will be the cut-off mark (pass score) of the OSCE station (Fig. 1).

SPSS version 20 is used to analyse the scores and to generate the linear regression line of the scores against the global rating category. The overall pass score after applying BRM standard setting of both examinations was compared with the arbitrary pass mark of 50 %.

3 Results

There were 41 and 58 students who sat for the EOS 5 and EOS 7 examination, respectively. Both EOS 5 and EOS 7 consisted of 12 OSCE stations respectively. The reliability of EOS 5 OSCE was 0.68 which were considered as acceptable, while the EOS 7 OSCE reliability was 0.52, which is moderate (Table 1). The mean OSCE score for EOS 5 and EOS 7 was 61.5 and 63.7, respectively, but became lower after BRM standard setting, which were 57.4 and 59.1, respectively. The

Table 1 Overall OSCE performance and reliability for EOS 5 and EOS 7

Assessment	EOS 5		EOS 7	
Standard setting	Before BRM std setting	After BRM std setting	Before BRM std setting	After BRM std setting
Total number of candidates	41	41	58	58
Mean	61.5	57.4	63.7	59.1
Median	63.0	59.0	64.1	59.5
Mode	64.0	60.0	63.4	53.6
Standard deviation	8.9	8.4	7.8	7.2
Minimum	40.0	37.0	40.9	38.2
Maximum	76.0	71.0	76.6	71.7
OSCE reliability (Cronbach alpha)	0.675		0.52	

Table 2 Difference in OSCE mean score pre- and post-standard setting by BRM

Assessment	Number of candidates	Mean OSCE score		*p* value	95 % confidence interval
		Before BRM std setting	After BRM std setting		
EOS 5	41	61.5	57.4	0.001	3.79–4.35
EOS 7	58	63.7	59.1	0.001	4.39–4.37

difference in means score after standard setting was significant ($p = 0.001$) for both EOS examinations (Table 2).

The EOS 5 and EOS 7 OSCE cut-off (pass score) marks were 56.4 and 57.7, respectively, after applying BRM standard setting. For each subject-specific station, the cut-off (pass score) varies from 45–65 marks after applying BRM standard setting. As compared to the original arbitrary pass mark of 50, the differences are statistically significant ($p = 0.001$) (Table 3).

BRM was able to identify more candidates who performed below the cut-off mark. Two additional candidates failed the EOS 5 OSCE and three additional candidates failed the EOS 7 OSCE if BRM was applied; the failure rate increased by 4.8 and 5.1 %, respectively (Table 4).

4 Discussion

This is a pilot study to evaluate the effect of BRM standard setting on OSCE performance in the end of semester examinations in clinical training at the Taylor's University School of Medicine (TUSOM). The medical course in TUSOM is new where currently there are only four cohorts of clinical students. OSCE remains the major mode of assessment for clinical competency in the clinical phase, and it is

Table 3 Cut-off score for each OSCE station in EOS 5 and EOS 7

Station	Cut-off (pass mark) without standard setting	Cut-off (pass mark) with BRM standard setting	*p* value	95 % confidence interval
EOS 5				
Obstetric and gynaecology	50	61.2	0.001	−11.258 to −3.942
Internal medicine	50	57.3		
Paediatrics	50	48.6		
Surgery	50	63.3		
Overall OSCE	*50*	*57.6*		
EOS 7				
Ophthalmology	50	57.8	0.001	−10.392 to −5.008
Internal medicine	50	57.1		
Surgery	50	62.2		
Anaesthesiology	50	53.6		
Overall OSCE	*50*	*57.7*		

Table 4 Number of candidate who have scored below cut-off (fail) mark

Assessment	Without standard setting	Fail rate (%)	With BRM standard setting	Fail rate (%)	Difference (%)
EOS 5	4	9.8	6	14.6	4.8
EOS 7	3	5.2	6	10.3	5.1

used for decisive purposes, which is to decide whether a candidate passes or fails the clinical assessment. It is therefore crucial to have a robust process which is defensible and controllable to determine the standard necessary to define a pass score in OSCE (Harden et al. 1992).

The results of this study show that BRM standard setting is feasible method for the end of semester OSCE assessment. It was able to identify additional poor performers who would have passed if the arbitrary 50 % cut-off mark was used as the pass score. When OSCE is used as a competency assessment tool in clinical training, a candidate who passes the OSCE is certified as having attained the minimal required clinical competency. It is therefore very crucial to make sure that the standard required has to be fair and absolute. Without standard setting, candidates who are borderline performers may obtain a pass score, but when BRM standard setting was applied, they were identified as having failed the OSCE.

Although the mean OSCE score and proportion of candidate who have passed or failed varied after applying standard setting, this is a normal expected finding and

should not be viewed as an issue. As shown in this study, additional candidates who have failed after applying BRM standard setting are the ones who have not achieved minimal competency; thus, it is crucial to identify them and provide them necessary remedial guidance early in the course of study.

5 Conclusion

The BRM standard setting is feasible, reasonable and defensible method of standard setting for the OSCE. We recommend BRM for all OSCEs in Taylor's Clinical School.

References

Ben-David, M. F. (2000). Standard setting in student assessment. *Medial Teacher, 22*, 120–130.

Cusimano, M. (1996). Standard setting in medical education. Academic Medicine, *71* (Suppl. 10), S112–20.

Harden, R. M., Hart, I. R., & Mulholland. H. (Eds.). (1992). Approaches to the assessment of clinical competence part I. In *Proceedings of the 5th Ottawa International Conference on Medical Education and Assessment*; Dundee, Scotland, September 1–3, 1992.

Kramer, A., et al. (2003). Comparison of a rational and an empirical standard setting procedure for an OSCE. *Medical Education, 37*, 132–139.

Mash, B. (2007). Assessing clinical skills-standard setting in the objective structured clinical exam. *South African Family Practice, 49*, 5–7.

Newble, D., & Dawson, B. (1994). Guidelines for assessing clinical competence. *Teaching and Learning in Medicine, 6*, 213–220.

Newble, D., Jolly, B., & Wakeford, R. E. (Eds.). (1994). *The certification and recertification of doctors. Issues in the assessment of clinical competence*. Cambridge: Cambridge University Press.

Part III
Assessment in a Technology-Rich Learning Environment

Assessing Factors Affecting Students' Acceptance and Usage of X-Space Based on UTAUT2 Model

Wei Wei Goh, Siew Fun Tang and Chee Leong Lim

Abstract Collaborative classroom has started to replace traditional classroom in education. Flexible learning space promotes an engaging, active and learner friendly learning environment. There are many discussions about their design, architecture, effectiveness and acceptance to the students and lecturers. It is undeniable that collaborative learning space is able to support various types of learning styles of students. The purpose of this study is to examine the factors that affect students' acceptance and usage of X-Space based on the Extended Unified Theory of Acceptance and Use of Technology (UTAUT2) model. Data were collected from 370 students from five schools of the Foundation Programme. An exploratory factor analysis (EFA) was employed to analyse the data. This study focused on Foundation students, and the result should not be used for generalization. In the future, research can be expanded to focus on degree and postgraduate students. Lastly, with the result obtained in this research, it contributes to universities and institutions that plan to introduce and implement similar collaborative classroom facilities. This research can be used as a guide for them to design and implement technologies-rich collaborative classroom.

Keywords Collaborative learning spaces · UTAUT2 model

W.W. Goh (✉) · S.F. Tang · C.L. Lim
Taylor's University, Subang Jaya, Malaysia
e-mail: weiwei.goh@taylors.edu.my

S.F. Tang
e-mail: SiewFun.Tang@taylors.edu.my

C.L. Lim
e-mail: CheeLeong.Lim@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_6

1 Introduction

Physical classroom in the university plays an important role in determining the overall student success academically (Wilkinson 2001). The usage and design of learning spaces have stimulated a great deal of discussion in higher education contexts. The correct use of infrastructure in the classrooms can create an effective learning environment in which students are able to learn comfortably. Layout of learning space has evolved from the assembly-line approach during the industrial age to the flexible and movable furniture setting (Ramsden 2011). Flexible learning space resulted in more student-centred classes as compared to traditional classroom which is more teacher-centred learning. Research showed that space quality impacts on educational experience and helps to create an optimal learning experience. Innovative classroom design offers an engaging classroom that will facilitate active and collaborative learning.

Since students learn in a diverse way, university should look into ways to provide flexible learning spaces to accommodate the different learning styles. Sound learning theory and ergonomic principles can be applied to create a conducive learning space in order to improve student learning and performance. Learning space should be mapped to learning and teaching activities to optimize the engagement between students and lecturers. Many researches have been conducted to explore the design and use of collaborative learning spaces and how these spaces impact on students' learning. University tends to focus on technological modifications, students' space preferences and how the students use the spaces for (Ramsden 2011). Some researches focus on the assessment of the learning spaces and how it changes and supports students' learning behaviour.

However, these studies place very little consideration into how students react to the collaborative learning spaces and what is the perception of the students in using the learning spaces? It is important to understand student's acceptance and usage of the space because this will determine the future design of the collaborative learning spaces in the university. Hence, it is clear that there is gap in the literature; thus, this study examines the factors that affect students' acceptance and usage of collaborative learning space. X-Space is the collaborative classroom facility used in this study. X-Space is a collaborative learning space that applies interactive technologies and flexible furnishings in the classroom. The main purpose of X-Space is to provide an ideal learning environment for students and to support active and collaborative learning.

1.1 Objective

The scope of this study is to examine the factors that affect students' acceptance and usage of X-Space based on The Extended Unified Theory Of Acceptance And Use Of Technology (UTAUT2) model. The objective of this research is to explore the

perception of students in using X-Space in term of the six main determinants, namely (1) performance expectancy, (2) social influence, (3) effort expectancy, (4) hedonic motivation, (5) price value and (6) facilitating conditions using UTAUT2 model.

2 Literature Review

2.1 Collaborative Learning Spaces

Collaborative learning space provides an ideal learning space for active and engaging learning. According to JISC (Todhunter 2015), learning spaces should be flexible, future-proofed, bold, creative, supportive and enterprising. Collaborative learning spaces should be able to accommodate current and evolving pedagogies, enable space to be rearranged and reconfigured. Furthermore, it needs to energize and inspire lecturers and students, be supportive to develop the potential of all students and make each space capable for different purposes.

Careful research and design of the classroom is important to ensure learning satisfaction among students and lecturers. The purpose of collaborative learning space is to ensure that students can fully participate in activities. This allows students to work independently on specific tasks and come together in a group to discuss project (Lippman 2015). Nowadays, learning space has greatly improved to increase more interactions between students and move away from didactic teaching.

Many studies have been conducted to design an effective learning spaces for library spaces (Wilkinson 2001) and collaborative cafe (Wang 2008). These studies applied learning theory and ergonomic principles to design the physical space and classroom layouts (Todhunter 2015) to achieve an optimal learning space. According to Wilkinson (2001), the first step to design a collaborative learning space is to define the main purpose of the space. Next is to determine what types of learning activities will be conducted in the learning space. Collaborative learning space can accommodate a wide variety of learning activities. Todhunter (2015) also agree that effective design of learning spaces must have a clear definition of learning activities in order to achieve the ideal learning outcomes. The previous studies showed that collaborative learning spaces have been very successful to allow students to connect with each other and with their lecturers (Folkins et al. 2015). Nowadays, students have different learning styles, and they do not like to sit in the same arrangement of the classroom for a long time (Kennedy 2013). Comfortable seating is important to optimize the focus of students and improve their performance in class (Jayaratne and Fernando 2009). Thus, a well-designed collaborative learning space promotes students satisfaction in studies.

2.1.1 X-Spaces

The collaborative learning space used in Taylor's University is called the X-Space. X-Space is defined as a collaborative learning environment that applies interactive technologies with the communication infrastructure. Physical attributes of X-Spaces such as lighting, tables, chairs, aesthetics, collaborative tool, connectivity and technology are taken into considerations during the design process. X-Space is purposely designed to facilitate diverse learning experience in a highly collaborative and engaging manner.

The circular table in X-Space allows students to form collaborative learning groups. Each table has access to its own high-definition flat panel displays. Students can project their presentation materials to the flat panel display using their own laptops, tablets and smartphones. High-definition multimedia interface (HDMI) and VGA cables are used to allow screen sharing between devices and the television monitor. Lecturers have complete control over the level of lighting, sound, projector and TV display. In summary, the flexibility of X-Space support active, collaborative and peer-to-peer approaches to learning.

2.2 UTAUT2 Model and Previous Studies

In this research, the Extended Unified Theory of Acceptance and Use of Technology (UTAUT2) model (Venkatesh et al. 2012) which is an extension of the UTAUT and TAM was chosen as the theoretical grounding for exploring factors affecting students' acceptance and usage of X-Space.

2.2.1 Difference Between UTAUT and UTAUT2 Model

The UTAUT is a unified model developed by Vankatesh et al. (2003) based on social cognitive theory. UTAUT model combines eight existing information technology (IT) acceptance research models. The UTAUT model uses four core determinants of usage and intention (performance expectancy, effort expectancy, social influence and facilitating conditions) alongside with four moderators (gender, age, experience and voluntariness of use) of key relationships.

According to Venkatesh et al. (2003), UTAUT model is believed to be all-encompassing compared to other technology acceptance model in evaluating and predicting technology acceptance. Credits were given to the UTAUT model for its ability to explain and predict behaviour intention and user behaviour regarding technology such as nurses using medical teleconferencing application (Biemans et al. 2005), e-government services (Alawadhi and Morris 2008), Internet banking (Cheng et al. 2008), adoption of the EMR system and use by doctors and use of Moodle (Hsu and Chang 2013).

However, UTAUT model does not take into consideration perceived usefulness and perceived ease of use as determinants of usage. User's acceptance of a new technology integrated into their environment is expected to be different based on the technology type, its use and the organization involved. The previous studies on various technologies acceptance in organizations have applied UTAUT with extensions/integration (Venkatesh et al. 2012). Therefore, an improved UTAUT2 model allows its predictors to be changed. In the study by Venkatesh et al. on consumer use of technology, they incorporated TAM model which focuses on perception, attitudes and beliefs to measure the effect of behavioural intention on technological use.

However, since the students' setting is different from the generic information system, this study deemed the model to be more accurate in treating behavioural intention as another determinant rather than as actual behaviour due to the fact that the utilization of X-Space in Taylor's University is still in its pilot stage. Further, there are new emerging studies that suggest behavioural intention predicts the actual behaviour with regard to technology use (Kriek and Stols 2010; Lau and Woods 2008). Additionally, many past studies have used behavioural intention in explaining technology acceptance among students (Wong et al. 2013; Khan 2011).

Hence, in the context of this study, the extension/integration was reformed, and the theoretical model comprises six main determinants of intention and students' preference to use X-Space: (1) performance expectancy, (2) social influence, (3) effort expectancy, (4) hedonic motivation, (5) price value and (6) facilitating conditions to find out students' perspectives of X-Space which were integrated as the determinants.

We did not consider moderating effect of gender, age, experience and voluntariness in this study. This is because the participant in this study comprises university students, whom gender, age and experience are similar. They participated in this research voluntarily.

From the theoretical framework in Fig. 1, the following hypotheses were proposed:

H_1 Performance expectancy positively affects students' intentions to use X-Space;
H_2 Effort expectancy positively affects students' intentions to use X-Space;
H_3 Social influence positively affects students' intentions to use X-Space;
H_4 Facilitating conditions positively affect students' intentions to use X-Space;
H_5 Hedonic motivation positively affects students' intentions to use X-Space;
H_6 Price value positively affects students' intentions to use X-Space;
H_7 Habit positively affects students' intentions to use X-Space;
H_8 Facilitating conditions of X-Space positively affects students' use behaviours of actually using X-Space;
H_9 Habit positively affects students' use behaviours of actually using X-Space;
H_{10} Students' behavioural intentions to use X-Space positively affect students' use behaviours of actually using X-Space.

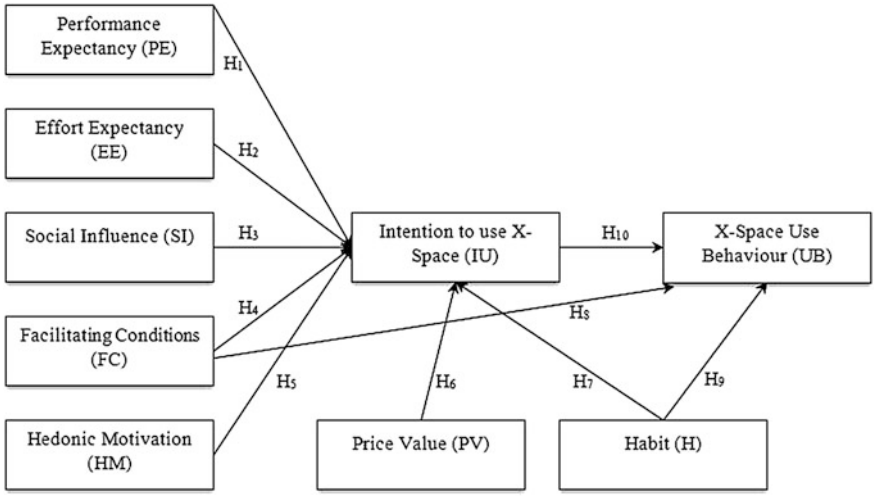


Fig. 1 Research framework

3 Research Methodology

This study is based on the UTAUT2 model by Venkatesh et al. (2012), which has six determinants. An online questionnaire was created to address the ten hypotheses. After data screening and cleaning, 327 responses for this study were used for data analysis. The questionnaire was posted online through the university e-learning management system.

The first section in the questionnaire includes questions that measure the respondents' perception in using X-Space. The second section includes the respondents' demographic and experiences in using X-Space. All the questions were measured using 5 items likert scales with responses ranging from "strongly agree" to "strongly disagree". Respondents do not take more than 10 min to answer the questionnaire. All the questions were checked for clarity and appropriateness before emailing out.

Data were analysed statistically using exploratory factor analysis, reliability testing and multiple linear regression.

4 Results

Respondents were taking Foundation programmes from five schools in the university: 47 % of the respondents were from school of hospitality, tourism and culinary arts, followed by school of business, 24 %, and school of computing and IT, 20 %. The respondents were between 18 and 20 years of age with 147 (45 %)

were males and 180 (55 %) were females. All respondents have used the X-Space before. 282 (86.2 %) and 45 (13.5 %) students have been using the X-Space for one semester and two or more semesters, respectively. 84 % of the students use the X-Space for the first time. 55 % of the students used X-Space for more than four hours in a week, and 34 % of the students used X-Space for 2–4 h in a week.

An exploratory factor analysis (EFA) was performed to reduce the large number of variables (items) to a smaller set of underlying factors that summarizes the essential information contained in the variables. The detailed explanation of the analysis and its interpretation are presented in the following section.

To determine the underlying factors, principal axis analysis with Promax rotation is employed as an indicative test to determine whether the 25 items are tapping onto the same construct. Promax rotation provides a far more interpretable solution because the difference between high and low loadings is more apparent in the pattern matrix which has fewer complex variables and a simpler structure. The loadings in the pattern matrix represent that unique relationship between the factor and the variable. The resulting factor correlation matrix indicates that all the factors appear to be lowly to moderately related. The Barlett's test of spherity was significant and that the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.930, far greater than 0.6. An inspection of the anti-image correlation matrix revealed that all the measures of sampling adequacy were well above the acceptable level of 0.5. Using items 1–24, if eight components were extracted, then 73.7 % of the variance would be explained. The overall reliability statistics (Cronbach's alpha) is 0.953. All constructs were found to be highly reliable by the used standards. UB has only one time, hence no Cronbach's alpha value (Table 1).

To test the hypotheses in this study, multiple linear regression (with "Enter" method) is employed. In predicting the intention to use X-Space (IU) among students, PE, EE, SI, FC, HM, PV and H are used as the predictors.

Table 2 shows the descriptive statistics for the variables understudy. 62.5 % of the variability in IU is accounted by PE, EE, SI, FC, HM, PV and H. In Table 3, the F test indicates that the model is a good fit for the data. Thus, the independent variables PE, EE, SI, FC, HM, PV and H can be used to explain the dependent variable IU. The results also demonstrate that PE, HM, PV and H have an influence on the intention to use (IU) with HM as the strongest predictor followed by H, PE and PV. Based on the statistical analysis carried out, hypotheses H_2 (Effort expectancy), H_3 (Social influence) and H_4 (facilitating conditions) are rejected. Hence, the hypotheses H_1 (Performance expectancy), H_5 (Hedonic motivation), H_6 (Price value) and H_7 (Habit) are supported at 95 % confidence level. Based on Table 4, the equation for the regression line is as follows:

$$\begin{aligned} \text{IU} = {} & 0.463 + 0.189\ \text{PE} + 0.033\ \text{EE} + 0.016\ \text{SI} + 0.054\ \text{FC} \\ & + 0.27\ \text{HM} + 0.119\ \text{PV} + 0.255\ \text{H} \end{aligned}$$

In predicting the X-Space, use behaviour (UB) among students, IU, FC and H are used as the predictors. Table 5 shows the descriptive statistics for the variables

Table 1 Reliability for constructs understudy

Construct	PE	EE	SI	FC	HM	PV	H	IU
Number of items	4	4	2	3	3	2	2	4
Cronbach's alpha	0.858	0.862	0.866	0.857	0.947	0.905	0.858	0.930

Table 2 Descriptive statistics

Variable	Mean	Std. deviation
IU	4.1575	0.73973
PE	3.9786	0.66225
EE	4.0107	0.62835
SI	3.5443	0.81230
FC	3.9093	0.70802
HM	4.1244	0.78424
PV	3.7569	0.95997
H	3.8471	0.83511

Table 3 Analysis of variance

Model	Sum of squares	*df*	Mean square	*F*	Sig.
Regression	111.576	7	15.939	76.103	0.000[a]
Residual	66.813	319	0.209		
Total	178.389	326			

[a] 0.05

Table 4 Regression coefficients

Model	Unstandardized coefficients		Standardized coefficients	*t*	Sig.	95.0 % confidence interval for *B*	
	B	Std. error	Beta			Lower bound	Upper bound
(Constant)	0.463	0.183		2.526	0.012	0.102	0.824
PE	0.189	0.055	0.169	3.440	0.001	0.081	0.297
EE	0.033	0.058	0.028	0.579	0.563	−0.080	0.147
SI	0.016	0.038	0.017	0.418	0.677	−0.059	0.091
FC	0.054	0.053	0.052	1.022	0.308	−0.050	0.159
HM	0.270	0.044	0.286	6.093	0.000	0.183	0.357
PV	0.119	0.033	0.154	3.548	0.000	0.053	0.185
H	0.255	0.041	0.288	6.211	0.000	0.174	0.336

understudy. 6.2 % of the variability in UB is accounted by BI, FC and H. In Table 6, the F test indicates that the model is a good fit for the data. Thus, the independent variables IU, FC and H can be used to explain the dependent variable UB. The results also demonstrate that only H has a positive influence on the use

Table 5 Descriptive statistics

Variable	Mean	Std. deviation
IB	3.4526	0.66234
IU	0E−7	0.97223691
FC	3.9093	0.70802
H	3.8471	0.83511

Table 6 Analysis of variance

Model	Sum of squares	d*f*	Mean square	*F*	Sig.
Regression	8.813	3	2.938	7.071	0.000[a]
Residual	134.202	323	0.415		
Total	143.015	326			

[a] 0.05

Table 7 Regression coefficients

Model	Unstandardized coefficients		Standardized coefficients	*t*	Sig.	95.0 % confidence interval for *B*	
	B	Std. error	Beta			Lower bound	Upper bound
(Constant)	2.673	0.300		8.897	0.000	2.082	3.264
IU	−0.014	0.054	−0.021	−0.259	0.796	−0.120	0.092
FC	−0.009	0.064	−0.010	−0.138	0.890	−0.135	0.118
H	0.212	0.061	0.267	3.481	0.001	0.092	0.331

behaviour (UB) of students. Hence, the hypothesis H_9 is supported at 95 % confidence level. Based on Table 7, the equation for the regression line is as follows:

$$\text{UB} = 2.673 - 0.014\ \text{IU} - 0.009\ \text{FC} + 0.212\ \text{H}$$

5 Conclusions

In summary, H_1, H_5, H_6, H_7 and H_9 are supported. Factors that affect students' intention to use X-Space are performance expectancy, hedonic motivation, price value and habit. Habit also positively affects students' use behaviour in using X-Space.

Understanding the factors that affect students' acceptance and usage of X-Space will enable policy makers to better design and implement collaborative classroom facilities. The result obtained provide useful guide for university to design and implement technologies-rich collaborative classroom. The limitation of the current study is that the study is focusing on Foundation students only, and it is difficult to make generalization. For further research, we intend to extend this study to

investigate the factors that affects lecturers' acceptance and usage of X-Space. In conclusions, it is hope that the results can contribute to the existing knowledge of ergonomics in learning spaces and thus improve students' performance in their studies through active and collaborative learning.

References

Alawadhi, A., & Morris, A. (2008). The use of the UTAUT model in the adoption of E-government services in Kuwait. In *Proceedings of the 41st Hawaii International Conference on System Sciences*, Hawaii, 2008.

Biemans, M., Swaak, J., Hettinga, M., & Schuurman, J. G. (2005). Involvement matters: The proper involvement of users and behavioural theories in the design of a medical teleconferencing application. In *Proceeds of GROUP'05*. Sanibel Island, Florida, USA, Nov 6–9, 2005

Cheng, D., Liu, G., & Qian, C. (2008). On determinants of user acceptance of internet banking: A theoretical framework and empirical study. In *Advanced management of information for globalized enterprises*, AMIGE 2008.

Folkins, J. W., Friberg, J. C., & Cesarini, P. A. (2015). University classroom design principles to facilitate learning. *Planning for Higher Education Journal., 43*(2), 45–62.

Hsu, H., & Chang, Y. (2013). Extended TAM model: Impacts of convenience on acceptance and use of Moodle (online submission).

Jayaratne, I. L. K., & Fernando, D. N. (2009). *Ergonomics related to seating arrangements in the classroom: Worst in South East Asia? The Situation in Sri Lankan school children* (pp. 409–420). IOS Press.

Kennedy, M. (2013). *Filling the space* (pp. 36–39). American School & University www.Asumag.com

Khan, H. (2011). Acceptance of technology among students in Pakistani universities: A test of technology acceptance model. *Journal Of Information Systems Technology & Planning, 4*(8), 52–60.

Kriek, J., & Stols, G. (2010). Teachers' beliefs and their intention to use interactive simulations in their classrooms. *South African Journal Of Education, Abstract, 30*(3), 439–456.

Lau, S., & Woods, P. C. (2008). An investigation of user perceptions and attitudes towards learning objects. *British Journal Of Educational Technology, 39*(4), 685–699.

Lippman, P. C. (2015). Designing Collaborative spaces for schools. *T.H.E. Journal*, 39–44.

Ramsden, B. (2011). Evaluating the impact of learning space. *Reference Services Review, 39*(3), 451–464.

Todhunter, B. J. (2015). An examination of the views of key stakeholders on the development of learning spaces at a regional university. *Journal of Facilities Management, 13*(2), 204–222.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478.

Venkatesh, V., Thong, J. I. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly, 36*(1), 157–178.

Wang, Z. (2008). Smart spaces: Creating new instructional space with smart classroom technology. *New Library World, 109*(3/4), 150–165.

Wilkinson, M. E. C. (2001). Designing the electronic classroom: Applying learning theory and ergonomic design principles. *Library Hi Tech, 19*(1), 77–87.

Wong, K., Teo, T., & Russo, S. (2013). Interactive whiteboard acceptance: Applicability of the UTAUT model to student teachers. *Asia-Pacific Education Researcher, 22*(1), 1–10. doi:10.1007/s40299-012-0001-9

An Adaptive Tool for Learning

Jer Lang Hong

Abstract Various approaches have been made to make students adaptively learn through online learning. With the advancement of Internet technologies and the wide application of e-learning tools, it is certainly possible to provide effective learning tools for younger generation of learners. The younger generation of learners are more Internet savvy, and they tend to learn through online material which provides easy-to-use menus and functionality. However, few online tools provide adaptive learning materials to these generations of learners, and most of them are rather static in nature and provide simplistic functions. In fact, studies have shown that different learners have different learning abilities, and thus, they require a different set of learning materials. In this paper, we proposed a novel adaptive learning tool which could effectively gauge the user's learning behavior and adapt the content material to suit his needs. Our preliminary study shows that the users show positive response to our tool.

Keywords Web 2.0 · E-learning · Adaptive behavior

1 Introduction

The introduction of Web 2.0 has enabled Web browsing and navigation easier. From the traditional static Web to dynamic server-based Web, users can now generate their own content by posting their opinions through various sources such as blogs and forums (Figs. 1, 2, and 3). With the advancement of Internet technologies, Web users are able to experience a more personalized Web browsing and cater for their usage and needs. Recently, researchers incorporate learning tools

J.L. Hong (✉)
Taylor's University, Subang Jaya, Malaysia
e-mail: JerLang.Hong@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_7

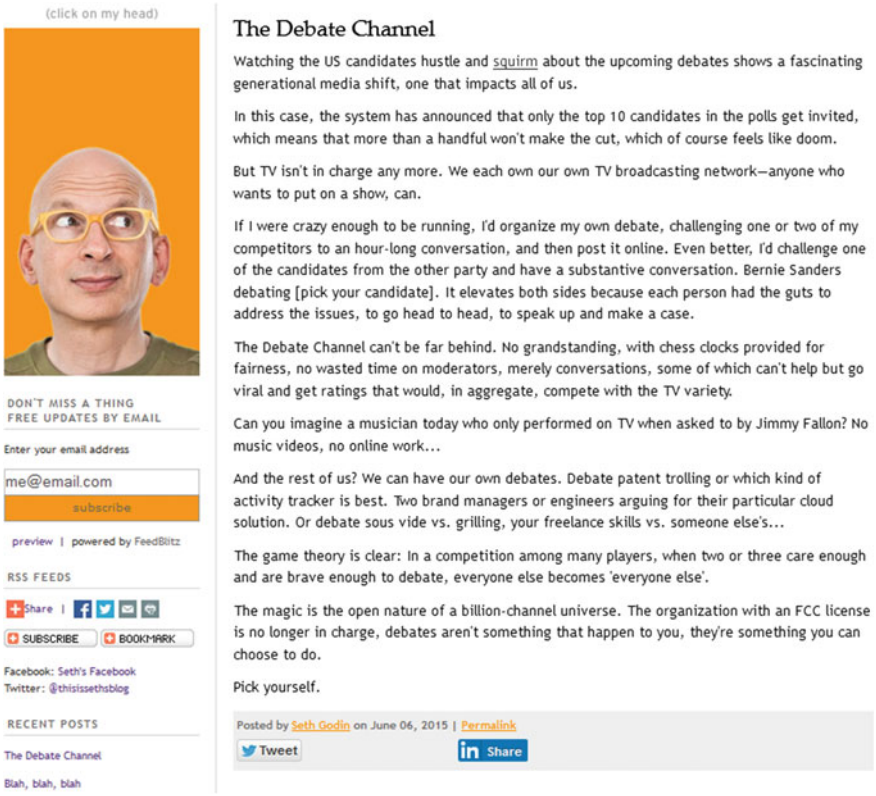


Fig. 1 Blog site

using the dynamic features of Web browsing. For example, in the recent Graduate Record Examination (GRE), an online adaptive examination is presented whereby the examination difficulty level is set according to the user answer (Fig. 4). Though different users answered different sets of questions according to their ability and they obtained different scores at the end of the examination, GRE has managed to show that the generation of questions using a set of predefined rules is able to differentiate the users' skills accurately. Inspired by the adaptive nature of GRE examination online tools, we proposed a novel adaptive tool for e-learning whereby users are able to learn through dynamically generated contents by browsing the Web. Statistics have shown that there is an increase in adoption of e-learning tools. In fact, research has shown that e-learning is proven to improve individual learning skills, especially the younger generation.

The paper is divided into several sections. The next section describes research related to ours, while the subsequent section explains the methodology in detail. Then, we present our experimental results, and finally, we conclude our work.

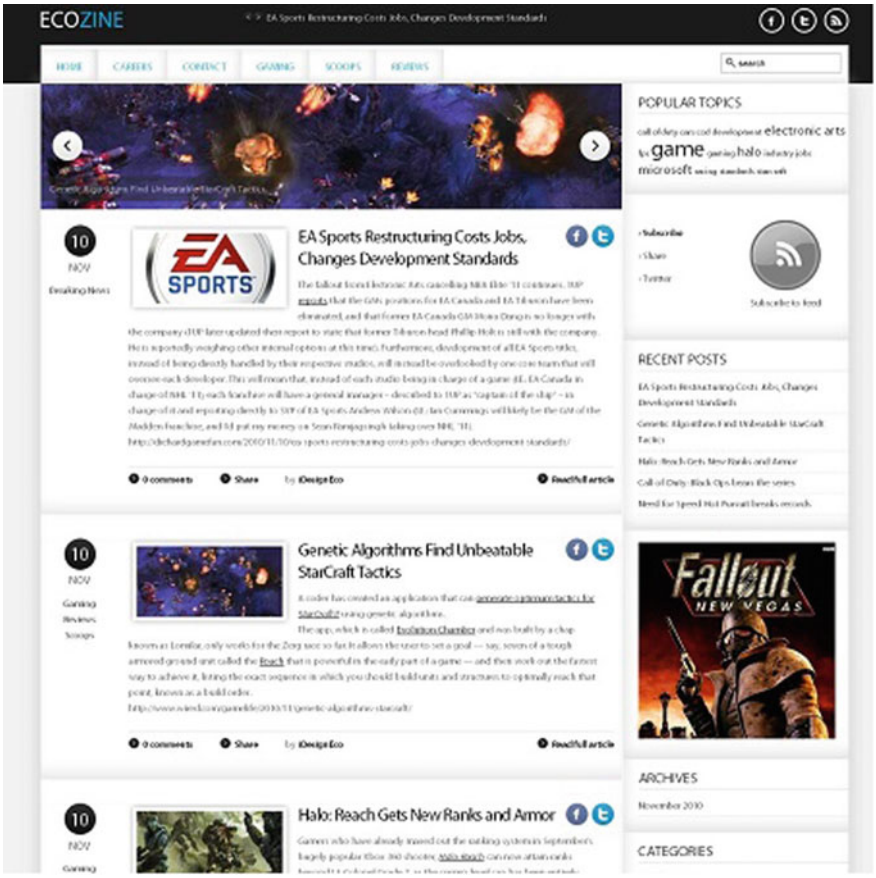


Fig. 2 WordPress example

2 Related Work

Snow develops an adaptive e-learning tool by combining perspective of differential psychologist, experimental psychologist, personality psychologist, and cognitive scientist. Recently, work has been developed in examining the ATI issue using computer as the environment (Maki and Maki 2002; Sternberg 1999). Shute proposed a model where different users with different learning skills learn through an e-learning environment (Shute 1993). Examples of relevant characteristics include incoming knowledge and skills, cognitive abilities, personality traits, learning styles, and interests. To maximize the system performance, one should leverage these different skills and make full use of it. This includes observing the degree of learning and the extent where it can further be developed.

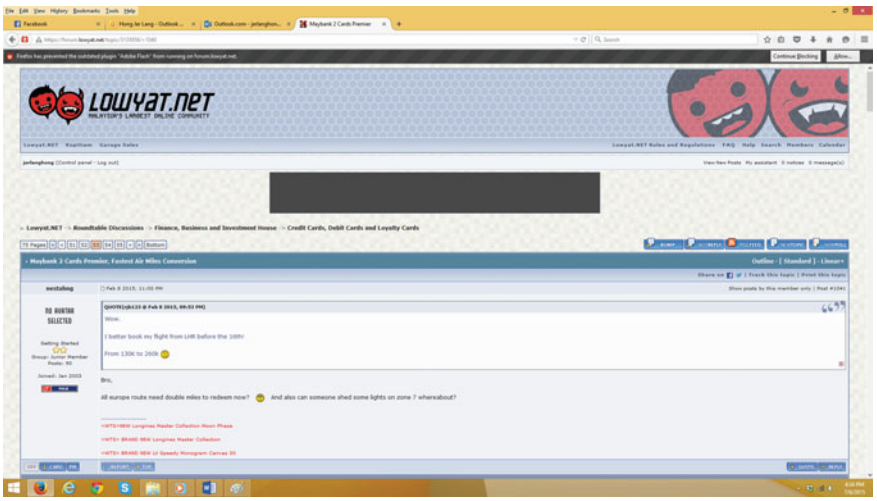


Fig. 3 Lowyat forum

Fig. 4 GRE sample questions

103. Concerning the transcription of a *lac* operon, the synthesis of mRNA starts from which part of the operon?
A. From the beginning of the promoter
B. From the middle of the promoter
C. From the beginning of the operator
D. From the middle of the operator
E. From the beginning of the *lac* Z gene

104. Self-splicing, which does not rely on protein-based enzymatic activities, is found in the splicing of:
A. Double-stranded RNAs.
B. mRNA precursors.
C. tRNA precursors.
D. rRNA precursors.
E. snRNAs.

Furthermore, individuals differ in how much they learn and what they learn, and the different outcomes of their learning reflect the differences in learning processes. The implications mapping the different skill sets of users have been discussed in the literature of (Ackerman 1996, 2003). Recent advances in cognitive science, psychometrics, and technology have make it possible to assess higher-level learning skills (Hambleton 1996; Mislevy et al. 1999) compared to earlier traditional methods such as paper-and-pencil multiple-choice tests, new

assessments for complex cognitive skills involve embedding assessments directly within interactive, problem-solving, or open-ended tasks (e.g., Bennett and Persky 2002; Mislevy et al. 2001).

3 Proposed Methodology

3.1 Overview

Basically, our system comes with an intuitive yet user-friendly interface (Fig. 5). The interface provides easy-to-navigate menus as well as dynamically generated contents. The system is equipped with backend engine which is able to process the user's request, learn, and adapt to the user's needs. The system also has local storage which is able to store every user's activity and keep a log of it. The idea behind the adaptive learning system is threefold. First, we are of the opinion that different users have different interest and learning curve. Therefore, providing a highly dynamic learning material suits the user's needs more. Secondly, the users are able to grasp the learning material and learn progressively should the materials presented dynamically. Finally, adapting the learning materials promotes more user engagement, and it helps to understand the user's needs better.

3.2 Interface

For the interface part, the user will go through easy-to-navigate menus to select his preference of learning materials. Once the user has selected his preference, a

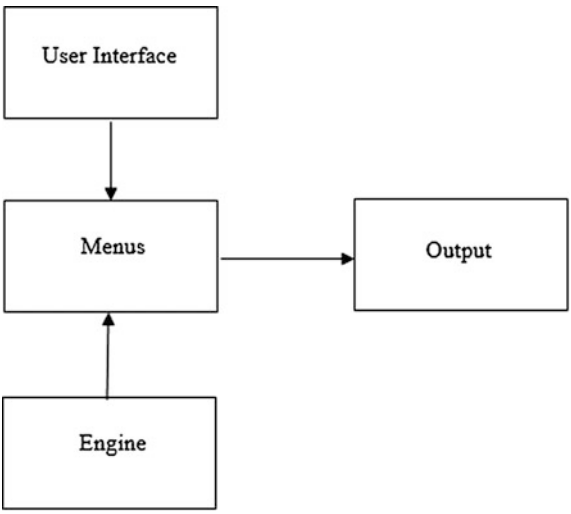


Fig. 5 Architecture of our system

start-up menu will be given where an overview and tutorial on how to use and navigate the system are provided. Then, the learning material is provided to the user in animated forms. The user can then navigate through the learning materials by simple browsing.

3.3 Adaptive Engine

Our system incorporates a backend engine which could dynamically adapt to the user's browsing behavior. Leveraging on the common user's behavior such as clicking, scrolling, and typing behaviors, our engine is able to deduce the typical user interaction, hence deducing the interest of the users. Inspired by machine learning and its ability to deduce user's activity, we incorporate Markov chain and genetic algorithm to learn the user's activity and adjust the browsing experience accordingly.

A thorough investigation and careful examination show that user's activity is a complex problem, and hence, a proper machine learning is required. A graph of user's activity is constructed, and a careful analysis is deduced. Our observation shows that the higher number of clicks generally indicates higher interest among users, though in some cases this may not be necessarily true. A scrolling behavior normally indicates better interest among users.

3.4 Output

The system outputs a log file and stores the activity of users in XML form. Users' activities are important for future usage as well as further analysis. When the user logs in to the system again in the future, his personal data and activity are loaded from the log file and personalization will be carried out where necessary.

4 Experiments

We conduct experiments to gauge the effectiveness of our system. We randomly select 20 different users from wide range of background and skills to test our system. The demographic of our users is as.

As is shown by the Tables 1 and 2, we choose random sample of users from diverse background. Preference was taken to select users who are IT savvy and in the age group of 20–25. Users are given instructions on how to use the system as well as briefed on the setup. They are given a grace period of 2 h to use the system and give feedback on it. A survey form is given to them to fill up the necessary details.

Table 1 Users' age group

Age range	Number
15–20	3
20–25	10
25–30	3
30–35	2
>35	2

Table 2 Skill set of users

Skills	Number
Engineering	3
Computing	10
Business	5
Others	2

Table 3 User-friendliness test

Rating	Number
Poor	0
Not satisfactory	2
Neutral	3
Satisfactory	12
Excellent	3

Table 4 Performance

Rating	Number
Poor	0
Not satisfactory	3
Neutral	3
Satisfactory	10
Excellent	4

Table 5 Learning experience

Rating	Number
Poor	0
Not satisfactory	1
Neutral	5
Satisfactory	9
Excellent	5

On the whole, users find our system to be user-friendly, and our system comes with an interactive GUI, together with easy-to-navigate menus (Table 3). Some of the users commented on the color and font size of the system, whereby a more appropriate color scheme can be used with an adjustable font size. The system is able to run effectively without any crash incident reported, and neither bug is detected (Table 4). Table 5 shows the learning experience of the users. As shown in

the table, users generally learn new knowledge through the easy-to-navigate menus as well as the interactive interface. A few comments have been given by the users over the fixed navigation provided (e.g., the sequence of menus seems to be rather constant after navigating).

5 Conclusion

We present a novel adaptive learning tool whereby the system is able to dynamically adapt to different sets of user activity and response. The system is equipped with personalization tool where the user behavior is captured and various relevant responses are given based on the users' need. Experimental test shows that our system is responsive to the users' need and is able to adapt well across wide range of user's responses.

References

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence, 22*, 227–257.

Ackerman, P. L. (2003). Aptitude complexes and trait complexes. *Educational Psychologist, 38*, 85–93.

Bennett, R. E., & Persky, H. (2002). Problem solving in technology-rich environments. In Qualifications and Curriculum Authority (Ed.), *Assessing gifted and talented children* (pp. 19–33). London, England: Qualifications and Curriculum Authority.

Hambleton, R. K. (1996). Advances in assessment models, methods, and practices. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 889–925). New York: American Council on Education/Macmillan.

Maki, W. S., & Maki, R. H. (2002). Multimedia comprehension skill predicts differential outcomes of web-based and lecture courses. *Journal of Experimental Psychology: Applied, 8*, 85–98.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (1999). *On the roles of task model variables in assessment design* (CSE Tech. Rep. No. 500). Los Angeles: University of California, Center for the Study of Evaluation, Graduate School of Education & Information Studies.

Mislevy, R. J., Steinberg, L. S., Almond, R. G., Haertel, G., & Penuel, W. (2001). *Leverage points for improving educational assessment* (CSE Tech. Rep. No. 534). Los Angeles: University of California, Center for Studies in Education/CRESST.

Shute, V. J. (1993). A comparison of learning environments: All that glitters. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 47–74). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Sternberg, R. J. (1999). *Thinking styles*. New York: Cambridge University Press.

Integrating Social Networking Sites with E-Learning

Jer Lang Hong

Abstract Recent advancement has led to the rapid growth in social networking sites. From the traditional methods of communicating via mail and online chat, users are now able to socially connect with each other globally via social networking sites. For example, users using Facebook are able to chat with their friends, post message, and view the news feed regularly. In this paper, we hope to integrate Facebook into e-learning so that users are able to use Facebook to learn their subject materials. Inspired by the many features provided by Facebook, we are of the opinion that the newsfeed, friends listing and application platform can be integrated into a learning platform. Our key observations showed that users read news feed very regularly and chat with their close friends frequently. Thus, creating a Facebook application that utilizes the news feed and friend listing features fully can ideally make a powerful learning platform for the user. Through the comprehensive API provided by Facebook, we are able to harvest the learning materials and gauge the users' learning ability.

Keywords Web 2.0 · Social networking sites · E-learning

1 Introduction

Social networking sites have enjoyed a steady and rapid growth since the late 1990s. With the introduction of Facebook (Fig. 1), Twitter (Fig. 2), Google+ (Fig. 3), and Pinterest (Fig. 4), these social networking sites have since obtained a tremendous amount of new users, and to facilitate the rapid surge of online users, these sites incorporate and introduce new features convenient for the users. Many business domains and governmental organizations have since relied on the data made available from these sites. Apple, for example, relies on data obtained from Facebook for its marketing strategy. Data available from social networking sites are

J.L. Hong (✉)
Taylor's University, Subang Jaya, Malaysia
e-mail: JerLang.Hong@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_8

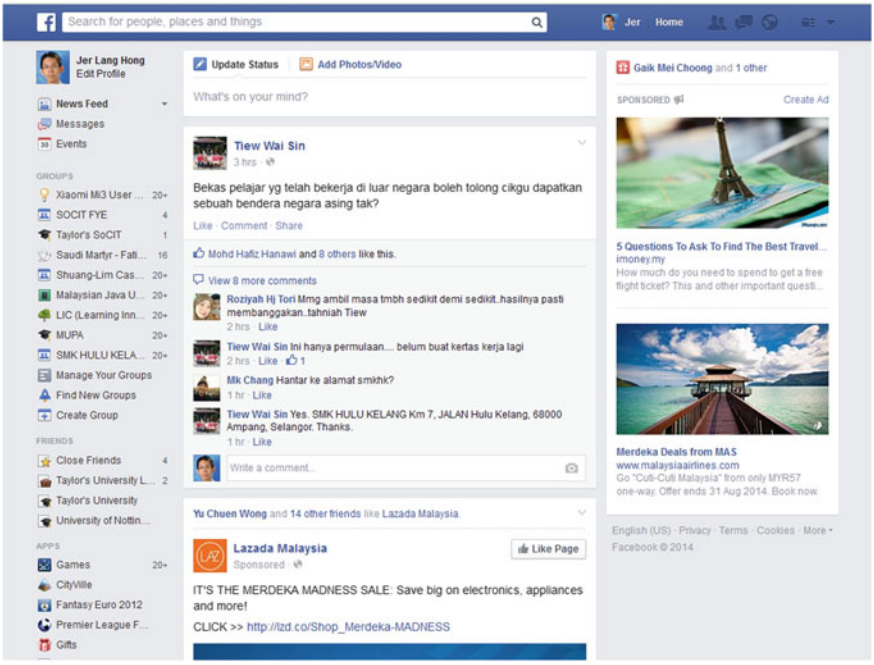


Fig. 1 Facebook example

also useful for other applications, such as poll prediction and counter intelligences. For example, in the recent election of Barrack Obama, the US government uses the data available from social networking sites to obtain feedback on the public view of the current administration. The citizens' right, concern, and their opinion are used to further strategies in the election campaign.

Recent advancement in social networking sites has enabled the users to socially connect and interact with each other. Users can easily connect to their social networking platforms using their smart devices in any location of their choice. These social networking sites not only serve the users need, but also provide third-party integration to facilitate application support. Many researches have been done to integrate and use social networking sites such as business analytics.

On the other hand, e-learning technologies have also grown, and this growth has been tremendous that students nowadays learn through online material rather than traditional textbook. Various advancements have been made to facilitate e-learning processes, such as the introduction of e-book and digital classroom. Coupled with the fact that students nowadays are IT savvy and has full access to social networking platforms, and it is therefore essential to incorporate e-learning methods in social networking sites.

In this paper, we proposed a novel social networking platform which provides full integration with e-learning technologies. After thorough investigation and

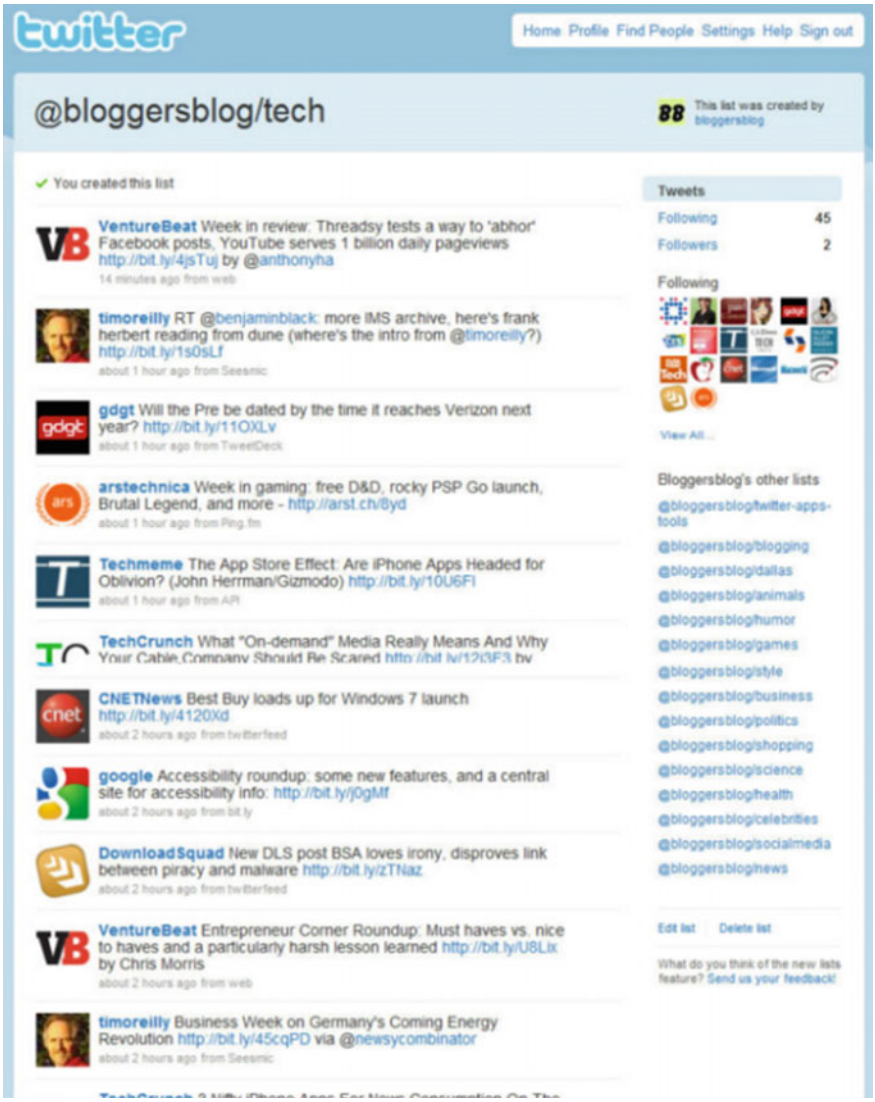


Fig. 2 Twitter example

careful examination, we find that it is possible to integrate Facebook with e-learning tools. Inspired by the extensive libraries and application support provided by Facebook, we aim to develop a Facebook application that utilizes the primary features such as Friend List and News Feed which would eventually provide and facilitates e-learning process. Our tool provides state-of-the-art e-learning features which coupled well with Facebook platform.

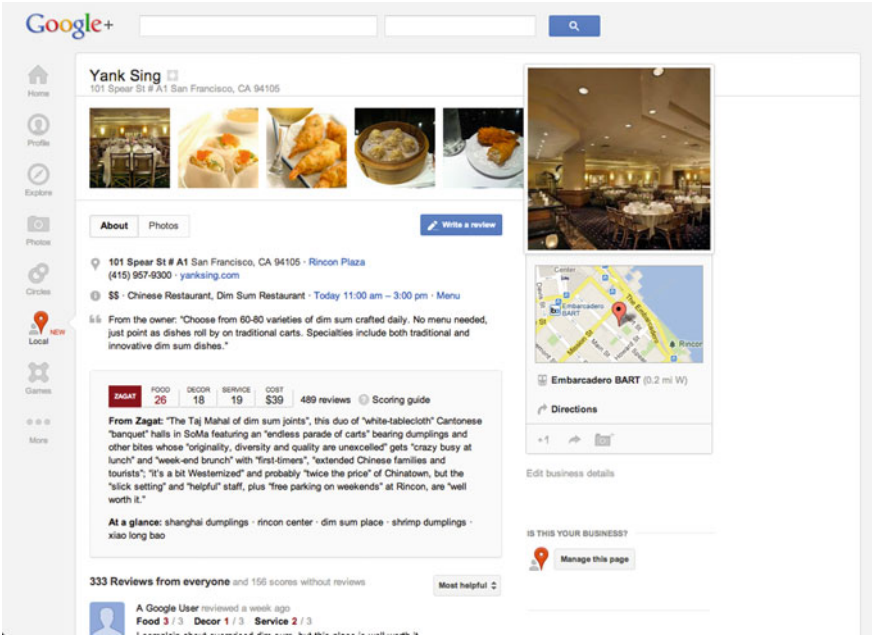


Fig. 3 Google+ example

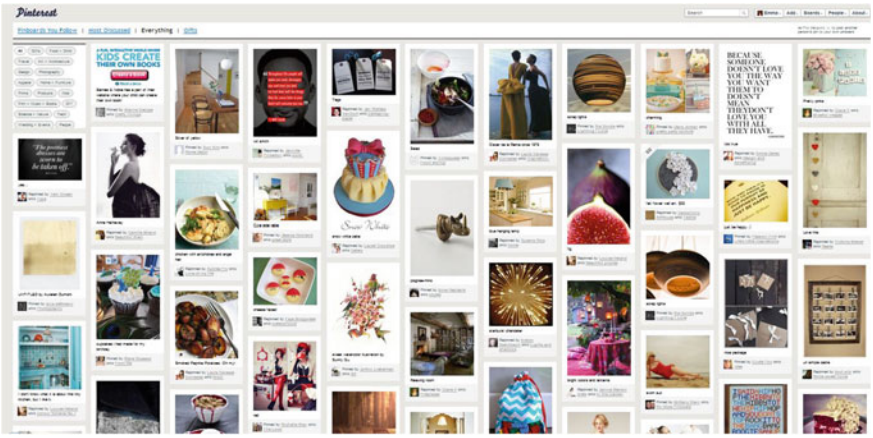


Fig. 4 Pinterest

The paper is divided into several sections. The next section describes research related to ours, while the subsequent section explains the methodology in detail. Then, we present our experimental results, and finally, we conclude our work.

2 Related Work

In the middle 2000s, Wiki has been recognized by the education sector that offers collaboration and constructive learning for the field of education (Klobas 2006). Wiki is also widely used in various disciplines such as teaching and learning, with application across various levels of education stages: (1) primary schools (Desilets and Paquet 2005), (2) secondary schools (Lund and Smordal 2006; Lombard 2007), and even in (3) higher level education (Augar et al. 2004; Cubric 2007; Doolan 2007; Wheeler et al. 2008).

There are various types of learning activities being planned and implemented using Wiki, for example, evaluation and review of journal articles (Forte and Bruckman 2006; Cubric 2007), storytelling (Desilets and Paquet 2005), essay writing (Wang et al. 2005; Forte and Bruckman 2006), and Wiki-based glossary of technical terms (Glogoff 2006; Cubric 2007). In 2000, five-stage model (Salmon 2000) was originally developed to support the role of e-moderator and scaffold students on a step-by-step basis to learn how to interact with other users online. It focused more on the role of tutors than the role of students.

Garrison et al. developed COI in 2000. The original conception of COI was to improve the practice of users in evaluating computer science conferences and also to help to design, facilitate, and direct higher order learning. It is an online tool with student participation; however, it does not offer a stage-by-stage guideline to augment interaction and learning for the users to use Web 2.0 technologies.

Four-stage online presence (4-SOP) model was developed by Goh (2010), and it is derived from theoretical lens of the five-stage model developed by Salmon (2000), Garrison et al. COI (Garrison et al. 2000). The themes are based on empirical data obtained from three case studies conducted based on UK context. This model consists of four stages and three online tools with user participation to provide a step-by-step guide in introducing the Web 2.0 technologies to the students. The distinctive feature of this model is to assist lecturers in planning, monitoring and reviewing the teaching and learning process at appropriate stage. 4-SOP model also focuses on critical thinking skills as the expected learning outcomes.

3 Proposed Methodology

3.1 Overview

We integrate our e-learning tool with that of Facebook using Facebook application development. Using their extensive API lists and the libraries' support, we develop a Web-based application that fully integrates with Facebook. The entire application is written in HTML and SDK provided by Facebook. Careful consideration is given to the UI design and theme, and in particular, we stick to the blue theme and the

font style provided by Facebook. We use the functionalities of the Facebook platform by linking to their API via Web services.

3.2 Features

After careful examination of other successful e-learning tools, we identify the features in other existing e-learning tools and incorporate them into our Facebook application. In particular, we aim to implement an interactive e-book in our Facebook application as well as adaptive learning which could effectively adapt to user's behavior upon using the application. Our application provides an intuitive yet simple user interface where the users can easily navigate through the menus to access the learning materials. We provide step-by-step navigation menus and animative icons to make our application attractive. On the other hand, we develop a backend adaptive engine which could effectively adapt to the user's behavior. Various factors for adapting to the user's behavior are used, such as number of clicks, scrolling behavior, and history of access. A weigthing function and formulas are used to determine the user interaction with the application. The weightage changes according to the user's action and interactivity.

3.3 Facebook Integration

To make it fully integrate with Facebook, we develop our application to make full use of Facebook features such as friend lists and newsfeed. We incorporate third-party integration with Facebook by using friend lists to promote our tools as well as learning activity. On the other hand, we adapt our learning content to that of user's newsfeed. We are of the opinion that news feed generated at the users' site can provide valuable information on their interest and it helps us to understand better the users' needs. Information retrieval techniques and segmentation techniques are used to extract and partition the data into logical segments. Then, we use ranking and filtering techniques for data display, where relevant data are taken from the list and display in the screen.

4 Experiments

We conduct experiments to gauge the effectiveness of our system. We randomly select 20 different users from wide range of background and skills to test our system. The demographic of our users are as follows:

As shown in the two tables above (Tables 1 and 2), we choose random sample of users from diverse background. Preference was taken to select users who are IT

Table 1 Users' age group

Age range	Number
15–20	3
20–25	10
25–30	3
30–35	2
>35	2

Table 2 Skill set of users

Skills	Number
Engineering	3
Computing	10
Business	5
Others	2

Table 3 User-friendliness test

Rating	Number
Poor	0
Not satisfactory	3
Neutral	4
Satisfactory	10
Excellent	3

Table 4 Performance

Rating	Number
Poor	0
Not satisfactory	3
Neutral	4
Satisfactory	8
Excellent	5

savvy and in the age group of 20–25. Users are given instructions on how to use the system as well as briefed on the setup. They are given a grace period of 2 h to use the system and give feedback on it. A survey form is given to them to fill up the necessary details.

Table 3 provides the user-friendliness Test on our application. The majority of users feel that it is simple to use our application and our application provides intuitive interface as well as easy to navigate menus. Table 4 gives the performance of our system. As far as we are concerned, our application does not suffer any crash incident and it is also stable due to the tightly coupled integration with Facebook platform. Most of the users acknowledge the effectiveness of our learning tool, especially the adaptive learning techniques with respect to user's newsfeed (Table 5).

Table 5 Learning experience

Rating	Number
Poor	0
Not satisfactory	2
Neutral	4
Satisfactory	10
Excellent	4

5 Conclusion

We have developed a novel e-learning tool that integrates with Facebook. Our application integrates with Facebook functionalities such as Friend Lists and Newsfeed. Not only it provides an intuitive interface as well as an adaptive backend engine that fully understands user's need. The e-learning is tightly coupled with Facebook platform, providing ease of use and convenience for the users. The platform works by having access to Facebook API and uses Web services to fetch and manipulate data. Our study shows that our application is highly effective in adopting a user's learning experience. Users generally are happy with our system, and they find that our system is user-friendly and engaging.

References

Augar, N., Raitman, R., & Zhou, W. (2004). *Teaching and learning online with Wikis*. *Proceedings of the 21st ASCILITE Conference*, (pp. 95–104). Perth.

Chong, M. L., Tan, B. H., & Abdullah, M. H. (2011). Wiki for co-writing a science dictionary. *GEMA Online. Journal of Language Studies*, *11*(3).

Cubric, M. (2007). Wiki-supported blended learning. University of Hertfordshire Annual Learning and Teaching Conference.

Desilets, A., & Paquet, S. (2005). *Wiki as a tool for web-based collaborative story telling in primary school: A case study*. *In Proceedings of Ed-Media*. Montreal Canada.

Doolan, M. (2007). *Collaborative working: Wiki and the creation of a sense of community*. Second International Blended Learning Conference (pp. 72–85).

Forte, A., & Bruckman, A. (2006). From Wikipedia to the classroom: Exploring online publication and learning. In *Proceedings of the 7th International Conference of the Learning Sciences*. Bloomington, IN.

Garrison, D., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, *2*(2–3), 87–105.

Glogoff, S. (2006). The LTC Wiki: Experiences with integrating a Wiki in instruction. In S. Mader (Ed.), *Using Wiki in education*.

Goh, W.W. (2010). The use of web 2.0 technologies in developing and enhancing students' critical thinking skills in higher education: A qualitative study (PhD thesis, University of Derby).

Klobas, J. (2006). *Wikis: Tools for information work and collaboration*. Oxford: Chandos.

Lombard, F. (2007). *Empowering next generation learners: Wiki supported inquiry based learning? European practise based and practitioner conference on learning and instruction*. Maastricht.

Lund, A., & Smordal, O. (2006). *Is there a space for the teacher in a Wiki?* In *International Symposium on Wikis* (pp. 37-46). Odense, Denmark

Prensky, M. (2001). Digital natives, digital immigrants. Retrieved 1 Apr 2007, from http://www.marcprensky.com/writing/

Salmon, G. (2000). *E-moderating: The key to teaching and learning online*. London: Kogan Page.

Wang, H., Lu, C., Yang, J., Hu, H., Chiou, G., & Chiang, Y. (2005). *An empirical exploration of using Wiki in an English as a second language course*. In *Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05)* (pp. 155–157). Kaohsiung Taiwan.

Wheeler, S., Yeomans, P., & Wheeler, D. (2008). The good, the bad and the Wiki: Evaluating student generated content for collaborative learning. *British Journal of Educational Technology*, *39*(6), 987–995.

Zakaria, M. H., Watson, J., & Edwards, S. L. (2010). Investigating the use of Web 2.0 technology by Malaysian students. *Multicultural Education and Technology Journal*, *4*(1), 17–29.

Zin, A. M., Idris, S., & Subramaniam, N. K. (2006). Improving learning of programming through e-learning by using asynchronous virtual pair programming. *Turkish Online Journal of Distance Education*, *7*(3).

A Case Study on Twenty-First-Century Students Skills Assessment Using Morfo

Chandra Reka Ramachandiran and Malissa Maria Mahmud

Abstract Emerging technologies and world globalization has provided unlimited opportunities for the new discoveries in the twenty-first-century learning. In today's classroom, the students are required to use educational and instructional technologies to synthesis the new acquired knowledge, collaborate with peers, solve problems, and make decisions to be successful in the current complex, diverse, and global community. It is not an easy task to achieve sustainability in education as these digitally experienced students learn differently and have new vernaculars. In response to these changes, this research aims to assess the 4Cs in the twenty-first learning skills via Morfo, an online tool used to create avatars. The 4Cs comprise critical thinking, creativity, collaboration, and communication. A total of 18 undergraduate students from a Private Higher Education Institution (PHEI) participated in this case study. The results of the study indicated that this assessment efficiently assesses all the 4Cs for an introductory computing course. It is also noticed that the respondents are able to engage well in twenty-first-century learning tasks that optimize their classroom performance.

Keywords Twenty-first-century learning · Assessment · 4Cs · Morfo

1 Introduction

The emergence of the technology over the decade has changed the way we teach and learn in the twenty-first-century. President Clinton (2005) once quoted:

C.R. Ramachandiran (✉)
Taylor's University Lakeside Campus, Subang Jaya, Malaysia
e-mail: chandrareka@hotmail.com

M.M. Mahmud
Sunway University, Bandar Sunway, Malaysia
e-mail: malissam@sunway.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_9

> Literacy is not a luxury; it is a right and a responsibility. If our world is to meet the challenges of the 21st century, we must harness the energy and the creativity of all our citizens.

Correspondingly, there has been a significant shift from the manufacturing era to the digital era. The information highway, backbone of the digital era, provides a learning platform that is more centralized and thus, promotes globalization. Besides that, the emerging technologies also provide unlimited opportunities for many new discoveries to equip the students better to face the challenges in the twenty-first century. In the twenty-first-century learning classrooms, the students are required to use educational and instructional technologies to synthesis the new acquired knowledge, collaborate with their peers, solve problems, and make decisions to be successful in the current complex, diverse, and global community. The success of such learning style lies in being able to communicate, share, comprehend, and apply the acquired information to solve the complex problems. To add on, students are also required to adapt to the changes in the information age and use the power of technology to create new knowledge as stated by Anderson (2008).

Today, e-learning tool is one of the emerging needs of information age, whereby it replaces the traditional teaching method for the distant students and simultaneously encourages collaborative learning (McArdle et al. 2004; Abdul-Kader 2008). To add on, Rosenberg (2001) affirms that e-learning application is an open system that integrates the access to the information and purposeful communication into a dynamic and intellectual learning community. Therefore, e-learning tools are used widely in the twenty-first-century learning classroom.

1.1 Problem Statement

Education is the process of acquiring knowledge or information. Over the years, it is noticed that the knowledge is growing exponentially and it has impacted on the diversified career specialization. At the same time, the emerging technologies are also the root cause of this phenomenon and this is due to the rapid growth of the knowledge via the information highway. Therefore, the need arises for the students to keep abreast with this new change in the twenty-first century. It is important to identify the appropriate learning styles and assessment methods to assess the essential adaptation skills required to address the new changes in the twenty-first century. The advancement in technology has helped students to acquire more knowledge and provided the facility to learn at their own pace. It has changed the teaching method, whereby the educators play the role of an instructor and promote technology-integrated teaching modules. However, it is not an easy task to achieve sustainability in education in this digital era as these digitally experienced students learn differently and have new vernaculars. Their needs and pace of learning is much diversified than before. Furthermore, educators must be skilled and trained to assess these young minds using the right assessment tool that can capture the 4Cs

such as creativity, communication, collaboration, and critical thinking skills which is within the twenty-first-century skills framework.

1.2 Objectives

In response to these changes, this research aims to assess the common core state standards of the twenty-first-century skills framework, 4Cs via Morfo, an application used to create avatars. In accordance with the objectives of the research study, two hypotheses were identified.

- **Twenty-first-century learning skills hypothesis**

  H1: the Morfo assessment successfully captures the 4C components exhibited by students
  H0: the Morfo assessment failed to capture the 4C components exhibited by students

- **Students performance hypothesis**

  H2: students obtain higher grades in a twenty-first-century learning classroom that incorporates the Morfo assessment
  H0: students do not obtain higher grades in a twenty-first-century learning classroom that incorporates the Morfo assessment.

2 Twenty-First-Century Learning Skills (4Cs)

The framework for twenty-first-century learning is a product that describes the skills a student will need to acquire to thrive in the current complex, diverse, and global community. It is noted that the common core state standards of the twenty-first-century skills framework 4Cs need to be integrated into the classrooms to produce students who are adequately prepared to face the new challenges of the digital era (Van Roekel 2008). Adding on, researcher Wilborn (2013) reported that the twenty-first-century 4C outcomes were clearly supported by assessments, curriculum, and professional development. The mentioned outcomes were designed to promote positive engagement and were align with the partnership for the twenty-first-century learning and innovation skills (Partnership for twenty-first-century Skills 2006). Figure 1 clearly depicts the four essential learning and innovation skills referred to as the 4Cs. They are the critical thinking, communication, collaboration, and creativity.

Critical thinking is one of the key components highlighted in various discussions to twenty-first-century classroom teaching and learning. It is said that the critical thinking includes inductive and deductive reasoning techniques, as well as making

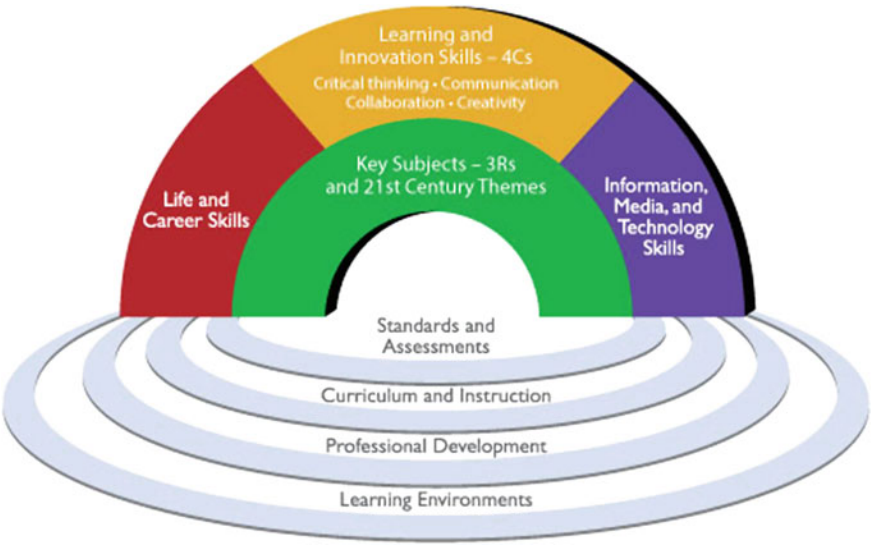


Fig. 1 Framework for twenty-first-century learning (Partnership for 21st Century Skills 2009)

the right analyses, inferences, and evaluations (Facione et al. 1995). The importance of such component is merely to understand the academic content and to relate it to the real-life scenario. Consequently, it will enhance their career performance in the workplace. Past research also suggests that for an organization to compete in the global economy, it needs workers who will think out of the box to continuously improve its products, processes, or services and they are well equipped to give back to the society (Wagner 2010; Peredo and McLean 2006).

On the other hand, communication comprises clarity of information, sharing of information, and the balance among the participants or communicators (Mohr et al. 1996). According to Pellegrino and Hilton (2013), communication is vital to promote teamwork and lies at the core of empathy, trust, conflict resolution, and negotiation. Therefore, effective communication is important to map various ideas together and to facilitate good networking skills with peers or clients. Due to this, there has been an increased focus on this skill in schools. Similarly, collaboration is also given importance in schools as it is known as an additional competencies related to conflict resolution, decision making, problem solving, and negotiation (Lai 2011). Currently, the information highway allows virtual teams to collaborate and spread their ideas around the world. To keep abreast with the globalization and the advancement of technology, employees are expected to have various skills related to effective communication and collaboration as it will enable them to effectively collaborate from any location. Hence, communication and collaboration are identified as an essential twenty-first-century competency by almost all of the organizations who are seeking competent employees.

Many educators and employers believe that creativity is the most outstanding competencies in the twenty-first century. However, it is noticed that this term has

various definitions such as unusualness, appropriateness, and transformation in varying combinations (Jackson and Messick 1971). Over the years, creativity has gained increasing focus in educational institutions globally. For example, Malaysia has devoted resources to fostering more creativity in their teaching and learning styles. Importance is given to creative projects done and the assessment rubrics do integrate creativity as a core assessment component. According to Setiadi and So (2013), the creativity expert Epstein has identified the four competencies that are related to creative expression. The competencies are capturing new ideas, challenging ourselves to solve a problem, broadening our student's interest, and finally associating with diversified people. He also affirms that the most important of these competencies is capturing the ideas and therefore suggests that the educators must provide students with the opportunity to capture their ideas on a daily basis in an idea folder or known as the idea box.

3 Research Methods

To address the objectives of this twenty-first-century learning skill-related research, Morfo was implemented to assess the 4C competencies. According to the official Morfo Web site, Morfo was ranked number 2 under the free Photo App for iPad category. To add on, Morfo was also identified as the top-selling photograph application tool on the new Windows Phone 7 platform, with an average review score of 4.5 out of 5 stars. Furthermore, Morfo also has the capacity to capture any character and convert it to a life-like 3D character.

The respondents for this research consist of 18 undergraduate students from a private higher education institution (PHEI). They were selected based on their level of computer literacy. The selection criteria for the respondents are that they must understand the nature of Web and multimedia technologies. Therefore, this assessment was integrated into a computer proficiency module for business and social science majors. The characteristics of the selected respondents are listed below:

- Gender: Male (67 %); Female (33 %)
- Age group: between 18 and 22 years
- Computer Literacy: Intermediate level of computer proficiency.

The assessment rubric was adapted from the Washington State Office of Superintendent of Public Instruction (OSPI) Web site. OSPI is the state education agency for the State of Washington. This organization implements the laws related to education in the United States of America. It is a holistic assessment on the 4Cs and the rubric is entitled twenty-first-century Skills Standard Rubrics. Since this rubric is being widely used in the USA, therefore, it is reliable and has been validated by the experts from this domain of research. Table 1 summarizes the assessment rubric that has been adapted to the current education system in the PHEI in Malaysia.

Table 1 4C assessment rubric

Learning and innovation: Morfo				
	4 Exceeds standard	3 Meet standard	2 Approaching standard	1 Not at standard
A. Creativity				
Use a wide range of idea creation techniques	Engages in multiple idea creation techniques and fully participates by offering insightful questions and listening well to others in the group	Engages in idea creation techniques and participates by offering ideas	Did offer minimal ideas during the idea creation, but listened actively to other group members	Did not actively participate in idea creation
Creates new and worthwhile ideas	Consistently develops new and valuable ideas using both existing and new knowledge, as well as existing and new resources	Develops new and valuable ideas using both existing and new knowledge, as well as existing and new resources	Develops new and valuable ideas using existing knowledge and resources	Does not attempt to develop new and valuable ideas
Develop, implement, and communicate new ideas to others effectively	Develops, implements, and communicates new ideas to others effectively in a variety of conditions	Develops, implements, and communicates new ideas to others effectively	Develops, implements, and attempts to communicate new ideas to others	Develops new ideas, but does not effectively implement or communicate ideas with others
B. Critical thinking				
Use various types of reasoning as appropriate to the situation	Uses various types of reasoning as appropriate to the situation in a variety of conditions	Uses various types of reasoning as appropriate to the situation	Uses reasoning as appropriate to the situation	Attempts to use reasoning as appropriate to the situation
Effectively synthesizes and makes connections between information and arguments	Is able to apply the connections between information and arguments in order to support a perspective	Is able to understand and make the connections between information and arguments	Is able to understand there is a connection between information and arguments, but is not able to determine what they are on their own	Does not attempt to understand the connection between information and arguments
C. Communication				
Use communication for a range of purposes (e.g., to inform, instruct, motivate, and persuade)	Effectively uses communication to inform, instruct, motivate, and persuade on multiple occasions using both verbal and nonverbal communication	Uses communication to inform, instruct, motivate, and persuade	Communicates only to inform or instruct	Communicates only to inform and does not demonstrate communicating for other purposes

(continued)

Table 1 (continued)

Learning and innovation: Morfo				
	4 Exceeds standard	3 Meet standard	2 Approaching standard	1 Not at standard
Utilize multiple media and technologies, and know how to judge their effectiveness as well as assess their impact	Worked creatively to craft a comprehensive product using multiple media and technologies and thoughtfully reflected on the effectiveness and impact of the product	Crafted a product using multiple media and technologies and reflected on the effectiveness and impact of the product	Crafted a product using multiple media and technologies but did not effectively reflect on the effectiveness and impact of the product	Attempted but did not complete crafting a product using multiple media and technologies and did not effectively reflect on the effectiveness and impact of the product
Communicate effectively in diverse environments	Communicates effectively with others in diverse environments using both verbal and nonverbal communication	Communicates effectively with others in diverse environments	Communicates with others in a diverse environment, but fails to communicate effectively with others	Fails to communicate with others in a diverse environment
D. Collaboration				
Demonstrate ability to work effectively and respectfully with diverse teams	Consistently listens to others. All statements, responses and body language, are respectful and appropriate. Always listens to, shares with, and supports the efforts of others	Listens to, shares, and supports others. Statements and responses are respectful and appropriate body language was exhibited	Most statements, responses, and body language are respectful; occasionally had a negative tone. Does not always listen to, share with, and support the efforts of others	Statements, responses, and/or body language were consistently not respectful. Rarely listens to, shares with, and supports the efforts of others
Assume shared responsibility for collaborative work, and value the individual contributions made by each team member	Motivates all members to share in contributions equally by valuing all members' ideas and contributions.	Participates and contributes to group's work equally. Values all members' ideas and contributions.	Attempts to share responsibility of groups' work, but ends up completing most of the work, without utilizing input of others in group.	Either does most or very little of the group's work and does not share or respect others' ideas

The assessment rubric in Table 1 clearly defines the 4 standards of respondents. A respondent is rated 4 if they exceed standard, 3 if they meet the standard, 2 for approaching standard, and 1 is the lowest rating for not being at standard. The 4Cs are clearly defined to ease the assessment process. The rubric was given to the undergraduate students from the PHEI from various specializations from semester two onwards. The past experiences using ICT in the PHEI enables them to provide rapid response compared to the first-semester students. The research was conducted in the X-Space Collaborative Learning Classroom and precise briefing was given highlighting the objectives of the assignment and the accuracy in the assessment using the given rubric. This research was conducted in a form of a project assessment for the introductory to computing module. The 18 respondents brought their own devices and used the assessment rubric as a guideline to achieve the required standards. The respondents were then divided into 3 groups to promote positive interactions and to increase the visibility of the 4C competencies. The duration given to complete the task was 8 h of classroom contact hours. Each respondent was asked to interact with their peers and to capture their image and to reproduce a 3D character using the downloaded Morfo application. They were instructed to create 3 avatars or Morfo virtual agents and to submit the best-fit virtual agent for evaluation. Figure 2 depicts the various virtual agents created using the Morfo application by the students.

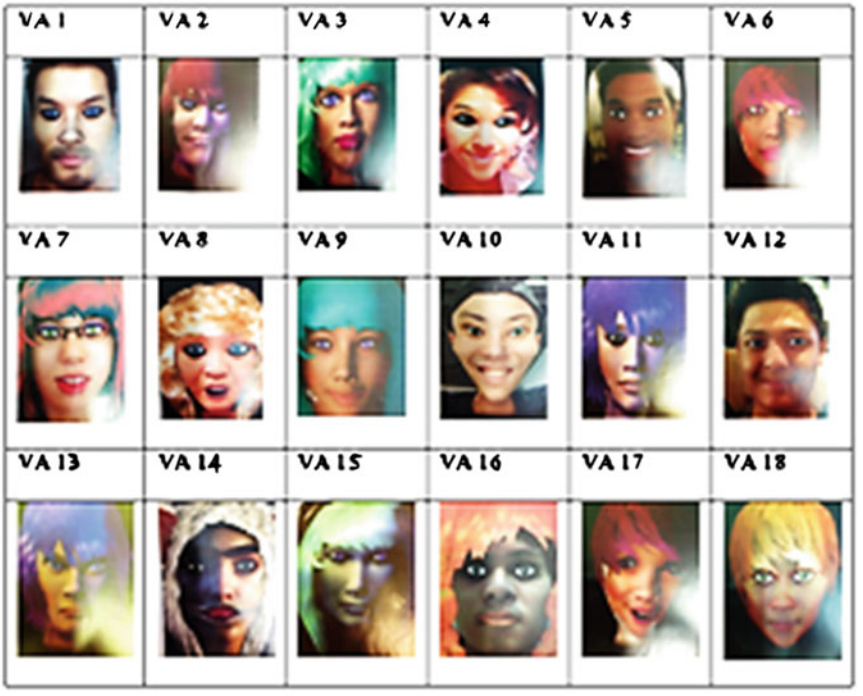


Fig. 2 Morfo virtual agents

Data analysis is the second phase of the research. This is a process, whereby all collected data from the survey are analyzed using the Data Analysis ToolPak, an integrated plug-in tool in Microsoft Excel. Analyses involved calculation of the mean value of the virtual agent assessment and the 4C competencies.

4 Analyses and Results

The PHEI student's active participation in completing the Morfo task indicated there is a positive trend of interaction among the 18 respondents. A total of 16 students (88 %) of PHEI students scored high ratings in the Morfo assessment. The majority of the respondents enjoyed the 8 h of continuous assessment as he agrees that the twenty-first-century Assessment method as an enjoyable learning experience for him. On the other hand, 2 students showed poor participation as they missed classes and did not portray positive interaction with their peers. Table 2 depicts the mean scores obtained for the 4 competencies assessed using Morfo assignment.

Table 2 depicts the mean value obtained by each respondent for the 4 vital competencies measured in this study. The 4 competencies are creativity, critical thinking, communication, and collaboration. 16 students scored more than 3.0 and that denote that they meet the standard for this assessment. To add on, the total score obtained by the respondents for each competency is presented in Table 3.

Table 3 illustrates the 4 different competencies captured in this assessment. It is noted that the majority of the respondents had high ratings for creativity and

Table 2 4C assessment

Student ID	Morfo ID	Mean 4Cs
1	VA01	3.6
2	VA02	3.4
3	VA03	3.3
4	VA04	3.8
5	VA05	3.0
6	VA06	3.8
7	VA07	3.6
8	VA08	3.3
9	VA09	3.5
10	VA10	3.5
11	VA11	3.3
12	VA12	3.3
13	VA13	3.3
14	VA14	2.1
15	VA15	3.0
16	VA16	2.3
17	VA17	3.2
18	VA18	3.3

Table 3 Twenty-first-century students' skills assessment score

Student ID	Morfo ID	Creativity	Critical thinking	Communication	Collaboration	Score
1	VA1	10	8	11	7	36
2	VA2	9	8	11	6	34
3	VA3	9	6	12	6	33
4	VA4	12	8	10	8	38
5	VA5	9	6	9	6	30
6	VA6	12	8	11	7	38
7	VA7	12	7	9	8	36
8	VA8	12	6	9	6	33
9	VA9	12	8	9	6	35
10	VA10	9	6	12	8	35
11	VA11	9	6	12	6	33
12	VA12	9	6	12	6	33
13	VA13	12	6	9	6	33
14	VA14	7	4	6	4	21
15	VA15	9	6	9	6	30
16	VA16	9	4	6	4	23
17	VA17	9	6	9	8	32
18	VA18	9	6	12	6	33

Fig. 3 Morfo virtual agents VA4 and VA6 (high ratings)

communication. Collaboration has the lowest ratings among the respondents. The findings clearly support the hypothesis H1 in the twenty-first-century learning skills hypothesis, in which all the 4C competencies are captured in a single assessment via Morfo.

Among the 18 Morfo virtual agents produced by the 18 respondents, VA4 and VA6 have topped the list for having the highest ratings for the 4C competencies. Both the agents had a total score of 38 from the overall score of 40 marks. Figure 3 depicts the 2 most outstanding Morfo virtual agents with high ratings.

On the other hand, to test the Students Performance Hypothesis, the overall performance for the class was compared with the previous semester score for the same module. It is noted that the respondents who have participated scored better grades in comparison with their seniors. The findings support the H2 hypothesis because they understood what they have practiced during the Morfo assessment and it cultivates the learning interest within them. Figure 4 captures the comparison of grades between the experimented group and the control group. The experimented group comprises 18 respondents and the control group comprises 16 students from the previous semester who are not assessed via the twenty-first-century skills assessment. The maximum and the minimum score are high for the experimental group compared to the control group. Besides that, the mean value depicts the overall student's performance. The experimental groups mean score is 81.01 % where else the control group's mean is 68.06 %. Therefore, hypothesis H2 is supported.

5 Discussion and Conclusion

It is clear that the aim of this study was to identify the efficacy of 4Cs assessment method and the effect of the twenty-first-century learning skills toward their overall performance. The two hypotheses tested are the twenty-first-century learning skills hypothesis and the students performance hypothesis. Table 4 summarizes the findings from the tested hypothesis.

5.1 Twenty-First-Century Learning Skills Hypothesis

The Morfo assessment successfully captures the 4C competencies exhibited by students. This is clearly depicted in Fig. 5. All the 4 competencies are mapped into a single assignment and it is assessed accurately. The outstanding Morfo outcome or virtual agents VA4 and VA6 have the highest score for the 4 vital competencies that comprises creativity, critical thinking skills, communication, and collaboration. The findings also support the findings by Wilborn (2013) who asserted that

Table 4 Summary of hypothesis

Tested hypotheses	Research findings
Twenty-first-century learning skills hypothesis H1: the Morfo assessment successfully capture the 4C components exhibited by students	H1 is supported as the Morfo assessment successfully captured the 4C competencies
Students performance hypothesis H2: e-students obtain higher grades in a twenty-first-century learning classroom that incorporates the Morfo assessment	H2 is supported as the experimental group showed an improved performance

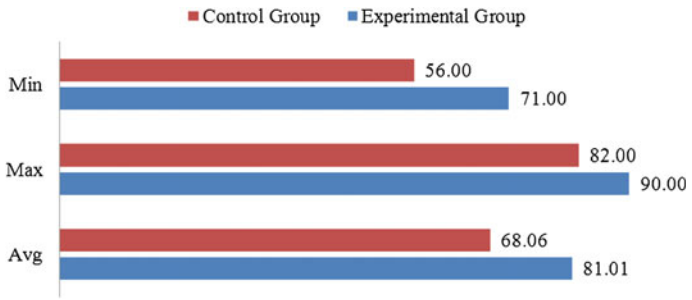


Fig. 4 Comparative analysis of students' performance

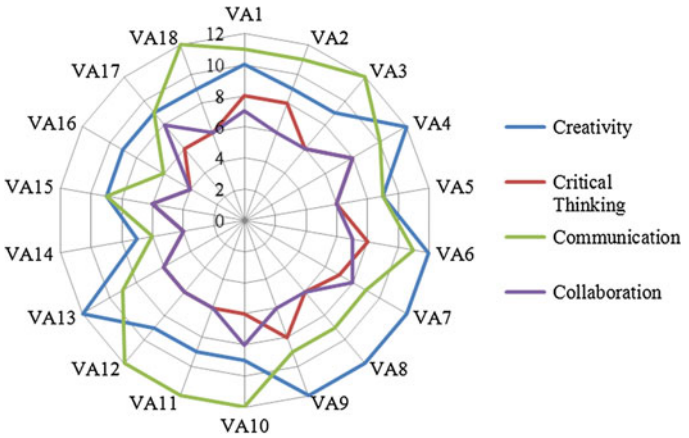


Fig. 5 Reflection of 4Cs in Morfo assessment

twenty-first-century 4C outcomes are clearly supported by effective assessments or classroom activities.

5.2 Students Performance Hypothesis

The research supports the affirmation by Setiadi and So (2013) that these competencies are capable of capturing new ideas, challenging ourselves to solve a problem, broadening our learning interest, and finally associating with diversified people. Hence, there is an increase in the student's performance due to their interest in learning. A respondent also quoted that he enjoyed the lesson as it was a new learning experience for him and he has improved his social skills and made many new friends. The overall class performance showed a positive increase in the mean value of 81 % in comparative to 69 % the previous semester.

In relation to the past review by Rosenberg (2001) who affirms that e-learning application integrates the access to the information and purposeful communication into a dynamic and intellectual learning community, it is important for educators to ensure the efficiency of a teaching-learning tool for assessment. The educators must ensure the infrastructure supports the activity as most of the assessments are technology driven. Align with the rapid ICT development, the students as well as the educators must be at par with the current development to achieve the desired learning outcomes via the right assessment methods as mentioned by Wagner (2010). In this study, Morfo successfully captured the learning objectives as well as managed to assess the 4C competencies of the twenty-first-century learning skills. The research findings also explored the new advancement in assessment in an optimum and conducive learning environment such as X-Space Collaborative Learning classroom for students in the higher education institution.

Though there are many researches being conducted in the area of teaching and learning space, there is a demand to design an assessment that caters to twenty-first-century students who are very technology driven. Other dimensions in assessment should be highlighted to improvise the effectiveness of the delivery medium in the twenty-first-century learning classroom. The findings may also vary from one module to another as different modules have different learning outcomes. Therefore, the role of an informative educator is important to identify the most efficient assessment method and this can only be attained through experimental learning. Conversely, it is a difficult task to understand the psychology of the students from a diversified society such as the Malaysians. Therefore, more research is required to identify the most accurate assessment method to effectively capture the 4C competencies at various education institutions.

References

Abdul-Kader, H. M. (2008). E-learning systems in virtual environment. In *IITI 6th International Conference on nformation & Communications Technology, 2008. ICICT 2008* (pp. 71–76). IEEE.

Anderson, R. E. (2008). Implications of the information and knowledge society for education. In *International handbook of information technology in primary and secondary education* (pp. 5–22). US: Springer.

Facione, P. A., Sánchez, C. A., Facione, N. C., & Gainen, J. (1995). The disposition toward critical thinking. *The Journal of General Education*, 1–25.

Famous Literacy Quotes. (2005). Retrieved On November 26, 2005 from http://www.readfaster.com/LiteracyQuotes.asp

Jackson, P. W., & Messick, S. E. (1971). Toward a theory of creativity. *Psychology: Adapted Readings*, 155.

Lai, E. R. (2011). *Collaboration: A literature review* (Vol. 2). Pearson Research Report.

McArdle, G., Monahan, T., Bertolotto, M., & Mangina, E. (2004). A web based multimedia virtual reality environment for e-learning. *Proceedings of Eurographics*, *4*(1), 9–13.

Mohr, J. J., Fisher, R. J., & Nevin, J. R. (1996). Collaborative communication in interfirm relationships: moderating effects of integration and control. *The Journal of Marketing*, 103–115.

Partnership for 21st Century Skills. (2006). Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century workforce.

Partnership for 21st Century Skills. (2009). P21 framework definitions. Retrieved from http://www.p21.org/storage/documents/P21_Framework_Definitions.pdf

Pellegrino, J. W., & Hilton, M. L. (Eds.). (2013). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academies Press.

Peredo, A. M., & McLean, M. (2006). Social entrepreneurship: A critical review of the concept. *Journal of world business*, *41*(1), 56–65.

Rosenberg, M. J. (2001). *E-learning: Strategies for delivering knowledge in the digital age* (Vol. 3). New York: McGraw-Hill.

Setiadi, N. J., & So, I. G. (2013, August). Assessing creativity skill development in art and design among undergraduate students: Implementing creative potential simulation software to capture creativity-relevant personal characteristics. In *2013 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)* (pp. 268–272). IEEE.

Van Roekel, N. P. D. (2008). Technology in schools: The ongoing challenge of access, adequacy and equity. *National Education Association, Washington DC*. Retrieved from http://www-infocris.isea.org/assets/docs/PB19_Technology08.pdf

Wagner, T. (2010). *The global achievement gap: Why even our best schools don't teach the new survival skills our children need–and what we can do about it*. Basic Books.

Wilborn, J. W. (2013). *Teacher self-efficacy: Common core state standards within a 21st century skills framework* (Doctoral dissertation, Liberty University).

Real-Time Assessment with Nearpod in the BYOD Classroom

Tee Wee Jing and Wong Seng Yue

Abstract This study aims to study the effectiveness of assessment in relationship to student satisfaction in a bring-your-own-device (BYOD) classroom. Specifically, we propose a real-time assessment model using Nearpod in a BYOD classroom. We implemented this real-time assessment model for two consecutive semesters for Foundation in Computing (FIC) students at School of Computing and IT (SoCIT), Taylor's University. In this study, the researchers discussed theoretical and design principles of real-time assessment implementation and how technology-rich learning environment could potentially support to meet these principles. Quantitative and qualitative approaches were used as research methodologies, which included a learning satisfaction survey, interviews, and classroom observations. The research findings provided insight into how faculty conceptualization of assessment methodology drove implementation and showed that real-time assessment in higher education could contribute to student satisfaction.

Keywords Real-time assessment · BYOD

1 Introduction

Our world today is in the golden age of technological innovation, where we embrace daily change by integrating and synergizing with digital technologies. The bring-your-own-device (BYOD) phenomenon began in the workplace, and this trend is currently moving into the classroom of universities. BYOD (Bradley 2011)

T.W. Jing (✉) · W.S. Yue (✉)
School of Computing and IT, Taylor's University Lakeside Campus, Subang Jaya, Malaysia
e-mail: WeeJing.Tee@taylors.edu.my

W.S. Yue
e-mail: SengYue.Wong@taylors.eu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_10

is a concept that allows employees to utilize their personally owned smart devices, i.e., laptops, tablets, smart phones, etc. to connect to company servers, access privileged company information and applications, in order to complete tasks for their organizations. BYOD is used to describe the same practice applied to students using personally owned smart devices in higher education. It is a trend nowadays that higher educational institutes are observing tendency of students and lecturers bringing their laptops, smart phones, and tablets as a resource for enhancing their learning and teaching experience. With the BYOD trend knocking on the doors of the classrooms in universities, it is important for educators to determine the best model on how to integrate BYOD practices into learning and assessment.

2 Literature Review

Generation Y is digital native and grown up with computer games, and they are now studying in higher education. The BYOD (Pope et al. 2005) movement has deep implications toward today's university students, who have embraced technology as a whole. The BYOD idea is that with the advancement of technology, students can access online learning tools and interactive resources ubiquitously with smart devices. Students will not only learn in the lecture halls, but instead use their own devices, i.e. laptops, smart phones, tablets, etc. for lifelong learning purposes.

Classroom management is the recent research interest among educators in universities, with the aim of creating interactive, engaging, and stimulating learning environment. Classroom management (Wang et al. 1993) has the most direct impact on students' learning satisfaction and achievements. Good lecturers (Adelman and Taylor 2005) are able to design good classroom learning activities, with real-time assessment and feedback in order to achieve successful learning outcome and increase students' satisfaction. In classroom management, democratic leadership (Miller and Pedro 2006) with mutual respect and understanding among lecturers and students also plays an important role to ensure students are ready and willing to exchange ideas and increase motivation for learning.

Real-time assessment (McConnell et al. 2003) via smart devices in a BYOD classroom is one feasible method to enhance effectiveness and satisfaction in classroom learning experience. When a lecturer can query and collect real-time responses from every student in the classroom, the lectures can know instantly what students understand, and more important, which concepts they fail to grasp. This real-time feedback enables the lecturers to spend more time to focus on problem areas in order to improve the learning outcomes.

Smart devices (Jelemenská et al. 2011) can also help augment traditional learning experiences. Recently, state-of-the-art technology has been developed to create an interactive environment in the classroom (Pai and Borba 2012). In fact, the use of smart devices in a BYOD classroom creates an interactive and engaging

environment that facilitates students' learning processes (Banister 2010). This model also provides an interactive classroom with real-time feedback (Enriquez 2010). One of the effective and efficient software for real-time assessment is Nearpod, which is developed across all major mobile and computer platforms. It is a synchronous presentation application that allows learners to access real-time quizzes, lecture slides, videos, drawings, votes, and so on, through an interactive presentation created and controlled by the lecturers during the learning experience.

3 Methodology

The research methods used were quantitative and qualitative approaches. For quantitative approach, data were collected from learning satisfaction survey from all students in the classes. As for qualitative approach, we conducted interviews with all students in the classes, and also carried out classroom observations. This research was conducted through a case study involving foundation-level students at the School of Computing and IT, Taylor's University. We implemented this real-time assessment model for two consecutive semesters for FIC students, comprising 25 and 22 students from the August 2014 semester-1 and January 2015 semester-1 intakes, respectively. "Java programming" that is part of the course "Introduction to Computing" was the topic chosen for the study.

For the implementation of real-time assessment model using Nearpod in a BYOD classroom, lecturers used iPads and Nearpod to manage learning contents on students' smart devices synchronously. This model combines presentation, collaboration, and real-time assessment tools into one integrated solution. Students received content and submitted their responses on the Nearpod application; lecturers monitored the classroom activities and received the results of each student in real time. Next, lecturers feedback the results to all students and had a discussion with students to go through the topics that the students were still not familiar with.

4 Results and Discussion

From the quantitative data collected from the learning satisfaction survey done by students, overall results showed encouraging feedbacks from students regarding real-time assessment with Nearpod in the BYOD classroom in order to enhance their learning experience. The scale of this survey is from a scale of 5 (Strongly Agree), 4 (Agree), 3 (Neutral), 2 (Disagree) to 1 (Strongly Disagree). The feedbacks are shown in Table 1.

From qualitative aspect, data from interview with students showed that the majority of students are motivated to learn more while they are engaging in real-time assessment with Nearpod in the BYOD classroom. Students feel that this

Table 1 Feedbacks from learning satisfaction survey

Feedbacks from students in the class	5 (%)	4 (%)	3 (%)	2 (%)	1 (%)
Real-time assessment with Nearpod in the BYOD classroom is highly interactive, more engaging, and enriches students' learning experience	57.45	36.17	6.38	0.00	0.00
Real-time assessment with Nearpod in the BYOD classroom with real-time feedback is an effective and efficient learning and teaching method to improve students' understanding of the topic	48.94	40.43	8.51	2.13	0.00
Real-time assessment with Nearpod in the BYOD classroom with real-time feedback increases the students' interest to learn, which in turn enables them to achieve better academic results	44.68	38.30	14.89	2.13	0.00

real-time assessment model is highly interactive, more engaging, and enriches students' learning experience. This is supported by below feedbacks from students:

- "Very nice assessment method to replace the current way of how we handle classroom."
- "Interesting. Should be done more often. I love interactive classrooms."
- "It's great that everyone can participate in this kind of learning environment!"
- "I think it is quite interesting as the lecturers are able to engage the users through various activities which makes teaching more fun/ attractive."

Observation from lecturers showed that the real-time assessment with Nearpod in the BYOD classroom with real-time feedback assesses student knowledge of the subject in real time and thus provides an opportunity for discussion within the classroom, which help to achieve better learning outcomes. This was supported by the fact that less motivated students, who had less engagement in the beginning, had showed a significant improvement in their learning attitude after they got involved in the real-time assessment. On top of that, our observation also showed that real-time assessment actually motivates the students and inspires them to work harder and smarter. This is because the students desire to have better results due to peer pressure and influence during the process of real-time feedback, where everyone in class can see their results.

5 Conclusion

The preliminary results in this study have demonstrated that real-time assessment using Nearpod in a BYOD classroom can enhance learning experience and contribute to student satisfaction. In fact, the use of the smart devices with Nearpod increases the interaction between the lecturers and the students rending the class more dynamic and tailored in real time to the students' needs. These initial results

will be followed by a more in-depth evaluation of the implementation of real-time assessment using Nearpod in a BYOD classroom conducted over a course length to confirm that these positive results will contribute to the students' interest to learn, which in turn enable them to achieve better academic results.

References

Adelman, H. S., & Taylor, L. (2005). Classroom climate. In S. W. Lee, P. A. Lowe, & E. Robinson (Eds.), *Encyclopedia of School Psychology*. Thousand Oaks, CA: Sage.

Banister, S. (2010). Integrating the iPod touch in K 12 education: Visions and vices. *Computers in the Schools, 27*(2), 121–131.

Bradley, T. (2011, December 20). Pros and cons of bringing your own device to work. *PCWorld*. Re-trieved from http://www.pcworld.com/article/246760/pros_and_cons_of_byod_bring_your_own_device_.html

Enriquez, A. G. (2010). Enhancing student performance using tablet computers. *College Teaching, 58*(3), 77–84.

Jelemenská, K., Cicák, P., & Dúcky, V. (2011). Interactive presentation towards students' engagement. *Procedia-Social and Behavioral Sciences, 29*, 1645–1653.

Pai, D., & Borba, G. (2012). The role of digital technologies for the innovation of the learning experience in the university classroom, *Strategic Design Research Journal*, *5*(2) (May–August 2012).

Pope, M., Hare, D., & Howard, E. (2005). Enhancing technology use in student teaching: A case study. *Journal of Technology and Teacher Education, 13*(4), 573–618.

Miller, R., & Pedro, J. (2006). Creating respectful classroom environments. *Early Childhood Education Journal, 33*(5), 293–299.

McConnell, D., Steer, D., & Owens, K. (2003). Assessment and active learning strategies for introductory geology courses. *Journal of Geoscience Education, 51*, 205–216.

Wang, M. C., Haerted, G. D., & Wallberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63*(3), 249–294.

Partner in Speak: Peer Coaching and Rating Through Online Synchronous Communication for ESL Learners

Sujatha Krishnan and Mohd Ridhwan Abdullah

Abstract Peer coaching is emerging as a way to enhance the educational experience amongst learners. However, ESL learners tend to have some inhibition when speaking with peers that are of higher proficiency than themselves in face-to-face conversations. It is essential to know whether they exhibit the same characteristics when speaking to peers via synchronous-based online communication. Thus, this research explores ways learners assess their peers when they are communicating with peers of similar and higher competency in English via synchronous online speaking sessions. This study was conducted for six weeks and the participants had to complete two thirty-minute sessions twice a week with a partner of similar and higher language proficiency. The learners then assessed their partners through a questionnaire after each session. Furthermore, the participants also had to write a short reflection about their completed session. The results were analysed using SPSS. The overall findings indicate that students benefit from online synchronous communication.

Keywords Peer coaching · Peer assessment · Online synchronous communication · ESL

1 Introduction

The classroom today is a very different setting as student-centred learning takes the limelight as opposed to teacher-centred activities. Some popular student-centred activities are group discussions and pair-work-based activities. Educators try to make the classroom more interactive and interesting for the learners, thus resulting in a huge variety of learner-based activities used. As stated by Li (2016), the use of technology in the classroom enables more student-centred learning which is essential to ensure that the lessons remain relevant to the students that are digital

S. Krishnan (✉) · M.R. Abdullah
Taylor's University, Subang Jaya, Malaysia
e-mail: Sujatha.Krishnan@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_11

natives. As more learner-based tasks are created, it becomes a norm that students are seen to be able to collaborate and learn with each other without much intervention from the teacher. This has opened numerous possibilities for students to learn with their peers instead of only relying on the teacher or Internet. According to the Oxford English Dictionary, a peer is 'a person of the same age, status or ability as another specified person'. Hence, these peers are students that are on an equal plane as they strive towards achieving a common task or goal.

Secondly, looking at the current educational scenario, the classroom experience is getting more virtual nowadays. The students can actually get their tasks completed during real class time or even out of it. Therefore, another factor that can make a big difference in online-based communication is if it is a synchronous or asynchronous mode that will be utilised by the students. Synchronous mode would require an immediate answer whereas an asynchronous mode will enable a delayed response. Educators also have to gauge the pros and cons of both modes of response to decide which suits their purpose better. This study aims to look at getting students to use social network-based software like Skype and Google Hangout to communicate with their peers in a target language outside the school hours.

The research seeks to address the following research questions (RQ):

1. How do students assess their own language proficiency in an online-based communication?
2. How do students assess their peers' language proficiency in an online-based communication?
3. What are the noticeable differences between the assessment of students for themselves and their peers?

2 Literature Review

2.1 Peer Coaching

Coaching is seen as a positive way to ensure that the student that is lacking in some aspect is given the assistance required. However, there are qualms in terms of who should be providing the assistance required—teachers, parents, the Internet or peers? The old paradigm would suggest that the sources of knowledge would be the older, wiser person, but with the current paradigm shift, there seems to be a preference for peer mentoring (Parker et al. 2008).

Beaumont et al. (2012) discussed the benefits and challenges of an online-based peer learning programme. Their study looked at an existing peer coaching that was campus-based named peer-assisted study session (PASS) and compared it with an online peer coaching scheme known as online peer-assisted learning (OPAL). Their participants were first-year students enrolled in an engineering course—Engineering Systems Design 2 (ESD2) and a second-year accounting course—Intermediate

Financial Accounting (IFA) students as these courses were notably difficult and required extra help by the students. The study was conducted for one semester using the Adobe Connect platform with five students taking part in the focus group. The overall results indicated that the participants found OPAL useful and would consider using it again for their future courses. The study also highlighted the challenges that should be considered in order to make this a more successful programme by considering the willingness of the participants to use the programme which would directly impact the attendance thus the success of the programme. Also, other equally important aspects would be times for the sessions conducted, training for the coaches, and sustainability of the programme for the intended learning activities and promotions of the programme.

As stated by Hrastinski and Stenbom (2013), the student-to-student online coaching can be done at all education levels in and outside the academic setting. The key point to note is that the learning is still controlled by the student as they decide when to log in to access the system. Therefore, the learners using the system should be not only comfortable using the system but also feel that it is beneficial to their learning.

According to Kulkarni et al. (2015), there are several merits of peer feedback. The students are given the opportunity to view their work from an assessor's perspective, thus they are also more aware of the solution, insights and approaches used while conducting peer assessment. These online students are able to connect with each other regardless of their demographic background especially when viewing course materials, looking at data or being involved in discussions on the subject matter.

Thus, learners benefit when they interact with their fellow classmates as they are able to relate better with each other in terms of the difficulties and the issues faced when completing their assigned work. This forms an online interaction which is beneficial to both speakers as they are able to improve themselves with comments given to each other.

2.2 Synchronous Online Communication

Synchronous communication requires the users to interact as if they are in a face-to-face conversation. According to Hrastinski (2008), synchronous communication requires the individuals communicating to be working at 'real time'. This approach enables the users to 'explore' concepts and also helps to improve the amount of information they remember after the interaction.

However, as stated by Lai and Zhao (2006), some benefits of using text-based online chat platforms versus the online synchronous interactions are learners get to 'review' their interactions by noticing the errors done earlier from texts and the corrective feedback given by their counterpart, thus enabling them to pay more attention to the errors made and correct them. This approach also mitigates the problem of speakers having lack of time to respond, resulting in their lack of

confidence and coherence as the response period increased with the text-based platforms versus using face-to-face interaction or synchronous speaking that requires almost immediate responses. Therefore, learners are able to monitor their usage of language better.

Nevertheless, as stated by Satar and Özdener (2008), the text-based interactions lack the pragmatic information of the speaker such as intonation and gestures. This puts learners at a disadvantage as they are not aware of the speaker's intended meaning. Moreover, the text-based interactions would be more formally constructed as compared to the speaking-based synchronous sessions which would be useful for learners to especially in settings outside the classroom.

In a study conducted by Shadiev et al. (2015), synchronous communication helped to foster a better understanding of another person's culture that was totally diverse from the participant. This is essential as students not only learn to practice a language, but they also learn about another person's culture, beliefs and other aspects that may be different in everyday life.

2.3 Peer Rating

Past studies of the quality of peer rating has mainly focused on its reliability and validity. According to Burke (1969), there were some apprehensions with regard to the ability of students to rate each other fairly, but a large body of research shows a great degree of agreement between student and teacher ratings in various disciplines of higher education, as reviewed by Falchikov and Goldfinch (2000). Research on the reliability of peer rating in group work, however, is limited. In other words, most research pertaining to the quality of peer rating has been devoted to the agreement between student and teacher ratings, as opposed to the consistency of ratings amongst students themselves. As the reliability and validity of peer rating have always been questioned, Centra (1975) concluded in his study, in most cases, peer rating is not reliable, although Centra's (1975) study can be deemed irrelevant today. According to Kane and Lawler (1978) 'not only peer rating is the most useful compared to other methods for feedback purposes (peer nominations and peer rankings), it also produces the least valid, reliable, and unbiased measurements'. For this method of peer assessment, it consists of students rating other students on some type of behavioural scale. Each student is required to rate their peers by indicating his or her perceptions of how the peer being rated performed. Most methods proposed to account for individual efforts in group work rely on peer and self-rating of contribution to group work by students themselves (Johnston and Miles 2004). Although teachers may not be fully aware of the individual contribution of each student, students are well aware of how well each individual is doing.

3 Methodology

This section is divided to several parts, the first section looks at the framework used for the study and the second section looks at the specific methods which were used to carry out the study and the data collection process.

As shown in Fig. 1, Brindley's learner-centred system (1987, taken from Self/Peer-Assessment: Stress-Free Consciousness-Raising and reflection; Finch, 2011) is a versatile system that enables an entry point at any part of the system. This system can be used during the course as well as after the course because it is a cyclic system. Here, the learner is able to identify his or her needs and reflect on the methods that can be used to obtain the skills that they are lacking. Therefore, if the learner starts at the discussion level, the student can then use their peers or teacher's suggestions and negotiate the means that can help them to gain skills that they are

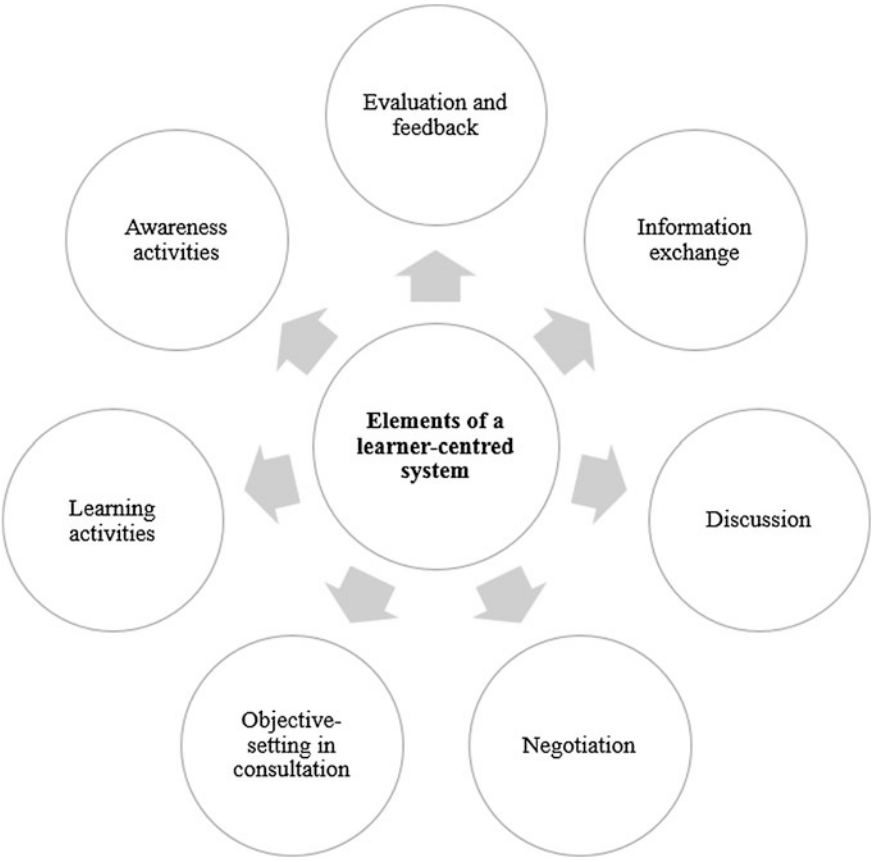


Fig. 1 Elements of a learner-centred system (Brindley 1984:77). *Source* Self/peer-assessment: stress-free consciousness-raising and reflection; Finch, 2011

lacking. They can then proceed to set the objectives they would like to achieve and then participate in learning activities with their peers or teacher that can help them obtain their objectives. Next, they can participate in awareness activities that can help them reflect in terms of their learning and achievement. This can then be followed with the evaluation and feedback from peers or teacher which can help the learner to identify whether they have accomplished the initial skill they wanted to achieve.

For this study, the researcher empowered the students to identify the skills that they would like to improve when they communicate with their peers. The students are also given the freedom to decide how they would like their session to be with their peers (either question–answer-based session or discussion-based session) as the only aspect that is controlled is the topic of discussion for each session. The students connect to each other via synchronous mode and they also can get updated content via the Internet synchronously while doing an online speaking session. However, their contact with the teacher is in asynchronous mode, as the teacher is not available for consultation when the students are interacting with each other. Nevertheless, the teacher can still monitor the sessions asynchronously via the recorded sessions done. The students also submit a questionnaire after each session is completed. They also need to submit a short reflection of how they felt the session went after each session is completed with their peers.

This study uses the mixed method as this was seen as an effective method to obtain qualitative and quantitative data. For the quantitative data, the information obtained is from questions based on the Likert scale as well as open-ended questions that are follow up questions for the responses given. This approach is similar to the study on learner's attitude towards e-mentoring conducted by Dahalan et al. (2012) which used self-administered questionnaires and the data were later generated via SPSS version 18. In terms of the qualitative data, a similar study done by Nagel et al. (2009), as the researchers used the students blog postings, discussion posts, online quizzes and also examination essays for qualitative-based data. Thus, the researcher decided to emulate this approach and obtain the student's response after each session via a Google Doc, as interviewing each student after every session did not seem executable in such a short frame of time. This document is shared only between the student and the researcher. To ensure that the document is constantly updated, the researcher gave short comments after each posting was done by the student.

Another factor that is essential in the study was that the researcher had initially done a short assessment to decide on the students' proficiency levels before the synchronous online sessions. This information was then used to categorise students either as high-English-proficiency (HEP) or intermediate-English-proficiency (IEP) speakers. The IEP students were then paired with HEP students for Weeks 1, 2, 5 and 6. In weeks 3 and 4, the IEP students were paired with a similar IEP speaker and a HEP speaker was paired with a fellow HEP speaker. This was done to identify whether there were differences in the communication patterns between speakers of different proficiency, as well as to find whether there were any substantial differences in terms of rating partners of different proficiency.

3.1 Data Analysis

There were several types of data obtained from this study to ensure a comprehensive interpretation of the response from the participants of the study. One set of data obtained is from the feedback questionnaire completed by the students after each session. The items from the questionnaire consisted of data derived from the Likert scale were for items 7–11. This set of data was analysed using SPSS (16th edition). The first set of data was from the reflections done by students after they had a session with their respective partners. The reflections are done in Google Docs and only shared between the student and the teacher.

3.2 Instruments

The questionnaire used is adapted from the Teacher Evaluation by Students (2005) published by the Character Education Partnership. The questionnaire consists of 15 items which have multiple choice, scale and open-ended answer options. The first part of the questionnaire with items one to four is to identify the total hours they spend using online and face-to-face communication with their friends. For the second part of the questionnaire, items five to 15 give more insight into the overall online speaking session and the students' feedback of their partners with regard to helping them communicate in English. The second instrument is the Google Docs which can only be accessed by the student who is in ownership of the document and the teacher. Each student will update the document after each session is completed. The teacher will also be able to read and give comments based on the feedback given by the student about the session done.

3.3 Sample

The samples for this study are based on convenience sampling, as the researcher was able to instruct the students on the necessary steps and topics to be completed for each session. Six males and five females participated in the study. These eleven students were between the ages of 18–26. Overall, there were eight international students from countries such as China, Kazakhstan, Thailand and Myanmar and three Malaysians. The students' basic proficiency levels are at IELTS Band 4.5 and above up to a maximum of Band 5.5. Prior to the commencement of their studies, these students have not had any synchronous-based online communication classes. The study was carried out in a private university in Malaysia located near the city centre. The duration of the research was for six weeks from 31 January 2015 until 10 March 2015.

Table 1 Topics assigned for the synchronous online sessions

Week	Topic	Session time (min)	Due date
1	Favourite or new food in another's culture or country	30	31.1.2015
1	Favourite movie	30	31.1.2015
2	Travel-interesting places	30	13.2.2015
2	Languages—useful languages to learn	30	13.2.2015
3	Technology	30	20.2.2015
3	Transportation	30	20.2.2015
4	Weather	30	28.2.2015
4	Sports–health-related activities	30	28.2.2015
5	Music/singers/songs	30	6.3.2015
5	Photographs-selfie/social network	30	6.3.2015
6	Being successful	30	10.3.2015
6	Future plans	30	10.3.2015

4 Assignment of Topics

The study started with the researcher assigning topics to be discussed for a duration of six weeks, and the students were required to complete two topics per week for 30 min. Table 1 for the assigned topics and deadlines.

5 Findings and Discussion

5.1 Analysis Based on Instruments

The data analysis for this study consisted of two instruments utilised which is the questionnaire and the reflection. The first set of data from the questionnaire consists of 15 items with multiple choices, scale and open-ended items to better gauge the responses given by the participant. The second set of data is the students' reflection on Google Docs which consists of two to three lines stating their opinions about each session after it has been completed.

5.2 Questionnaire

The first item required students to state the number of hours they spend in a week communicating face to face with their friends. Six students indicated that they spent less than five hours communicating face to face with friends in a week. Four

students stated that they spent between six to ten hours in face-to-face communication with friends and one student spent above 10 h.

The second item was to identify the number of hours students spent online in a week. This included time spent using a laptop, mobile phone apps (like WhatsApp and WeChat) and other devices. Five students spend between six to ten hours a week online. Four students spent below five hours online and two students spent more than ten hours online in a week. After establishing the current pattern of students' behaviour in terms of their communication preferences and practices, the next part will look at the data in terms of the research questions of the study. For the purpose of this study, the researcher got the students to utilise Skype or Google Hangout, with a huge majority opting for Skype due to the familiarity of the software.

RQ 1

How do students assess their own language proficiency in an online-based communication?

One of the items in the questionnaire administered required the students to rate their own English proficiency. Using a simple Likert scale '1' represented Very Bad, while '5' was 'Very Good'. There were a total of 43 sessions done and the mean is 3.60. This means that most students generally rated themselves between (3) Average and (4) Good. Table 2 states the total sessions based on the students ratings of their own language proficiency.

RQ 2

How do students assess their peers' language proficiency in an online-based communication?

The next item asked students to rate their peers' English language proficiency using a similar Likert scale with the same descriptors. The mean for the sessions done is 4.04 as most students rated their peers a 4 or a 5. Interestingly enough, none of the participants rated their peers a Very Bad (1) or Bad (2). The highest number of was recorded for Very Good (5) with 17 entries and Average (3) with 15 ratings. This is closely followed by 11 for the Good (4) descriptor (Table 3).

Table 2 Students' self-assessment of proficiency

Descriptor	Number of ratings
1. (Very Bad)	0
2. (Bad)	2
3. (Average)	22
4. (Good)	10
5. (Very Good)	9

Table 3 Students' peer assessment of proficiency

Descriptor	Number of ratings
1 (Very Bad)	0
2 (Bad)	0
3 (Average)	15
4 (Good)	11
5 (Very Good)	17

RQ 3

What are the noticeable differences between the assessment of students for themselves and their peers?

In total, there were 43 sessions conducted throughout the six weeks study. As stated earlier, the researcher identified the students' level of proficiency before the research commenced, and used the information to pair the students with their partners accordingly. In weeks three and four, the students were paired with partners similar in proficiency with themselves. As stated in Table 4, 10 sessions were paired correctly as the students rated themselves and their partner the same while three sessions had students rate their partners higher even though their partners had the same proficiency level with themselves.

In weeks 1, 2, 5 and 6, the researcher paired a high-English-proficiency (HEP) speaker with a low-English-proficiency (LEP) speaker for the online speaking sessions. As shown in Table 5, the HEP students rated their LEP partners higher than themselves for four sessions which is inaccurate, as they were supposed to be the better speakers. They also rated their partners as having a similar level proficiency in 12 sessions out of the 16 sessions done. Interestingly, none of them rated their LEP partners as lower than themselves.

As shown in Table 6, the lower-English-proficiency (LEP) students had 14 sessions with their high-English-proficiency (HEP) partners. They rated their HEP partner correctly for 5 sessions and the other nine sessions were rated as having equal proficiency with themselves. None of the LEP students rated themselves lower than their HEP partners although this would have been the expected outcome.

Table 4 Students' self- and peer assessment during similar proficiency sessions in weeks 3 and 4

Descriptor	Number of ratings
Evaluated partner higher	3
Evaluated partner same	10
Evaluated partner lower	0
Total	13

Table 5 HEP students' self- and peer assessment during different proficiency sessions in weeks 1, 2, 5 and 6

Descriptor	Number of ratings
HEP Evaluated partner higher	4
HEP Evaluated partner same	12
HEP Evaluated partner lower	0
Total	16

Table 6 LEP students' self- and peer assessment during different proficiency sessions in weeks 1, 2, 5 and 6

Descriptor	Number of ratings
LEP Evaluated partner higher	5
LEP Evaluated partner same	9
LEP Evaluated partner lower	0
Total	14

In conclusion, it is clear that students were not judgemental towards their partner's proficiency in English as they did not rate their partners lower than themselves. Either as a HEP or LEP speaker, the students did not gauge their partners' proficiency levels as lower than themselves. This shows that through synchronous online communication, the students are less concerned about proficiency levels and more interested in the topic of conversation and making the session fun and interesting. To get a better insight in terms of their view of the sessions done, the next aspect that will be discussed is the reflection done by the students in Google Doc after each session.

5.3 Reflection from Google Docs

The second set of data that was collected was the personalised reflection by the participants after each session. Below are some of the responses given by the participants.

Response 1

> I think this is better than speaking with coach because we already knew each other. So, I felt its more free when talking. It also give me more information because i can know something about her country.
>
> (Female, Malaysian, HEP)

Response 2

> Johnny told me a lot of his country different news and information he wants to visit China. Also he is a good player. We talked some sciences and public transportation from country of each other. Also we change our opinions of girls who are like what kind of music and social network.
>
> (Male, Chinese, HEP)

Response 3

> I think this time we talk a lot more than last time, Sulu has prepared some questions to ask me, and I am also trying my best to think of something to reply to her. About travelling, we shared the places we want to go in the future and we both agree that travelling is fun and relaxing at the same time. Sulu seemed to be interested in China and Chinese culture, so I make some suggestions about going to China. About language, I was surprised to find that Sulu wanted to learn Chinese! So I offered to be her teacher when she is studying Chinese, but we have to admit that English is still the most useful language all around the world. you can find it useful while you are travelling and doing business. It may also save your life! It has been great fun talking to her and I have learned a lot. I think talking to the classmates surely is a great way to improve our speaking skills and share our information at the same time!
>
> (Female, Chinese, HEP)

Response 4

> I was happy talked with Tina again:-) When I talked with Tina via Skype, I felt calm, because I did not hesitate and It was interesting conversation.
> (Female, Kazakh, LEP)

Response 5:

> I talked with June Ko about future plans and how to be successful in the future. Our plans are same in some points and he knows precisely how to be successful person.
> (Male, Kazakh, LEP)

Response 6

> 'We have done 1 h of 2 topics, food and movie, I felt very relax to talk to my classmate and it will improve my speaking skills'.
> (Male, Malaysian, LEP)

In conclusion, the data from the responses show that students are comfortable to speak with their peers regardless of their dissimilar level of proficiency. These students try to maintain a good conversation and the lack of language skills do not inhibit them as much as it would if they were to converse face to face. Many students have also reached a comfort level which enabled them to trust that their peer will not be judging them for the grammatical errors they may be doing while the conversation is in progress. The findings of this study are similar to the findings of the studies by Satar and Özdener (2008), Jones and Gallen (2015), which stated that respondents found the online-based interactions in English less inhibiting because they were already familiar with their partners and have interacted face to face with each other before commencing on the online-based sessions.

6 Conclusion

Overall, this research was conducted to identify whether students were able to identify the skills they required to improve their English language proficiency, along with rating themselves and their peers accurately according to their notions of having a useful and successful online conversation. It is clear that students were more focused on the process and did everything they could to ensure that they peer had a good session and they were good partners for their peers.

6.1 Insufficient Number of Participants

Since the total number of participants in this study is only eleven students, it entails that the study is not able to generalise the findings discussed. More participants and

a longer duration are required in the next study to enable the research findings to be more substantial.

Data from participants

Also, the students assessing themselves and their sessions may not be totally reliable as they may state a session as fun because they discussed some items outside of the assigned topic. Therefore, to conclude it may be stated that students felt that they benefitted from the session due to the fact that they had to converse in English as their partner did not share the same mother tongue.

6.2 Implications of the Study

There are several implications from this research. The two main implications are that students' English language proficiency improved and the students better understood another culture foreign to their own.

English Language Proficiency

One of the main implications of this study is that students were able to practice speaking English outside the classroom setting without inhibition because they were conversing with their peers. These students normally speak English only in the classroom and occasionally around the campus for daily activities such as buying food, borrowing a book and other basic communication purposes. The extra time they spent is not only during the actual conversation itself as the students had to prepare some questions or read about the topic before speaking with their friends, thus increasing the amount of time they used English.

Understanding a Different Culture

Another interesting implication is that students actually found out more about another culture which is different from their own. Normally in class, students tend to interact and work with someone from the same culture or background as opposed to looking for someone with a diverse culture compared to their own. Here, the students were assigned a partner with a different culture; thus, they were given a chance to work with someone who normally they would not choose and with very diverse culture from their own, enabling them to be more aware of the similarities and differences they have in their respective cultures.

6.3 Recommendations for Future Research

The number of participants in this study was too small to be able to generalise, so a bigger sample is required for the next study. Also, there was not enough time to conduct individual interviews, and having these data will also help the researcher to

gain more insight about the students' opinions in terms of peer coaching in online-based synchronous communication.

References

Burke, R. (1969). Self-assessments and peer ratings. *Journal of Educational Research, 61*, 444–448.

Centra, J. A. (1975). Colleagues as raters of classroom instruction. *Journal of Higher Education, 46*(1), 321–337.

Dahalan, N., Hassan, H., & Atan, H. (2012). Student engagement in online learning: Learners attitude toward E-Mentoring. *Procedia-Social and Behavioral Sciences, 67*, 464–475.

Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Review of Educational Research, 70*, 287–322.

Hrastinski, S. (2008). Asynchronous and synchronous e-learning. *Educause Quarterly, 31*(4), 51–55.

Hrastinski, S., & Stenbom, S. (2013). Student–student online coaching: Conceptualizing an emerging learning activity. *The Internet and higher education, 16*, 66–69.

Johnston, L., & Miles, L. (2004). Assessing contributions to group assignments. *Assessment & Evaluation in Higher Education, 29*, 751–768.

Jones, M. H., & Gallen, A. M. (2015). *Peer observation, feedback and reflection for development of practice in synchronous online teaching*. Innovations in Education and Teaching International, (ahead-of-print), 1–11.

Kane, J. S., & Lawler, E. E. (1978). Methods of peer assessment. *Psychological Bulletin, 85*(3), 555–586.

Kulkarni, C., Wei, K. P., Le, H., Chia, D., Papadopoulos, K., Cheng, J., & Klemmer, S. R. (2015). Peer and self assessment in massive online classes. In *Design Thinking Research* (pp. 131–168). Springer International Publishing.

Li, Y. W. (2016). Transforming conventional teaching classroom to learner-centred teaching classroom using multimedia-mediated learning module. *International Journal of Information and Education Technology, 6*(2), 105.

Lai, C., & Zhao, Y. (2006). Noticing and text-based chat. *Language Learning & Technology, 10*(3), 102–120.

Nagel, L., Blignaut, A. S., & Cronjé, J. C. (2009). Read-only participants: A case for student communication in online classes. *Interactive Learning Environments, 17*(1), 37–51.

Satar, H., & Özdener, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *The Modern Language Journal, 92*(4), 595–613.

Shadiev, R., Hwang, W. Y., & Huang, Y. M. (2015). A pilot study: Facilitating cross-cultural understanding with project-based collaborative learning in an online environment. *Australasian Journal of Educational Technology*.

Teacher Evaluation by Students. (2005) National Schools of Character: Award Winning Practices, Published and distributed by the Character Education Partnership, http://www.rucharacter.org/file/Microsoft%20Word%20%20Teacher%20Evaluation%20by%20Students%281%29.pdf. Accessed on 12 Feb 2015

Beaumont, T. J., Mannion, A. P., & Shen, B. O. (2012). From the campus to the cloud: the online peer assisted learning scheme. *Journal of Peer Learning, 5*(1), 1–15.

Part IV
Assessing Professional Competencies and Institutional Initiatives

Assessment and Feedback in the Final-Year Engineering Project

Douglas Tong Kum Tien and Siow Chun Lim

Abstract The final-year project (FYP) is considered to be one of the most important learning experiences in an engineering undergraduate education. It is mandated by engineering accreditation bodies worldwide as a compulsory module in the engineering curriculum. For most fresh graduate engineers, their final-year project experience will be the most vivid of all their recent memories when they recall their engineering education experience. This is due to the immense challenge the students face by undertaking for the very first time an individual, highly complex, open-ended research work with tight deadlines to meet and high expectations to attain on their own, albeit under the guidance of their supervisors. As such the FYPs' delivery and assessment process offer strategic opportunities for engineering educators to add value to an undergraduate engineer's training provided that it is designed and implemented with a clear intent. This paper documents the assessment approach adopted by School of Engineering, Taylor's University, for the FYP module which is divided into the FYP1 and FYP2 modules covering two semesters. This assessment process is distinct from the practices at many other equivalent institutions in that it has as much as 6 assessment components for each module, assessors other than the supervisor contribute the bulk of the marks, timely feedback from assessors to students is fully integrated into the assessment process to facilitate learning and students are required to participate in a research conference which is also made into one of the assessment components. For this undergraduate conference, the students' FYP research papers are reviewed by external reviewers from other universities and assessed and awarded marks by them. A detailed description of each assessment component followed by an evaluation of how each assessment component adds value to the students' learning for both the FYP1 and FYP2 modules is discussed.

D.T.K. Tien (✉) · S.C. Lim
School of Engineering, Taylor's University, Subang Jaya, Malaysia
e-mail: DouglasKumTien.Tong@taylors.edu.my

S.C. Lim
e-mail: ChunLim.Siow@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_12

Keywords Final-year engineering project · Engineering education · Learning experience · Assessment for learning

1 Introduction

Assessment can be done in many different forms for different purposes. However, it has to be done in such a manner that allows the assessors to assess the students as to what and to what extent they managed to learn and also to allow students to demonstrate their learning to a certain extent (Brown and Smith 1997). Of late, learner-centered curriculum is rapidly replacing traditional teacher-centered curriculum. A learner-centered assessment must then be in place because curriculum and assessment are mutually complementary (Manitoba Education 2006). As higher education becomes increasingly accessible, diversity of learners is inevitable. This coupled with the paradigm shift to student-centered learning and further reinforced by the practice of outcome-based education leads to the emphasis on the attainment of learning outcomes.

According to the Boards of Engineering Malaysia, completion of final-year engineering project is one of the prerequisites for a student to be graduated as an engineer (Council 2012). This is a common practice across the globe (Ku and Goh 2010). The School of Engineering, Taylor's University, implements project-based learning as project work is considered pivotal in an engineer's training (Blicblau 2004; Al-Atabi et al. 2013). At Taylor's, students undertake project-based modules from their first semester until graduation. The distinguishing factor of the FYP from the other project-based modules is that the students are required to complete a research project individually for the first time. A systematic method to design the curriculum of FYP has been proposed by Fraile et al. (2010).

Assessment for project-based modules is conducted differently from examination-based modules. These assessments use purposefully designed rubrics to assess learning and tend to do so in a more holistic manner. Projects are intended not just to develop students' cognitive capabilities but also their affective faculties and in some cases even their psychomotor skills (Tong et al. 2015). In view of this, assessments for project-based modules offer a far greater range of options to affect students' learning and holistic development compared to examinations. Consequently, educators who recognize these possibilities will realize a variety of strategic opportunities to add value to their students' learning experience through well-designed project assessments centered on clearly intended learning outcomes.

Students who have never previously experienced an individual, highly complex, open-ended research work with tight deadlines to meet and high expectations to attain cannot help but be profoundly affected by their experience. Due to the nature of the module, elements of transformative learning tend to be present in an FYP. Transformative learning is defined as the process by which problematic frames of reference are transformed to make them more inclusive, discriminating, open, reflective and emotionally able to change (Mezirow 2009). The transformed frame of

reference is more likely to generate beliefs and opinions about one self that is closer to the truth or more justified to guide action. Such an experience if it was accepted positively will help create in the students their emerging identity as a competent engineer and researcher. Their frames of reference with regard to themselves and their capabilities will transform as they grow to successfully meet these challenges.

In this paper, the assessment approaches for FYP which span two semesters would be analyzed. There are 6 assessments each in FYP1 and FYP2. Timely feedbacks are furnished to the students after each assessment. Through this series of assessments, educators, be they assessors, project supervisors, or even the module coordinator are able to regularly gauge the progress of students and whether the students are attaining the learning outcomes. As the feedbacks are descriptive, students benefit from it because they know where and what to improve on a continual basis. Finally, discussion on how the assessments would contribute added values to the students' learning experience would be outlined.

2 Assessments for FYP1

The following sections detail the assessment components of FYP1 which are summarized in Table 1.

2.1 Meeting Records

Meeting Records are required for both FYP1 and FYP2. Meeting Records serve the purpose of assessment for learning as they are actually learning logs. Its purpose is to motivate students to meet their supervisors and elicit regular feedbacks. By so doing, it is also hoped that the final-year students would not feel isolated due to a lack of peer supports that they have been accustomed to in their prior years of their degree (Armstrong 1997; Popov 2003). Some institutions name this assessment as progress meeting. The meeting between supervisors and students is vital especially at the early phase of research as students need the motivation to kick-start their progress as well as guidance on integrating theory and practice (Bouki 2007; Rowley and Slack 2004). The weekly meetings also serve the purpose of instilling accountability and help sustain the motivation to maintain a constancy of effort.

Table 1 Assessment components for FYP1

Formal assessments	Learning domain	Percentage (%)
Meeting Records	Cognitive, affective	15
Initial Proposal	Cognitive	15
Abstract	Cognitive	5
Interim Report	Cognitive	45
Oral Presentation	Cognitive, affective	10
Engineering fair	Cognitive, affective	10

A serendipitous benefit that may result is that the students receive substantial mentoring and develop a level of personal relationship with an academic staff, their supervisor, to an extent that would not be possible in other types of settings.

Students are assessed based on the frequency of meeting as well as the percentage of task completion agreed upon at each meeting. At present, all meetings are conducted face-to-face. However, it may be worthwhile to also explore blending face-to-face supervision with electronic logs. Some engineering education researchers have experimented with electronic logs and the outcome seems to be encouraging (Manish et al. 2009; Malik 2008).

2.2 Initial Proposal

The purpose of the Initial Proposal is meant to expose FYP1 students to the art and science of crafting a research proposal. Although students are not required to apply for research grants, the main motivation behind this exercise is to educate them on the key elements required in preparing a research proposal. This would be useful should they apply for research grants in their postgraduate studies and beyond. Emphasis is placed on ensuring that students are clear with their research objectives, how to achieve the objectives, project timeline, and budgets required. As aforementioned, this is the first time students are undertaking a project individually. Therefore, this assessment would challenge them to plan for a one-year research project. Students are required to submit their Initial Proposal in the middle of the FYP1 semester.

This assessment component also serves as a driving force to transfer ownership of the project from the supervisor to the student. The main differentiator here is that this assessment is being assessed by two other lecturers in addition to an indirect assessment by the supervisor. In several universities in Australia, this assessment component is solely assessed by the project supervisor (Ku and Goh 2010). However, learning experience of students can be enhanced by eliciting feedbacks from different perspectives. It would also lead to a more impartial and more reliable assessment by having other lecturers to be the assessors.

A Gantt chart is required in the proposal. The main purpose is to encourage students to monitor the progress of their projects. It also equips students with time management skills so valuable in the working world. Through a Gantt chart, they are able to organize the time required to complete each specific set of tasks.

2.3 Abstract

In Abstract writing, students are assessed for the first time in their degree on how to summarize their research in a concise manner. The importance of an Abstract to

attract readership is highlighted. Abstract writing is challenging because it is widely acknowledged that writing a long story short is more difficult than writing a short story long. Such skill could also help students to prepare their future curriculum vitae in that it instills the skill of identifying what is important and then to convey that in concise manner. The Abstract is assessed by the same two lecturers who assessed the Initial Proposal and written feedback is provided.

2.4 Interim Report

Interim Report is the final written submission at expected at the end of FYP1. At this stage, the literature review is expected to be thorough. The research methodology is expected to be finalized leaving only the experimental or simulation works to be done in the following semester. It helps students to crystalize their emerging understanding of their research effort through a written report assessed by the same two lecturers who assessed their Abstract and Initial Proposal. The Interim Report is intended to serve as an effective midpoint check providing an opportunity for midpoint feedback before the students embark on FYP2. This assessment is conducted by the same two lecturers who assessed the Initial Proposal and written feedback is provided for this assessment as well.

2.5 Oral Presentation

This assessment differs from other assessments as aforementioned for FYP1 in that the feedback is immediate. Students present on what they have achieved in their project thus far. Then, the assessors elicit clarification and also give suggestions as to how to improve or refine their works. Thee assessors would have already been given a printed copy of the Interim Report one week before the Oral Presentation. By so doing, the assessors are better prepared to ask questions and give recommendations. Hence, the quality and relevance of feedback can also be enhanced. Again, the assessment would be guided by a rubric and apart from the verbal feedback during the presentation, the students would also receive a copy of the written feedback.

2.6 Engineering Fair

The students are required to participate in the end of semester Engineering Fair. This event showcases all the project-based modules achievements by students from semester 1–8. In this event, the FYP students need to conduct a poster presentation to two judges. One judge would be an academic staff of Taylor's School of Engineering while the other would be an invited judge from the industry with

significant engineering experience. Through this assessment, the students would further hone their presentation skills, learn to condense information into the limits of a poster, interact with an industry judge, and receive their advice and feedback. The students are expected to remain at their booths throughout the entire event and be ready to present their work not only to the judges but also to any interested person regardless whether they are another student, a lecturer, or a visitor. The Engineering Fair is the final assessment component of the semester and this highly visible event concludes the assessment cycle for the module. In addition, there is also a competitive element involved in that prizes are awarded for the best projects. For many students, the prize award is the highlight of the Engineering Fair. Both FYP1 and FYP2 students will participate in the Engineering Fair.

3 Assessment Components for FYP2

The following sections detail the assessment components of FYP2 which are summarized in Table 2.

3.1 Meeting Records and Engineering Fair Poster

These assessment components are similar to those in FYP1, and they serve the same purpose. However, the Engineering Fair Poster Presentation would be more meaningful for FYP2 students because they have now either a complete artifact or findings to present.

3.2 EURECA Conference Paper and Presentation

The main aim of requiring the students to prepare a manuscript capturing their research work and then presenting it at a conference is to give them early exposure

Table 2 Assessment components for FYP12

Formal assessments	Learning domain	Percentage (%)
Meeting Records	Cognitive, affective	15
Eureca Conference Paper	Cognitive	15
Oral Defense of Thesis	Cognitive	5
Thesis	Cognitive	45
Eureca Conference presentation	Cognitive, affective	10
Engineering fair poster	Cognitive, affective	10

to a conference setting usually considered as the privilege for postgraduate students. By including this as an assessment, it is expected enhance learning experience. It is worth mentioning that a select few of the best-quality papers would be published in a Scopus journal and achieving such feat would definitely add value to the students' confidence as well as academic qualification. Publishing a journal article at an undergraduate level could be the catalyst to prompt some students to further their studies to postgraduate level.

Students first submit their manuscripts which will then be forwarded to external reviewers from locally or abroad for rigorous assessment. By having this conference, students have the opportunity for their work to be reviewed and independently benchmarked against that of other universities. It sends the message to the students that their work is more than a regular FYP. Their work is actually seen as a worthwhile contribution to the body of knowledge and its standard is deemed to have met the rigors of academic research by external reviewers.

In the EURECA Conference, there are various categories of awards for best projects. There were six awards in the most recent EURECA with an additional best overall project award chosen from among the winners of these six awards. This overall winner received a fully sponsored short-term attachment at the National University of Singapore as part of the prize. Having the opportunity to do research placement at a foreign and more established university would further enrich the student's learning experience.

3.3 Oral Defense

The Oral Defense of FYP2 differs from Oral Presentation of FYP1 in that the students are not expected to prepare any presentation slides. They are only required to bring a copy of their Thesis and defend their work when asked specific questions by the assessors. The assessors, two lecturers from the school whose identity is unknown to the students up to that point, are given the student's Thesis one week ahead of the Oral Defense. An Oral Defense is a unique experience of any engineering undergraduate. However, the practice of not encouraging students to present from slides followed by questioning later is rare. The ultimate aim of this practice is to have the students take full ownership and responsibility over the authenticity of their work.

3.4 Thesis

The Thesis is the final piece of document to be submitted. A Thesis encapsulates all the works that students have undertaken in their research. This is the most significant piece of work the students are required to produce over the entire two semesters of FYP. As the students have already received feedback from the external

and internal reviewers for all of their preceding submissions, they are therefore expected to considerably improve the quality of their Thesis. Having feedbacks from the supervisor, internal and external assessors contribute significantly to their learning and development. Autonomy, authenticity, and ownership are what are expected to be valued by students upon the completion of the final-year Thesis.

In summary, throughout the entire two semesters of FYP, 12 assessments consisting of 10 different types are conducted. Their details are summarized in Table 3.

Table 3 Summary of FYP assessment components

Assessment	Purpose	Module learning outcomes	Remarks
Meeting record[c]	Accountability, self-discipline, receiving ongoing guidance and feedback, advice	FYP1: Build a research plan using project management tools FYP2: Apply project management tools to execute the research plan	Supervisor guides student and award marks based on the progress of work and frequency of meetings
Engineering fair[c]	Develop poster presentation skills, ability to condense info	FYP1: Compile the findings in both written and verbal form FYP2: Compile and present a final-year project Thesis and a peer-reviewed conference paper	An external assessor from industry and an internal assessor evaluate student's work and poster presentation and award marks Based on marks awarded, there are various categories of prizes awarded to students with the higher marks
Abstract[a]	Simulate conference participation First assessment component submitted together with the Initial Proposal during mid-semester of FYP1	Design and prepare research methodology	Two internal assessors evaluate award marks and provide timely written feedback
Initial Proposal[a]	First assessment component submitted together with the Abstract Forces students to engage in initial study Develops sense of ownership by having students submit a proposal to conduct their research	Formulate the scope and objectives of a particular research project Organize critical literature review Build a research plan using project management tools Design and prepare research methodology	Two internal assessors evaluate award marks and provide timely written feedback The feedback is channeled through the supervisor to guide/improve the student's work

(continued)

Table 3 (continued)

Assessment	Purpose	Module learning outcomes	Remarks
Interim Report[a]	Mid project summary and evaluation Feedback from other experts than supervisor	Formulate the scope and objectives of a particular research project Organize critical literature review Design and prepare research methodology Compile the findings in both written and verbal form	Two internal assessors evaluate, award marks and provide timely written feedback The feedback is channeled through the supervisor to guide/improve the student's work in the upcoming FYP2
Oral Presentation[a]	Presentation skills Feedback from other experts than supervisor	Compile the findings in both written and verbal form	Two internal assessors evaluate, award marks and provide timely written feedback The feedback is used to help students improve on their oral and PowerPoint Presentation skills
EURECA paper[b]	Experience real conference participation Benchmark paper against other universities through peer review Highlights the importance and contribution of the students work as it is deemed worthy of a peer-reviewed conference	Evaluate results using research-based knowledge and research methods which include experiment design, data analysis and the synthesis of information to provide conclusions Compile and present a final-year project Thesis and a peer-reviewed conference paper	An external reviewer cum assessor from a different institution nominated by the student's supervisor reviews and comments on the paper as conference reviewer He/she also award marks and gives written feedback as an assessor
EURECA Conference presentation[b]	Experience real conference participation Benchmark paper against other universities through peer review Highlights the importance and contribution of the students work as it is deemed worthy of a peer-reviewed conference	Evaluate results using research-based knowledge and research methods which include experiment design, data analysis and the synthesis of information to provide conclusions Compile and present a final-year project Thesis and a peer-reviewed conference paper	An external assessor from either academia or industry and an internal assessor evaluate the student's presentation In addition to marks awarded, there are different categories of prizes and the opportunity to publish in a SCOPUS Indexed Journal for the best papers

(continued)

Table 3 (continued)

Assessment	Purpose	Module learning outcomes	Remarks
Thesis[b]	Documentation, evaluation and communication of an entire 2 semesters' work through a Thesis This is the most significant piece of work produced by the students for assessment	Evaluate results using research-based knowledge and research methods which include experiment design, data analysis and the synthesis of information to provide conclusions Compile and present a final-year project Thesis and a peer-reviewed conference paper	Two internal assessors evaluate, award marks and provide timely written feedback The feedback is channeled through the supervisor. Students revise their Thesis based the feedback and submit a hardbound Thesis. The hardbound Thesis is not assesses as it is submitted after final exams
Oral Defense[b]	Experience unstructured presentation that requires impromptu responses The ability to defend one's work	Evaluate results using research-based knowledge and research methods which include experiment design, data analysis and the synthesis of information to provide conclusions Compile and present a final-year project Thesis and a peer-reviewed conference paper	Two internal assessors evaluate, award marks and provide timely written feedback The Oral Defense also helps the assessors to be able to assess student's work better. The Thesis is given to assessors a week before Oral Defense

[a]FYP1
[b]FYP2
[c]Both FYP1 and FYP2

4 Student Learning Experience

An initiative which is thought to be able to enrich students' learning experience is the dissemination of the each assessment rubric to students much earlier than the submission deadlines. By definition, a rubric entails the expectations for an assignment by articulating the levels of quality and describing the characteristic of each level of quality (Heidi et al. 2009). Utilization of rubrics allows educators to furnish focused feedback on continuous works on a regular basis. This catalyzes the achievement of the purpose of assessment for learning (Barbara 2003; Andrade 2000). Past researchers have concluded that rubrics are advantageous in terms of boosting students' self-efficacy (Quinlan 2006; Stiggins 2001).

It is the module coordinator's responsibility to ensure that the students would receive their feedback within two weeks after each assessment submission. The FYP module coordinator also helps to ensure that comments from the assessors are descriptive and clear for students to know what needs to be improved. The

rubrics are designed such that the assessment criteria are clearly furnished. Although the students are not permitted to know their actual marks, their supervisors work closely with them and advise them on how to incorporate assessors' feedback for further improvement.

Shafie et al. (2008) found that completing FYP has improved students' learning experience in terms of:

1. Time management,
2. Learning ability,
3. Problem-solving skill, and
4. Personal growth.

The findings above agree with the intended outcomes of the FYP assessments described in this paper. Time management was assessed in the Initial Proposal and Meeting Records through construction of Gantt chart and planning the weekly tasks. Students' learning ability and problem-solving skill are captured in all assessments. In terms of personal growth, the EURECA and Engineering Fair serve as unique platforms for the development of the said learning experience.

From another perspective, the FYP can also be viewed as an engineering communication course because written and oral communication is continuously emphasized through the various assessments. The Oral Assessments, which include Oral Presentation, Oral Defense, EURECA Conference Presentation, and Engineering Fair Poster Presentation, establish a dialogic environment (Artemeva et al. 1999) which allows students to gain meaning of what they did and hence elevate their learning experience through verbal negotiation with the assessors.

5 Conclusions

A description and evaluation of how the final-year engineering project module at School of Engineering, Taylor's University, had served the purpose of assessment for learning was presented. These assessments were designed in a way that allows the module coordinator, the project supervisors, and the students to plan and advance their learning. Students' progress with respect to the intended learning outcomes is able to be closely monitored through the current assessment structure. The reliability of these assessments is upheld by having many different forms of assessment conducted by both internal and external assessors, as well as by the supervisor. The continual feedback integrated into each and every assessment component makes this FYP assessment process a means of assessment for learning as the assessment itself is integrated with and promotes the students' learning. Finally, it can be said that the entire FYP experience challenges each student to develop holistically. These assessments as described in this paper are centered on clearly intended outcomes and when they are taken together, complement each other and serve to deliver a truly transformative learning experience for the emerging engineer.

References

Al-Atabi, M., Edwin, C., Satesh, N., Abdulkareem, S. A.-O., & Marwan, M. (2013). A blueprint for an integrated project based learning framework in engineering education: A case study at Taylor's university. *Journal of Engineering Science and Technology*, 8–18.

Andrade, H. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 13–18.

Armstrong, S. (1997). Dissertation supervision: Managing the student experience. In S. Brown (Ed.), *Facing up to radical changes in universities and colleges*. London: Kogan Page.

Artemeva, N., Logie, S., & St-Martin, J. (1999). From page to stage: How theories of genre and situated learning help introduce engineering students to discipline-specific communcation. *Technical Communication Quarterly*, 301–316.

Barbara, M. M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. *Practical Assessment, Research and Evaluation*.

Blicblau, A. (2004). Promotion of final year capstone projects. In *American society for engineering education annual conference and exposition.* Washington, DC: American Society for Engineering Education.

Bouki, V. (2007). Undergraduate computer science projects in UK: What is the point? In *Proceedings of the informatics education Europe II conference* (pp. 176–184). Thessaloniki: South-East European Research Center.

Brown, S., & Smith, B. (1997). *Getting to grips with assessment.* SEDA Publications.

Council, E. A. (2012). *Engineering programme accreditation manual*. Kuala Lumpur: Board of Engineers Malaysia.

Fraile, R., Arguelles, I., Juan, G. C., Gutierrez-Arriola, J. M., Godino-Llorente, J. I., Benavente, C., et al. (2010). A systematic approach to the pedagogic design of final year projects: Learning outcomes, supervision and assessment. *International Journal of Engineering Education*, 997–1007.

Heidi, A. L., Xiaolei, W., Ying, D., & Robin, A. L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *The Journal of Educational Research*, 287–301.

Ku, H., & Goh, S. (2010). Final year engineering projects in Australia and Europe. *European Journal of Engineering Education*, 161–173.

Malik, M. (2008). Work in progress: Use of social software for final year project supervision at a campus based university. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications.* Vienna.

Manish, M., Khusainov, R., Zhou, S., & Adamos, V. (2009). A two year case study: Technology assisted project supervision (TAPaS). *Engineering Education*, 76–83.

Manitoba Education. (2006). *Rethinking classroom assessment with purpose in mind.* Manitoba: Manitoba Education, Citizen and Youth.

Mezirow, J. (2009). An overview of transformative learning. In K. Illersis (Ed.), *Contemporary theories of learning* (pp. 90–105). New York: Routledge.

Popov, A. (2003). Final undergraduate project in engineering: towards more efficient and effective tutorials. *European Journal of Engineering Education*, 17–26.

Quinlan, A. (2006). *Assessment made easy: Scoring rubrics for teachers from K-college*. Lanham: Rowman and Littlefield Education.

Rowley, J., & Slack, F. (2004). What is the future for undergraduate dissertations? *Education and Training*, 176–181.

Shafie, A., Janier, J., & Herdiana, R. (2008). Students' learning experience on the final year project for electrical and electronics engineering of Universiti Teknologi PETRONAS. In *International conference on engineering education*, (pp. 1–5). Budapest.

Stiggins, R. (2001). *Student-involved classroom assessment.* Upper Saddle River, New Jersey.

Tong, D. K. T., Gamboa, R. A., & Namasivayam, S. N. (2015). Final year engineering project and its contribution to holistic education. In S. F. Tang & L. Logonnathan (Eds.), *Taylor's 7th teaching and learning conference proceedings* (pp. 575–591). Singapore: Springer Science +Business Media.

Tools for Research Management

Jer Lang Hong

Abstract Recent advancement in E-Learning tools has enabled learning techniques to be measured quantitatively and qualitatively. With the state-of-the-art tools, teaching engagement and student's learning can be gauged where improvements can be made through various feedbacks obtained from the tools. However, few works have been carried out to measure student's learning for research-related works. Though there exist certain standards and measurements in judging a research quality, most of these works are highly subjective and is subject to argument. In this paper, we propose a novel technique to effectively measure a student ability to conduct research effectively. Several benchmarks and standards are used to evaluate the learning curve, such as the ISI, Scopus, and Citations score. In addition to that, we measure the student's writing skills, reviewer feedback, their ability to conduct a proper literature review, as well as their analytical skills using various qualitative methods. Then, we combine the student's skills where a summary of skills can be obtained so that we can further evaluate and plan the student progress as well as improving the university research initiatives and policies.

Keywords Web 2.0 · Research methodology · E-Learning

1 Introduction

Recent advancement in E-Learning tools has made it possible for us to gauge the effectiveness of learning methods. From the traditional pen and pencil to exam paper-based questions, researchers (e.g., student, academician, industry researcher) have embarked upon IT technology whereby online tools are used to ease the user's learning experience. However, very few studies have been made to gauge student's learning skills in doing research. If we can effectively measure the student

J.L. Hong (✉)
Taylor's University, Subang Jaya, Malaysia
e-mail: JerLang.Hong@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_13

performance in research, it is certainly useful for us to monitor and formulate improvement plan so that students can effectively cope with their study life.

Most people have acknowledge that doing high-impact research is a non-trivial task, due to the complexity of work involved and the many challenges that lie ahead. Fortunately, most research, if not all, do have certain methodology whereby we can measure them qualitatively and quantitatively. Typically, research methodology involves finding a research area, conducting literature review, formulate the research problem, finding motivation, proposed solution, and experimenting on the given solution. Though one is not able to promise a working solution right ahead, but one can effectively monitor his task and formulate an improvement plan if a tool is developed (Fig. 1).

In this paper, we developed an E-Learning tool to effectively gauge a researcher learning experience in conducting a research. Several management tools have been developed, and some of the notable examples are content management system (Fig. 2), environmental management system (Figs. 3, 6 and 7), welfare management system (Fig. 4), and quality management system (Fig. 5). Various factors are taken into consideration for the measurement, both quantitatively and qualitatively. Some of the well-known measures which we can use as benchmarking criteria are the researcher's profile score, such as the ISI, Scopus, and Citations score. Though we do not guarantee a highly accurate solution to monitor researcher performances,

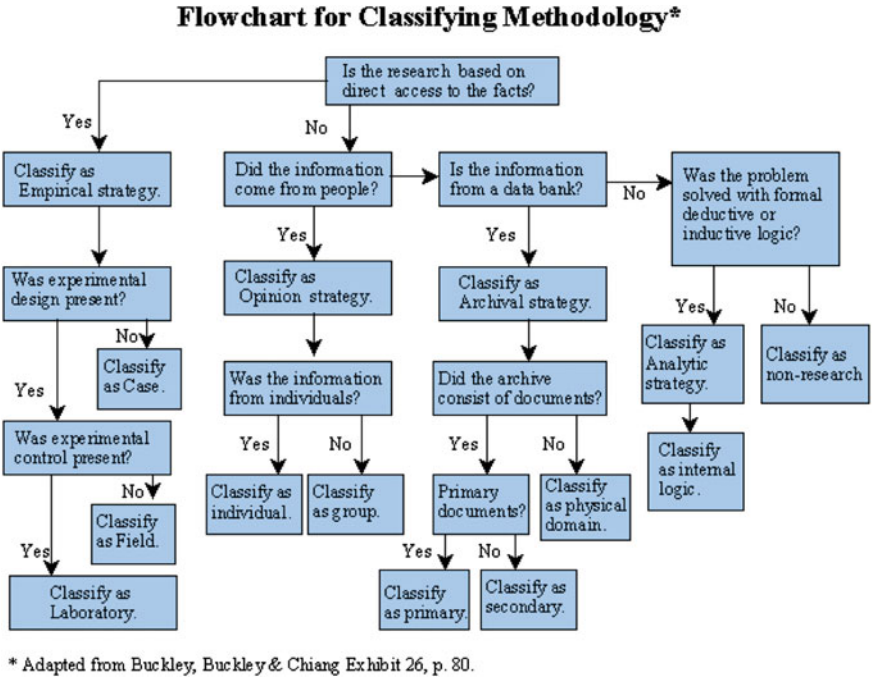


Fig. 1 Research methodology example

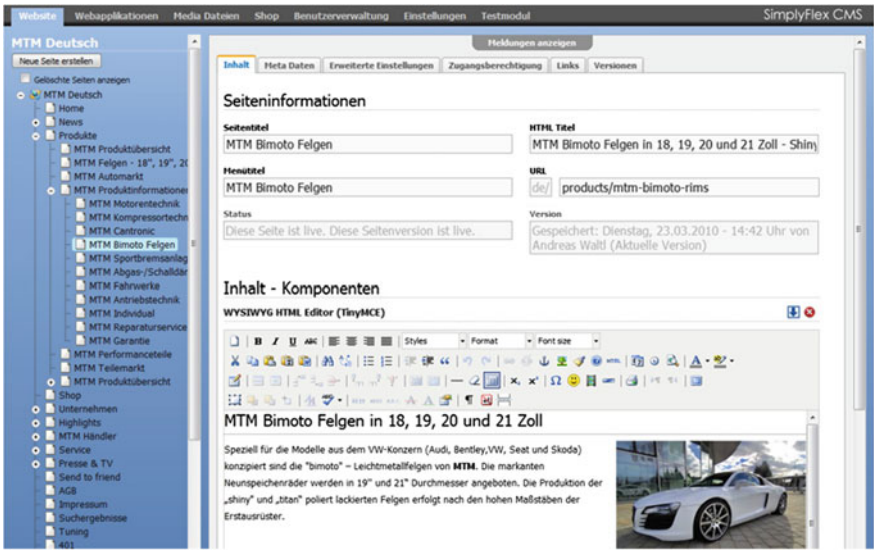


Fig. 2 Content management system

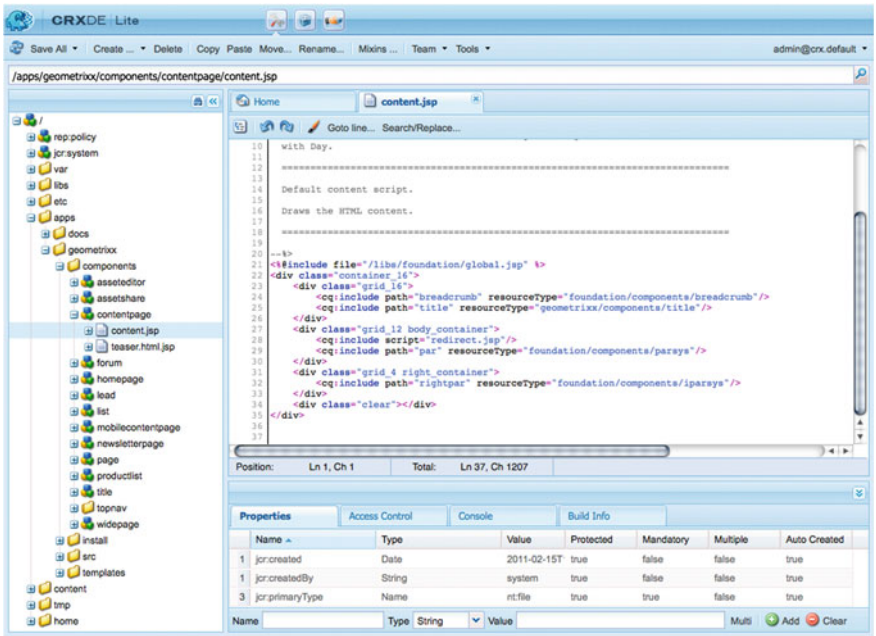


Fig. 3 Environmental management system

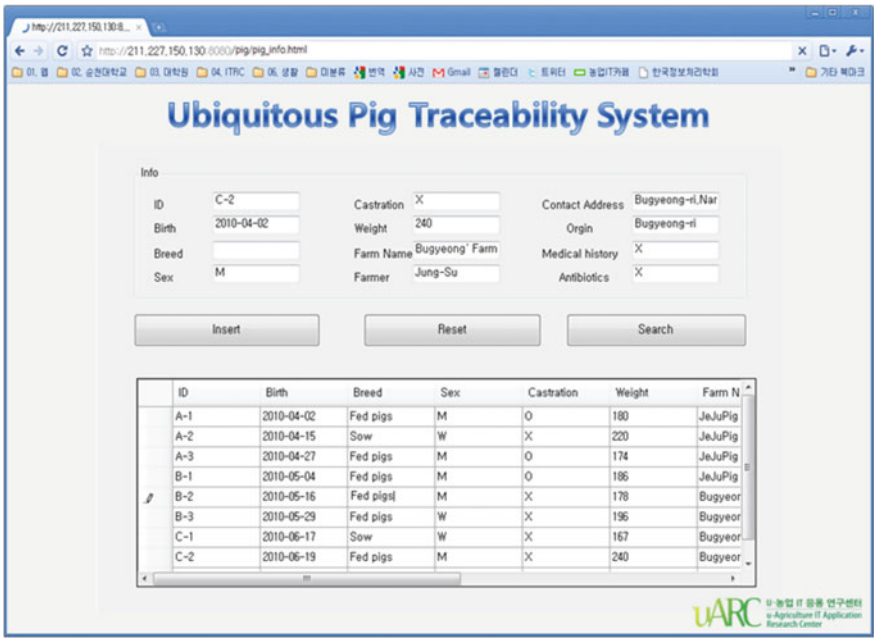


Fig. 4 Welfare management system

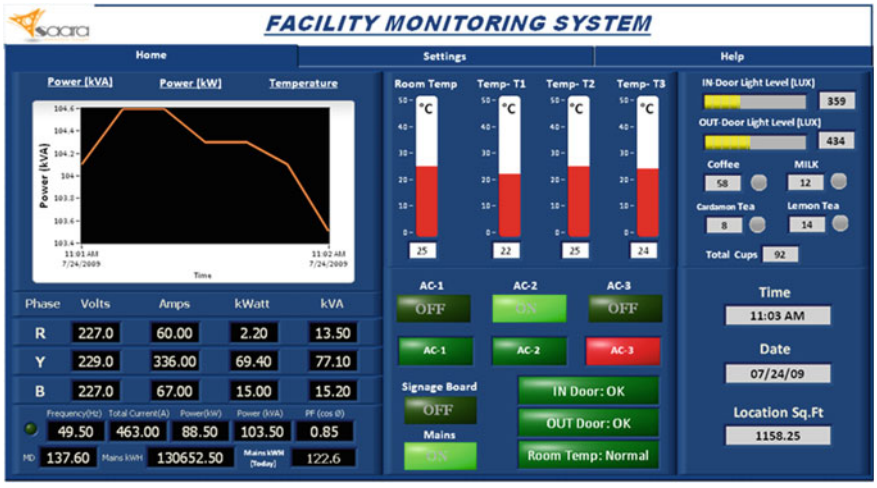


Fig. 5 Quality management system

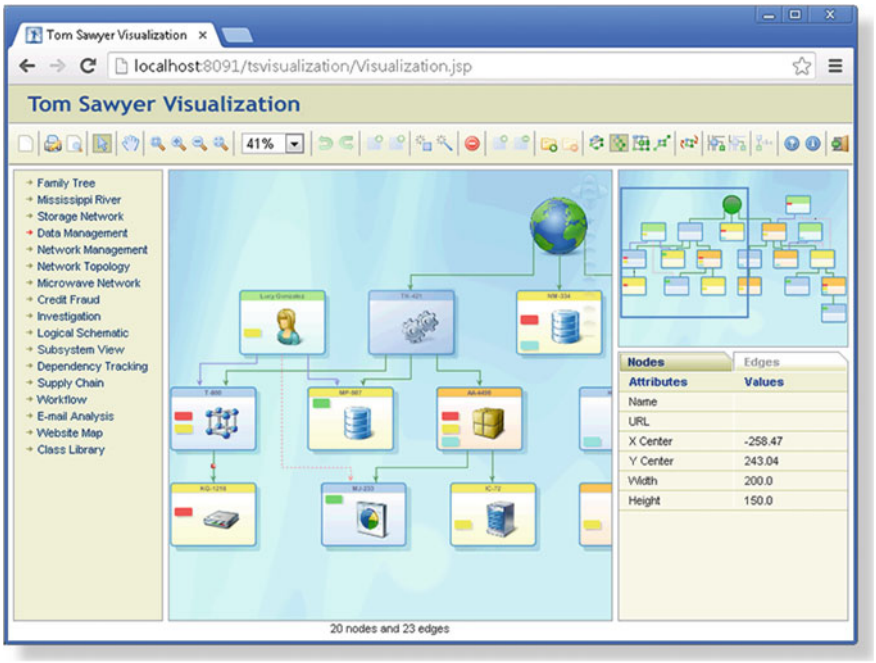


Fig. 6 Another environmental management system

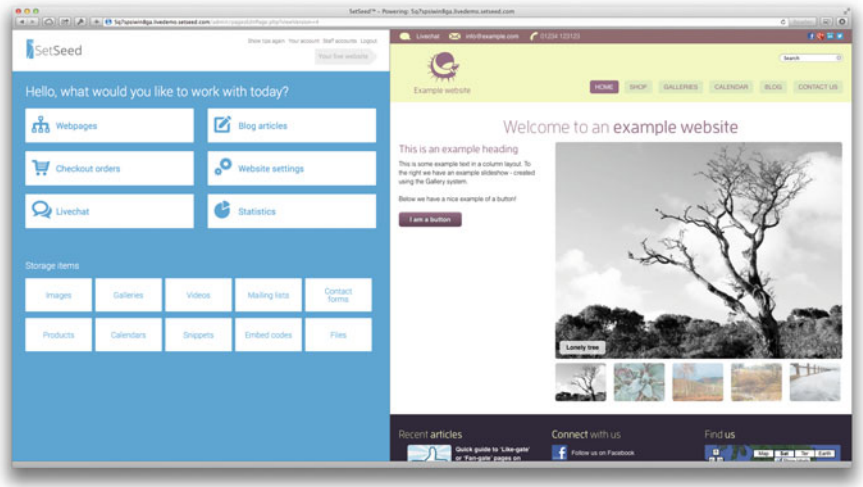


Fig. 7 Environmental management system (SetSeed)

we strongly believe that our proposed solution will serve as a useful tool to help a researcher in monitoring his progress, hence ensuring a successful completion of research studies.

The paper is divided into several sections. The next section describes research related to ours, while the subsequent section explains the methodology in detail. Then, we present our experimental results and finally we conclude our work.

2 Related Work

Research management involves several key process. Some of the key areas are such as literature review, publication record, and research profile. Sometimes, research management can become large that it requires a particular system for managing it. To achieve this, we require a manager to monitor and manage the system. Challenges in terms of leadership increase because of the intrinsic uncertainty in research projects. Such uncertainty requires a high level of autonomy of project participants with specific expertise on specialist domains. The manager knows best the technologies involved, the key requirement, and the many dependencies involved in it. Managers need to balance the various key requirements and dependencies in order to manage the system well.

The tasks involved in research are highly interdependent, and the need for managing two-way interaction between researchers working on different tasks is clear. As such, interaction and collaboration are important features of collaborative research projects, which impose substantial challenges on research managers. The nature of different researchers with different research profile and ability makes the research management system difficult to develop. In addition to that, there exists different discipline in research areas, where different criteria are used to judge the quality of research.

In 2007, Alvesson and Karreman (2007) develop an empirical method for theory development. Bachiochi and Weiner (2002) develop qualitative data collection and analysis in 2002. In 2007, Edmondson and McManus (2007) develop a methodology for research management research whereas Maxwell (1992) develops a method for understanding and validating research. In 1979, Van Maanen (1979) develops a research methodology for organizational structures.

3 Research Questions and Objectives

Based on our proposed system, we have therefore formulated research questions as stated below:

1. How would the system developed benefit the researchers in managing their research?

2. How do the system quantitatively and qualitatively measured a researcher progress?

Based on our research questions, we have embarked on the following research objectives:

1. The developed system will have intuitive User Interface where the users can use the system with ease without much difficultly. Other than that, the users will have easy to follow instructions guide as well as well formatted data.
2. The researcher progress could be effectively measured by conducting a user study. This user study can give deeper insight on the system efficiency in evaluating the researcher progress.

4 Proposed Methodology

We develop a system to effectively measure the student's ability to adapt to research environment. To achieve this, we conduct a comprehensive review on researcher skills and the various factors used to gauge their ability to learn. One of the important factors and metrics used to measure a researcher's skills is to measure the citations and h-index score. The other factors taken into consideration are the ability of the researcher to publish in high-impact venues. Others include the researcher ability to obtain grants as well as network with other researchers in the world. Finally, factors such as conducting literature review and developing a strong proposal will be included in this study.

To obtain the various statistics on a researcher, we develop a simple crawler that could effectively search through the Web for information. A simple query on the researcher's name as well as his profile will make the crawler more focused in its search. For some of the information such as citations and h-index, we use Google API to locate and find the researcher profile as connecting to the API and fetch the data directly is much more efficient than conventional Web search.

Once researcher profile is obtained, we develop a research management system (RMS) to keep track of researcher progress. This system is able to keep track of researcher progress which includes literature review, findings, and thesis writing. When a researcher has read a paper, he will key in his particular data and findings into the system. Using statistical methods, we analyze the pattern and determine his progress of literature review. Once a paper data has been entered, it will be cross-checked with other existing papers and a summary and updates of research profile and progress will be generated.

Once the system is implemented, we conduct user study where the users are selected and divided into group. A closed laboratory session where the system is set up is scheduled. Once the test is completed, we compile and analyze the data to gauge the effectiveness of our system.

We benchmark our system using standard available in the researcher community. Typically, a prolific researcher would have high h-index as well as high

citation scores. A well-established researcher will have good publications in addition to established network and grant projects. A student who progress well in his research will have substantial literature review, well-developed proposal, and a good number of high-impact publications.

5 Experiments

We conduct experiments to gauge the effectiveness of our system. A random users of 20 is selected from different background for the user study. The demographic of our users is as follows.

As can be seen for the two tables above (Tables 1 and 2), we choose random sample of users from diverse background. Since the system is computer-based, we select those in the computing group as well as having IT background. A set of instructions are given so that users are able to adapt and use the system properly. The time allocated for the users to use the system is 2.5 h. A survey form is given to them to fill up the necessary details.

Table 3 provides the user-friendliness test on our application. The majority of users feel that it is simple to use our application and our application provides intuitive interface as well as easy to navigate menus. However, some users find our

Table 1 Users age group

Age range	Number
15–20	3
20–25	12
25–30	3
30–35	1
>35	1

Table 2 Skills set of users

Skills	Number
Engineering	3
Computing	12
Business	4
Others	1

Table 3 User-friendliness test

Rating	Number
Poor	0
Not satisfactory	4
Neutral	6
Satisfactory	8
Excellent	2

Table 4 Performance

Rating	Number
Poor	0
Not satisfactory	2
Neutral	5
Satisfactory	8
Excellent	5

Table 5 Learning experience

Rating	Number
Poor	0
Not satisfactory	2
Neutral	6
Satisfactory	7
Excellent	5

system a bit hard to cope, due to the lack of knowledge they have in research management. Table 4 gives the performance of our system. Our system is able to generally run without crashes. Several bugs are reported, all of which have been fixed. Most users are generally happy with our tool, particularly the way it is able to monitor a research progress and manage it effectively (Table 5).

6 Conclusion

We have developed a novel research management system which could effectively keep track and monitor a researcher progress. Our system introduces many state-of-the-art features which is able to obtain researcher profile as well as monitoring their progress effectively. We integrate a Web search and easy to use menu as well as intuitive interface for users to upload their latest information and finding in research. We hope that our system will be useful for early career researchers to embark on their research without difficulty.

References

Alvesson, M., & Karreman, D. (2007). Constructing mystery: Empirical matters in theory development. *Academy of Management Review, 32*, 1265–1281.

Bachiochi, P. D., & Weiner, S. P. (2002). Qualitative data collection and analysis. In S. G. Rogelberg (Ed.), *Handbook of research methods in industrial and organizational psychology* (pp. 161–183). Oxford: Blackwell.

Edmondson, A. C., & McManus, S. E. (2007). Methodological fit in management field research. *Academy of Management Review, 32*, 1155–1179.

Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review, 62*, 279–300.
Van Maanen, J. (1979). Reclaiming qualitative methods for organizational research: A preface. *Administrative Science Quarterly, 24*, 520–526.

Inquiry-Based Instruction and Situational Interest of Students in Higher Education

Chan Yuen Fook, Suthagar Narasuman, Gurnam Kaur Sidhu, Yap Bee Wah, Lee Lai Fong and Siti Fairus Dalim

Abstract Inquiry-based instruction is an approach which enhances students' academic excellence by engaging them in the learning process through various classroom activities that build their confidence and understanding of the subject matter. Inquiry-based learning appreciates the sense of inquiry and investigation in which questions are answered and students are facilitated to memorize the information through the instructional materials. Alvarado and Herr (inquiry-based learning using everyday objects: hands-on instructional strategies that promote active learning in grades, Corwin Press, 2003) advocated that inquiry-based approach promotes intellectual engagement among students and motivates them to improve their performance. Besides that, Rotgans (Learn Instr 32(2):37–50, 2014) had also argued that student's interest toward their studies is the core element that drives their academic performance. Since preparing the situational interest of students is deemed as an important responsibility of lecturers, hence, the objective of this study was to identify the relationship between inquiry-based instruction and situational interest of students in higher education. In this study, survey questionnaires had been administered to students from different levels of programs, namely diploma, undergraduate, and postgraduate studies. Finally, a total of 470 complete questionnaires had been collected and analyzed by identifying the mean, standard deviation, and correlation between the variables. The results indicated that inquiry-based instruction has been implemented at a high level in higher education. Basically, students investigated had agreed upon their situational interest at a high level. The correlation analysis further indicated that a moderate, positive, and significant relationship between inquiry-based instruction and the situational interest of students in higher education. The findings indicated that teaching strategy adopted by lecturers is a crucial factor in improving interest in learning among students in higher education.

Keywords Inquiry-based instruction · Situational interest · Higher education

C.Y. Fook (✉) · S. Narasuman · G.K. Sidhu · Y.B. Wah · L.L. Fong · S.F. Dalim
Faculty of Education, Universiti Teknologi MARA Selangor, Puncak Alam, Selangor, Malaysia
e-mail: yuenfook@salam.uitm.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_14

1 Introduction

The teaching of twenty-first-century skills requires teachers to urgently reflect on their profession as these latent skills require teachers to perform more than their current practical skills and competencies (Nevgi and Niemi 2014). An effective teaching approach for the twenty-first-century generation involves transition from teacher-centered to learner-centered classroom interaction (Finsterwald et al. 2013). The transformational change demands the teacher, not only to play the role of the center of knowledge but also to play the role of facilitator who guides and leads the students toward developing meaningful learning in the classroom. Twenty-first-century learning expects students to build their own knowledge with strong support from the teacher through activities such as group discussion, case study projects, and experiments embedded within in a lesson. One way of compelling students to actively participate in the class is through inquiry-based instruction (Ebert-May et al. 1997).

Many researchers believe that a student-oriented teaching method will trigger students' learning interest in the classroom (Schraw et al. 2001; Krapp et al. 1992). In relation to this, being facilitators is common in an active learning approach, where teachers aid students to develop their understanding on conceptualizing knowledge (Kember 1997). Alvarado and Herr (2003) specified that inquiry-based instruction is significantly associated with active learning because it is teaching approach and assessment leads to the execution of active learning. Their statement is supported by Hanna and Dettmer (2004) who also advocate the proclivity of inquiry-based instruction leading to active learning. Inquiry-based instruction has the propensity to resolve the learning problems and trigger the interest for active learning among students. Hence, the Organisation for Economic Co-operation and Development (OECD) has proposed the application of inquiry-based approaches in classroom instruction, especially in the application of problem-based activities (Dostál 2015).

2 Literature Review

Inquiry refers to the process through which students develop the basic understanding of the concepts and ideas presented in a classroom setting to acquire knowledge (Colburn 2006). Inquiry-based learning (also enquiry-based learning in British English) starts by posing questions, problems, or scenarios—rather than simply presenting established facts or portraying a smooth path to knowledge. The process is often assisted by a facilitator. Inquirers will identify research issues and questions to develop students' knowledge or solutions. Inquiry-based learning also includes problem-based learning and is generally used in small-scale investigations and projects, as well as research. However, in principle, inquiry-based instruction is closely related to thinking skills and cognitive development (Dostál 2015). This

approach involves getting students to participate in various classroom activities that enhance their confidence and cognizance of learning materials. Invariably, this process enhances students' thinking skills and academic excellence (Colburn 2006).

As we know, instructors play an important role in inquiry-based learning, because they hold the responsibility of establishing the classroom environment and culture, apart from setting the guidelines for student participation in class discussion. Besides that, instructors are also responsible for initiating innovative problem-based activities, case studies, and solutions to problems and questions raised by students. In this sense, inquiry-based learning serves as an effective procedure that generates various questions from students as class discussions progress from one issue to the next. Subsequently, these discussions and the questions raised heighten students' interest and curiosity within the context of the study and help the learner to restructure their knowledge boundaries (Beerer and Bodzin 2004). In fact, instructional methods employed in inquiry learning have the potential to enhance critical thinking skills and solution-seeking strategies. For instance, when students face problems in understanding the subject matter, their curiosity levels are peaked and they begin to seek solutions and enquire and seek more knowledge about the subject to gain a better understanding (Warner and Myers 2009).

Lack of interest has been identified as one of the contributing factors that cause the decline in the quality of graduates (Jocz et al. 2014). According to researchers, interest is a very important factor as it serves as a catalyst in encouraging students to scale new heights of excellence in their field. Motivation levels plunge when students have no interest in what they are doing, and subsequently, they fail in achieving their educational goals. The findings from the research conducted by Jocz et al. (2014) showed that active learning has the potential to help students to gain interest in their studies. However, s*ituational interest is defined as temporary interest that arises spontaneously due to environmental factors such as task instructions or an engaging text* (Schraw et al. 2001). This definition has been supported by Krapp et al. (1992) who highlighted that s*ituational interest* is spontaneous, transitory, and environmentally activated. Indeed, Dewey (1913) was the first to emphasize the crucial role of interest in learning. Dewey believed that interest differed from effort and that it was interest, rather than effort per se, that led to deeper learning. Dewey made two basic assumptions about interest. First, he believed that interest must be present in the classroom to satisfy students' intellectual and personal needs. Second, he believed that interest could be fostered by providing students with a variety of materials and educational opportunities that promoted challenge and autonomy. Schraw et al.'s (2001) study had confirmed about the claims made by Dewey (1913). Schraw et al. (2001) have found that teachers view interest as their responsibility; in fact, teachers were supposed to be actively involved in selecting interesting materials and creating a classroom environment that increases students' interest. In contrast, students tend to adopt a more passive attitude about interest in which interest is something that happens to them, rather than something that occurs because of how they interact with classroom materials.

Situational interest increases learning when a task or to-be-learned information is novel (Hidi 1990), or when information is relevant to a task or learning goal (Schraw and Dennison 1994; Shirey 1992). Situational interest often precedes and facilitates the development of personal interest (Krapp et al. 1992). Situational interest appears to be especially important in catching students' attention, whereas personal interest may be more important in holding it. Schraw et al. (2001) also agreed that situational interest of students is of extreme significance because it is changeable and can be controlled by instructors to a certain extent. Furthermore, situational interest can increase the learning abilities of students. Thus, managing the situational interest of students is a crucial responsibility for instructors.

Based on the above assumptions, Rotgans (2014) has argued that student's interest toward their studies is the core element that drives their academic performance. The interest is basically a sort of motivation that enhances students' concentration toward their studies. Besides that, Ainley (2006) has also stated that development of strategies and practices that help in managing the interest of students is essential for the trainers and mentors because this leads to effective learning. Hence, Sweet et al. (1997) believe that promoting interest in the classroom increases students' intrinsic motivation to learn. Promoting interest of students in the classroom is crucial as the body of knowledge of several past studies has established a sound association between interest and students' motivation to learn. Based on this assumption, this paper has been designed to identify the relationship between inquiry-based instruction and situational interest of students in higher education.

3 Research Method

A descriptive correlational research design has been used to measure the levels of inquiry-based instruction practiced by lecturers and the situation interest attained by students in this study. This design was also instrumental in identifying the relationship between inquiry-based instruction and situational interest of students in higher education. In order to collect the relevant information needed for the study, a survey questionnaire which contains items of the two constructs and demographic factors had been administered to 475 students in a local tertiary institution. The questionnaire used for this study was adapted from Assessment Experience Questionnaire (AEQ) which was developed by Gibbs and Simpson (2003). The reliability of inquiry-based instruction ($r = 0.78$) and situational interest (0.85) has been identified at above average level with the Cronbach's alpha reliability analysis. The respondents selected were from three different levels of programs, namely diploma, bachelor's degree, and master's degree. The quantitative data collected were analyzed using descriptive and inferential statistics. Descriptive statistics such as mean and standard deviation were used to measure the levels of inquiry-based instruction and situation interest, while the Pearson's product–moment correlation was utilized in measuring the relationship between inquiry-based instruction and the

situational interest of students in higher education. Besides that, the findings from the open-ended questions in the questionnaire were analyzed thematically and used to support the quantitative findings.

4 Findings

Table 1 indicates the gender composition of the respondents. Four hundred and seventy-five students had responded to the questionnaire. The majority of the respondents consist of female students (76.0 %), while the rest were male students (21.1 %). A total of 14 respondents (2.9 %) did not provide any response indicating gender in the questionnaire.

Table 2 shows a breakdown of the programs the respondents were undertaking at the time the study was conducted. Among the 475 (100 %) respondents involved in the study, 92 (19.4 %) were undertaking diploma in TESL programs. A total of 213 (44.8 %) were undertaking various bachelor programs in education, while 156 (32.8 %) were undertaking various master's programs in education. Nonetheless, 14 (2.9 %) participants did not respond to this item.

Table 3 shows the descriptive statistics for students' appreciation of inquiry-based instruction practiced by their lecturers in higher education. Generally, the student respondents strongly agreed that their lecturers did ask about their viewpoints ($M = 4.32$, SD = 0.667) during instruction. Most of the student respondents also agreed that their lecturers allowed them to gain active control ($M = 4.31$, SD = 0.665), to encourage them to solve problems ($M = 4.30$, SD = 0.706), and to connect evidence ($M = 4.20$, SD = 0.696) in their learning.

Besides that, the findings indicated that their lecturers model critical thinking ($M = 4.16$, SD = 0.795), allow students to think ($M = 4.16$, SD = 0.742), provide clear instruction ($M = 4.12$, SD = 0.809), praise performance ($M = 4.12$,

Table 1 Gender of the respondents

Gender	Frequency	Percent
Male	100	21.1
Female	361	76.0
Missing	14	2.9
Total	475	100.0

Table 2 Study programs

Study programs	Frequency	Percent
Diploma	92	19.4
Bachelor's degree	213	44.8
Master's degree	156	32.8
Missing	14	2.9
Total	475	100.0

Table 3 Descriptive statistics of inquiry-based instruction practiced by lecturers

Inquiry-based instruction	*N*	Mean	Std. deviation
	Statistic	Statistic	Statistic
Ask viewpoints	469	4.32	0.667
Students gain active control	472	4.31	0.665
Encourage to solve problems	470	4.30	0.706
Encourage to connect evidence	468	4.20	0.696
Model critical thinking	469	4.16	0.795
Allow students to plan	471	4.16	0.742
Clear instruction	469	4.12	0.809
Praise performance	467	4.11	0.822
Facilitate process information	469	4.10	0.728
Conducive learning environment	469	4.08	0.733
Encourage to discover	472	4.08	0.740
Use technology	469	4.05	0.807
Prepare task	470	4.03	0.719
Clarify mistakes in learning	470	4.03	0.771
Reflect lesson and improvement	469	4.02	0.755
Collaborate beyond classroom	468	3.85	0.908
Design rubrics	467	3.85	0.870
Adjust grading process	468	3.79	0.821
Guided discovery	471	3.78	0.892
Encourage to work with experts	464	3.71	0.888

SD = 0.822), facilitate processing of information (M = 4.10, SD = 0.728), and provide a conducive learning environment (M = 4.08, SD = 0.733).

In fact, student respondents also agreed that their lecturers encourage them to discover (M = 4.08, SD = 0.740), use technology (M = 4.05, SD = 0.807), prepare task (M = 4.03, SD = 0.719), clarify mistakes in learning (M = 4.03, SD = 0.771), and reflect on lessons and improvement (M = 4.02, SD = 0.755). However, respondents only moderately agreed that inquiry-based instruction encourages students to collaborate beyond classroom (M = 3.85, SD = 0.908), and their lecturers design rubrics (M = 3.85, SD = 0.870), work with experts (M = 3.71, SD = 0.89), provide guided discovery (M = 3.78, SD = 0.89), and assist in adjusting grading processes (M = 3.79, SD = 0.82).

Table 4 depicts the situational interest experienced by students in their learning in higher education. The student respondents strongly agreed that inquiry-based instruction has certain positive impact on their situation interest in learning. The respondents agreed that active learning has become more interesting (M = 4.13, SD = 0.69), while inquiry-based instruction was practiced by their lecturer, and the feedback provided has motivated them to do better (M = 4.26, SD = 0.68) in their learning. Generally, the respondents like the course because their lecturer involved them in the learning (M = 4.11, SD = 0.775), and lecturer's questions have

motivated them to find the right answer ($M = 4.07$, SD = 0.737). They were interested due to guided discovery in inquiry-based instruction ($M = 4.07$, SD = 0.745); hence, most of the students looked forward to attending the classes ($M = 4.06$, SD = 0.801). However, the respondents only moderately agreed to enjoy learning by minor assignments ($M = 3.09$, SD = 0.87) and to like the course because feedback is provided ($M = 3.95$, SD = 0.85).

Table 5 portrays a moderate, positive, and very significant relationship between inquiry-based instruction and situational interest ($r = 0.604$, p-value < 0.001). The inferential statistical findings from Pearson's product–moment correlation were consistent with the findings from the descriptive data and interview data. For instances, three respondents claimed that their lesson had become more interesting after inquiry-based instruction had been practiced by their lecturers.

> I feel that there was a lot of improvement in learning and I myself have become more interested in learning when the lecturer was using inquiry-based instruction (Respondent 1).

Table 4 Descriptive statistics of situational interest experienced by student respondents

Situational interest	*N*	Mean	Std. deviation
	Statistic	Statistic	Statistic
Feedback motivated to do better	471	4.26	0.680
Active learning is more interesting	470	4.13	0.686
Like course because lecturer involves students	470	4.11	0.775
Lecturer's question motivated me to find right answers	471	4.07	0.737
More interested due to guided discovery	471	4.07	0.745
Look forward to attending classes	470	4.06	0.801
Like course because feedback provided	471	3.95	0.845
Enjoy learning by minor assignments	471	3.90	0.870

Table 5 Relationship between inquiry-based instruction and situational interest in higher education

Variables		Inquiry-based instruction	Situational interest
Inquiry-based instruction	Pearson's correlation	1	0.604**
	Sig. (two-tailed)		0.000
	N	434	431
Situational interest	Pearson's correlation	0.604**	1
	Sig. (two-tailed)	0.000	
	N	431	469

**Correlation is significant at the 0.01 level (two-tailed)

Hence, both the correlation data and interview data inferred that the lessons in the class had become more interesting after the implementation of inquiry-based instruction in the classroom. Obviously, inquiry-based instruction has developed situational interest among the students in the course. Another respondent also said that inquiry-based instruction had enhanced his situational interest in the course.

> All the inquiry-based instruction strategies practised by my lecturer had contributed to the enhancement of my interest in the course. A clear example was when my lecturer asked me to think and respond critically to the assignment given in his course (Respondent 2).

This finding indicates that inquiry-based instruction has contributed to students' situational interest. Another respondent also stated that she feels her situational interest has improved because of the inquiry-based instruction employed by her lecturer.

> Enhancement of situation interest may be due to teaching strategies. The inquiry-based instruction practised by my lecturer had made me understand the topic taught better (Respondent 3).

Therefore, based on the statistical and thematic analysis, the findings have established a moderate, positive, and very significant relationship between inquiry-based instruction and situational interest in the classroom.

5 Discussion

The findings showed that student respondents strongly agreed that inquiry-based instruction is widely practiced by their lecturers in higher education, and students had experienced high levels of situational interest in their course. Furthermore, the Pearson's product–moment correlation had indicated a moderate, positive, and very significant relationship between inquiry-based instruction and situational interest in their course. The findings collected from the interview have also supported the relationship between the two variables.

Students agreed that when lecturers provide feedback while practicing inquiry-based instruction, it can improve their motivation to learn and increase their situational interest. The findings have been supported by some earlier studies on inquiry-based instruction and situational interest. Ainley (2006) stated that situational interest frequently facilitates and develops the personal interest of students because it appears as the tool that catches the students' attention quickly. Furthermore, Hanna and Dettmer (2004) also concluded that inquiry-based learning leads to the execution of active learning because it augments an individual's interest with the study. This argument has been supported by the findings in this study. Felder and Brent (2009) further claimed that active learning can also be implemented through inquiry-based instruction and formative assessment because this

kind of evaluation leads to active teaching. The work of Felder and Brent (2003) specified that active learning expands the situational interest of students and subsequently results in improved interest and persistence in academic endeavors. Chen et al. (2001) concluded that situational interest of students can be improved through three strategies, namely providing learners with meaningful alternatives, cautiously choosing well-organized task, and helping students in accessing suitable background knowledge about the task. These strategies are achievable in the classroom through the application of inquiry-based instruction. According to Colburn (2006), inquiry-based learning is an approach used in active learning which involves the students in a series of learning procedures that enhance their confidence and understanding which subsequently stimulates academic excellence. It is clear now that through inquiry-based teaching, students are actively involved in class and this contributes to the increased situational interest on the subject matter. In fact, inquiry-based instruction can improve students' critical thinking skills by evoking their curiosity, as they face problems in understanding the subject matter (Beerer and Bodzin 2004: Warner and Myers 2009).

6 Conclusion

Many studies on situational interest have focused on effects of interest on learning and the characteristics of text that increase interest in reading. However, this study focused on the use of inquiry-based instruction among instructors and the situation interest attained by students in tertiary education. The study has revealed that there was a moderate, positive, and significant relationship between the two variables in the study. The findings also concluded that providing feedback to students has the potential to develop students' situational interest in learning and they were also able to pay more attention within a shorter time. Hence, there is a definite and obvious impact of inquiry-based learning, active learning, and situational interest on students because all of these approaches have the tendency to drive student's motivation, interest, and performance to higher levels within the context of their studies. Hence, this study advocates that instructors in tertiary education should evaluate their teaching strategies and approaches to ensure that learning objectives are met and student learning's interest has enhanced. The findings indicate that asking relevant questions and helping students to find their answers will enhance students' interest in learning in higher education.

Acknowledgement This paper is part of a research project funded by Fundamental Research Grant Scheme (FRGS), Ministry of Education, Malaysia, and Research Management Institute, Universiti Teknologi MARA.

References

Ainley, M. (2006). Connecting with learning: Motivation, affect and cognition in interest processes. *Educational Psychology Review, 18*(4), 391–405.

Alvarado, A. E., & Herr, P. R. (2003). *Inquiry-based learning using everyday objects: Hands-on instructional strategies that promote active learning in grades* (pp. 3–8). Corwin Press.

Beerer, K., & Bodzin, A. M. (2004). Promoting inquiry-based science instruction: The validation of the science teacher inquiry rubric (STIR) [Online]. Available at http://www.lehigh.edu/~amb4/stir/aets2004.pdf. Accessed on April 1, 2015.

Chen, A., Darst, P. W., & Pangrazi, R. P. (2001). An examination of situational interest and its sources. *British Journal of Educational Psychology, 71*(7), 383–400.

Colburn, A. (2006). An inquiry primer. *Science Scope*, *23*(6), 42–44.

Dewey, J. (1913). *Interest and Effort in Education*. Boston: Houghton Mifflin (Bath, U.K.: Chivers, 1969).

Dostál, J. (2015). Inquiry-based instruction concept, essence, importance and contribution. Olomouc: Palacký University. ISBN 978-80-244-4507-6, doi:10.5507/pdf.15.24445076

Ebert-May, D., Brewer, C., & Allred, S. (1997). Innovation in large lectures-teaching for active learning. *Bioscience Journal, 47*(9), 601.

Felder, R. M. & Brent, R. (2003). Learning by doing. *Chemistry Engineering Education*, *37*(4), 282–283. Available at http://www.ncsu.edu/felder-public/columns/active.pdf.

Felder, R. M., & Brent, R. (2009). Active learning: An introduction. *ASQ Higher Education Brief, 2*(4), 56–100.

Finsterwald, M., Wagner, P., Schober, B., Lüftenegger, M., & Spiel, C. (2013). Fostering lifelong learning—Evaluation of a teacher education program for professional teachers. *Teaching and Teacher Education Journal*, *29*, 144–155.

Gibbs, G., & Simpson, C. (2003). Measuring the response of students to assessment: The assessment experience questionnaire. In *11th International Improving Student Learning Symposium*, Hinckley, England.

Hanna, G. S., & Dettmer, P. A. (2004). *Assessment for effective teaching: Using context-adaptive planning*. Boston, MA: Pearson A&B.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research, 60*, 549–572.

Jocz, J. A., Zhai, J., & Tan, A. L. (2014). Inquiry learning in the Singaporean context: Factors affecting student interest in school science. *International Journal of Science Education, 36*(15), 2596–2618.

Kember, D. (1997). A reconceptualization of the research into university academics' conception of teaching. *Learning and Instruction journal, 7*(3), 255–275.

Krapp, A., Hidi, S., & Renninger, K. A. (1992). Interest, learning and development. In A. Renninger, S. Hidi, & A. Krapp (Eds.), *The in role of interest learning and development* (pp. 3–25). Hillsdale, NJ: Erlbaum.

Nevgi, A., & Niemi, H. (2014). Research studies and active learning promoting professional competences in Finnish teacher education. *Teaching and Teacher Education Journal, 43* (2014), 131–142.

Rotgans, J. I. (2014). Situational interest and learning: Thirst for knowledge. *Learning and Instruction, 32*(2), 37–50.

Schraw, G., & Dennison, R. S. (1994). The effect of reader purpose on interest and recall. *Journal of Reading Behavior, 26*, 1–18.

Schraw, G., Flowerday, T., & Lehman, S. (2001). Increasing situational interest in the classroom. *Educational Psychology Review, 13*(3), 211–224.

Shirey, L. (1992). Importance, interest, and selective attention. In A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 281–296). Hillsdale, NJ: Erlbaum.

Sweet, A. P., Guthrie, J. T., & Ng, M. M. (1997). Teacher perceptions and student reading motivation. *Journal of Educational Psychology, 90*, 210–223.

Warner, A. J., & Myers, B. E. (2009). What is inquiry-based instruction? [Online]. Available at http://edis.ifas.ufl.edu/wc075. Accessed on April 1, 2015.

Evaluation of the Effectiveness of Field Trips in the Teaching and Learning of Biosciences

Malarvili Ramachandiran and Saroja Dhanapal

Abstract Field trips can be defined as a type of experiential learning that gets students out of the traditional classroom setting into a new mode of learning. It is advocated that field trips are one of the most important things educators can provide for their students as they not only expand students' learning and experiences, but also increase students' knowledge and understanding of the world in which they live. Despite ample evidence shown through research on the effectiveness of field trips which emphasizes on hands-on, real life, and practical applications of learning, the dilemma that is faced by educators lies in providing proof of student learning for the current trend in education tends to emphasize mainly on assessments. Therefore, the objectives of this study are, firstly, to evaluate the effectiveness of field trips as an educational tool to enhance student's understanding of the subject taught and, secondly, to show whether report writing can be used as an assessment tool to evaluate the learning that takes place via field trips. Students pursuing an undergraduate degree program in biomedical science were used as the sample where they have participated in field trips which were then assessed via report writing. The results were analyzed using SPSS and the outcome indicated that there is an increased understanding of the subject among the students. The high percentage of scores in the report-writing task confirmed this finding. This was supported by a survey administered to the same sample which showed that students have a positive perception on the impact of the field trips on their understanding of the subject. The paper ends with recommendation for the incorporation of field trips in the curriculum to stimulate better understanding and increase motivation towards the learning of science. It is further advocated that report writing can be used as an effective assessment tool to evaluate this increased understanding.

Keywords Field trips · Experiential learning · Biosciences · Assessment

M. Ramachandiran (✉)
Taylor's University, Subang Jaya, Malaysia
e-mail: Malarvili.Ramachandiran@taylors.edu.my

S. Dhanapal
University of Malaya, Kuala Lumpur, Malaysia
e-mail: saroja.dhanapal@um.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_15

1 Introduction

Field trips may also be termed as an instructional trip, excursion, or educational journey. Krepel and Duvall (1981) define them as school or class trips with an educational intent, in which students interact with the setting, displays, and exhibit to gain an experiential connection to the ideas, concepts, and subject matter. Further, Tal and Morag (2009) described field trips as students' experiences outside the classroom at interactive locations designed for educational purposes. Behrendt and Franklin (2014) have divided field trips into two types: formal and informal. They distinguish the two by stating formal field trips consist of planned, well-orchestrated experiences where students follow a documented format while informal field trips are less structured and offer students some control and choice concerning their activities or environment. The significance of field trips in education in the twenty-first century has been summarized by Montgomery County Public Schools (2006) in a comprehensive manner:

> Field trips enable teachers to expand students' learning beyond the walls of the classroom into the vast community outside. They provide students with experiences that cannot be duplicated in the school, but are nonetheless an integral part of school instruction. … a field trip can best be described as a living laboratory in which learning is acquired through active, hands-on experience with the rich resources of the local community.

In reference to the role of field trips in the curriculum, Lei (2010b) explains that the connection between the field trip venue and the classroom links the field trip's experiential learning with prior experiences and learning from the classroom. She further added that each student observes the natural settings and creates personally relevant meaning to the experience and this helps the students play with concepts or activities often not possible in the classroom. In another article, Lei (2010a) asserted that earlier course content becomes relevant as students assimilate and accommodate new understanding and cognition. The purpose of field trips is well explained by Michie (1998), claiming that field trips may be planned for five purposes: to provide first-hand experience, to stimulate interest and motivation in science, to add relevance to learning and interrelationships, to strengthen observation and perception skills, and to promote personal (social) development. According to Touhidur and Spafford (2009), field trips are considered essential in biological science education. They stated that learning in biological sciences traditionally takes place in one or more of three different environments: the classroom or lecture theater, the laboratory, and the field (outdoors). They also asserted that in biological science education, the field trip is considered a key component of the curriculum, and when combined with lectures, problem sets, readings, and workshops, it provides vital experience necessary for converting students into successful professionals. Despite evidence pointing to the benefits of field trips, there is controversy on a number of issues: whether each unit in the biological sciences needs field trips, whether the learning outcomes of a field trip are worth the time and money, and are these superficial activities or bona fide learning experiences.

In discussing field trips, it must be noted that it is closely linked to experiential learning. Kolb (1983) defines experiential learning as authentic, first-hand, sensory-based learning where learning consists of grasping an experience and then transforming it into an application or result through experiential activities which enable students to explore, touch, listen to, watch, move things, dissemble, and reassemble. The Association for Experiential Education (2012) defined experiential learning as a methodology in which educators direct students to a specific experience and then guide the students through reflection to "increase knowledge, develop skills, clarify values, and develop people's capacity to contribute to their communities." According to Behrendt and Franklin (2014) and Kolb's (1983) learning cycle, learning experientially requires the learner to have an experience and then reflect, analyze, and test the idea to develop knowledge and to create another experience. It is important to note that teachers have been doing this since time immemorial through learning format in the formal classroom, through laboratories and projects. However, it is advocated that informal experiential learning can be an equally powerful learning tool with unique virtues (Behrendt and Franklin 2014).

Research on knowledge gain and learning that occurs during field trips is extensive (Hudak 2003; Kisiel 2006; Mawdsley 1999; Michie 1998; Nadelson and Jordan 2012; Scarce 1997; Scribner-MacLean and Kennedy 2007). Current research (Kisiel 2006; Martin and Seevers 2003; Knapp 2002) has shown that field trips are essential for many reasons and among the key reasons highlighted are that they provide real experiences related to all content areas, extend learning by expanding a child's world, and provide a framework for learning. Nabors et al. (2009) have listed four specific benefits of field trips: they enrich and expand the curriculum, strengthen observation skills by immersing children into sensory activities, increase children's knowledge in a particular subject area, and expand children's awareness of their own community. In addition, they also stated that field trips provide living laboratories where children acquire knowledge outside the realm of the regular classroom. Behrendt and Franklin (2014) also voiced similar views by stating that experiential learning at formal and informal field trip venues increases student interest, knowledge, and motivation. Scarce (1997), on the other hand, stated that field trips are experiential, authentic social events that create a new way of knowing an object, concept or operation. In line with this, it is held that such quality experiences lead to deeper learning and interest development (NRC 2009).

Despite the benefits that have been proven to exist from field trips, this mode of learning does pose some serious challenges to educators. Michie (1998) identified seven barriers to successful field trips: transportation; teacher training and experience; time issues such as school schedule and teacher's ability to prepare; lack of school administrator support for field trips; curriculum inflexibility; poor student behavior and attitudes; and lack of venue options. Further, according to Mawdsley (1999) and Scarce (1997), finding time for the trip and making arrangements for students who cannot make the trip adds tasks to an already-busy teacher schedule. Mirka (1970) and McKenzie et al. (1986) refer to three sources for the reason a gap exists between the educational potential of field trips and its realization which include logistic limitations, lack of adequate teaching/learning materials, and

teachers' unfamiliarity with the outdoor as a learning environment. These concerns have added on to the existing concern on assessing the learning outcomes of field trips.

It must be noted that the assessment process has become more onerous now than in the past as the various stakeholders in the process demand more transparency, increased reliability, and more accountability (Hornby and Laing 2003). According to them, this is intensified further due to the demands on staff which has significantly increased as a result of increased student numbers, Quality Assurance Agency guidelines, and the introduction of new university-wide systems such as the common grading scheme, anonymous, and double marking. Gibbs (1992) had highlighted the necessity for us to review existing practice since assessment methods designed for an era with a smaller and a more homogeneous student population is thus unlikely to be effective now. According to Feller and Stern (2007), various assessment methodologies have arisen from different disciplinary and multidisciplinary perspectives, measuring different aspects of scientific activity and addressing various science and technology policies and assessment questions. Report writing is becoming more common as a form of assessment in the various fields of sciences. In the context of assessing students' written work, Broad (2003) stated that the precise purpose of ascoring guide and rubric is to rebuild readers' evaluative frameworks so they can agree more consistently and more quickly. He also added that these tools are effective as impressively high correlations are reached among scorers, and students' writing can be judged with remarkable speed and relative ease.

Research on the effectiveness of field trips and assessment has been an ongoing area of research. However, it must be noted that literature review is still scarce in the two key areas: understanding students' perceptions on field trips' impact on their learning outcome and analyzing the effectiveness of report writing in evaluating students' learning during field trips. Thus, this research aims to fill these gaps. As such, the objective of this paper can be seen to be twofold. Firstly, it aims to examine the effectiveness of field trips as an educational tool to enhance student's understanding of the module taught, and secondly to analyze how report writing can be used as an assessment tool to evaluate the learning that takes place via field trips.

2 Methodology

This study adopted a quantitative research method as the strengths of this method are suited to the purpose of the study where the objective is to evaluate the effectiveness of report writing as a tool for assessing students' learning outcomes achieved from field trips. The population targeted for this study were undergraduate students from a biomedical science degree program in a reputable institution of higher learning in Malaysia. The data for the study were collected through two methods: a 37-item Likert scale survey questionnaire and a formal report-writing task assessed via a specified marking rubric.

The questionnaire entailed two parts: Sect. 1 was to collect demographic data and Sect. 2 was to identify the students' perceptions on the effectiveness of field trips as a learning tool. Section 2 was further divided into five subsections with questions addressing specific areas related to the study: Section A (Knowledge and Lifelong Learning), Section B (Soft Skills Development), Section C (Suitability of Learning Environment), Section D (Citizenship and Global Perspectives), and Section E (Student Satisfaction). Section 2 used a five-point Likert scale ranging from 1 to 5, Strongly Disagree to Strongly Agree. The researchers adopted the Likert-type scale as it can be easily read and understood by the respondents as well as enable the researchers to easily construct and manage the data (La Marca 2011). The questionnaires were distributed to 52 respondents comprising of second-year biomedical science students on return from the field trips. A purposive sampling method was used for this study, whereby biomedical science students undertaking two modules which incorporated field trips as part of the summative assessment were selected. In order to ensure confidentiality, the questionnaires were completed anonymously.

In addition to the survey, the respondents were also required to produce a report on the field trip addressing specific question comprising descriptions of services and processes observed during the field trip as well as to conclude on the positive-learning outcomes of the trip. These reports were then evaluated by the module leader using the prescribed marking rubric and to ensure validity and reliability of the evaluation, all the reports were moderated by a qualified peer. The data from the survey questionnaire and the reports were analyzed using SPSS version 23 to provide answers to achieve the objectives of the study.

3 Findings and Discussion

This section of the paper discusses the findings of the research according to the two research objectives: to examine the effectiveness of field trips as an educational tool to enhance students' understanding of the module taught and, secondly, to show how report writing can be used as an assessment tool to evaluate the learning that takes place via field trips. Prior to the discussion, a brief description of the demographics of the sample used in the study is given.

A total of 52 questionnaires were distributed among second-year biomedical science students comprising both males and females and the data collected via report scores were from the same respondents. All analyses in this study are based on these respondents. The demographics of the respondents are shown in Table 1.

Table 1 Gender distribution of respondents

	Frequency	Percentage
Male	18	35
Female	34	65
Total	52	100

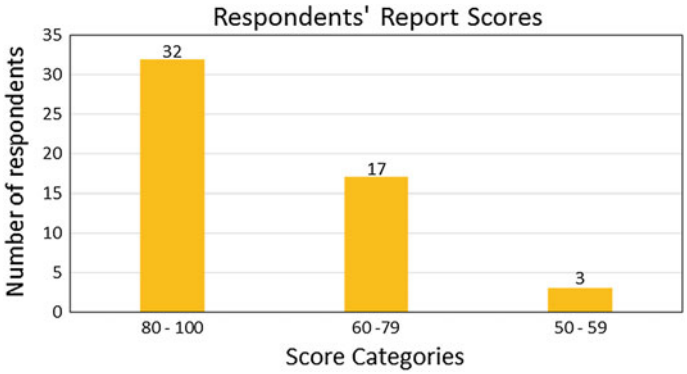


Fig. 1 Respondents' field trip report scores

The findings indicate that a total of 35 % (n = 18) of the respondents were males and 65 % (n = 34) were females. In reference to the ages of the respondents, the findings indicated that they ranged from the age of 19 to 24 with the mean age of 21.4.

The scores for field trip reports submitted by the students were tabulated and were further categorized for analysis. Scores obtained by the students were divided into three categories: high (80–100 %), moderate (60–79 %), and low (40–59 %). Figure 1 shows the results obtained by students for the individual submission of their field trip reports. Based on the analysis, it was noted that 61 % (n = 32) of students obtained a score between 80 and 100 % while 33 % (n = 17) obtained a score between 60 and 79 %. Both of these score categories fall into A and B grades, respectively. This indicates that although the students went for the field trip without much prior knowledge or information, they had managed to learn and gather information on-site and delivered it effectively in their reports. A minority of students (6 %) obtained a score of between 50 and 59 %. This is mainly due to the fact that students were unable to address all relevant questions pertaining to the trip.

The mean score obtained by the students was 78.4 % with a standard deviation of 1.4 %. The lowest and highest scores obtained by the students were 53.3 and 90.0 %, respectively. These findings were further compared to the students' perceptions on acquisition of knowledge and lifelong learning field trips which were obtained using the survey questionnaire.

In order to address the second objective which is to analyze how report writing can be used as an assessment tool to evaluate the learning that takes place via field trips, the results from the survey questionnaire will be discussed according to the respective subheadings: Section A (Knowledge and Lifelong Learning), Section B (Soft Skills Development), Section C (Suitability of Learning Environment), Section D (Citizenship and Global Perspectives) and Section E (Student Satisfaction).

3.1 Section A: Knowledge and Lifelong Learning

Figure 2 shows the findings of the respondents' perceptions of the knowledge and lifelong learning that was achieved from the field trips. The findings in Fig. 2 indicate that there is an overall positive perception toward the use of field trips in the teaching and learning process. Out of the 10 statements put forth to the respondents, only 4 statements showed a very minimal level of disagreement. This comprises A4: The field trip provides me a platform to clear my doubts from classroom on the subject, A6: I find learning during the field trip more effective than learning in classroom, A8: I am able to put theories into practice (hands-on experience and demonstration given) and A9: The trip enables me to share the knowledge gained with friends and family. The percentages of disagreement were 1.9, 3.8, 5.8, and 1.9 %, respectively. Clarification on this was sought through an informal session with all the respondents and some of them claimed that the duration of the field was too short to provide an avenue for them to clear doubts (A4) as well as have a hands-on experience on site (A8). In terms of comparing the effectiveness of field trips and traditional classroom learning, only a small percentage (3.8 %) felt that the classroom learning was more effective. The highest percentage of agreement (92.3 %) as to the benefits of field trips was indicated in A1 and A10 where the respondents agreed that the field trips helped them understand the lesson better and also the visual exposure strengthened their understandings. This finding is supported by the analysis of the report-writing scores which indicates high percentage of passing rates.

Table 2 indicates the respondents report scores with regard to their perceptions in Section A of the questionnaire. The table summarizes data from cross-tabulation

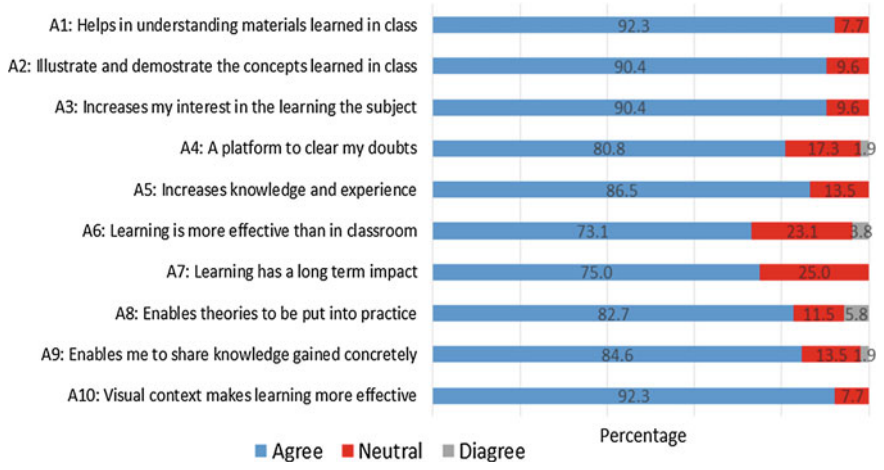


Fig. 2 Respondents' perceptions on knowledge and lifelong learning

Table 2 Comparison of respondents' report scores and perceptions on knowledge and lifelong learning

Statements	Findings		Score category (%)			Total
			80–100	60–79	50–59	(%)
A1: Helps in understanding materials learned in class	Neutral	% within A1	50	25	25	100
	Agree		64	32	4	100
	Strongly agree		60	35	5	100
A2: Illustrate and demonstrate the concepts learned in class	Neutral	% within A2	40	60	0	100
	Agree		58	37	5	100
	Strongly agree		68	25	7	100
A3: Increases my interest in the learning the subject	Neutral	% within A3	60	40	0	100
	Agree		58	42	0	100
	Strongly agree		67	19	14	100
A4: A platform to clear my doubts	Disagree	% within A4	100	0	0	100
	Neutral		44	56	0	100
	Agree		56	36	8	100
	Strongly agree		76	18	6	100
A5: Increases knowledge and experience	Neutral	% within A5	57	29	14	100
	Agree		54	42	4	100
	Strongly agree		74	21	5	100
A6: Learning is more effective than in classroom	Strongly disagree	% within A6	100	0	0	100
	Disagree		100	0	0	100
	Neutral		67	33	0	100
	Agree		46	54	0	100
	Strongly agree		64	24	12	100
A7: Learning has a long-term impact	Neutral	% within A7	62	38	0	100
	Agree		44	50	6	100
	Strongly agree		74	17	9	100
A8: Enables theories to be put into practice	Strongly disagree	% within A8	0	100	0	100
	Disagree		100	0	0	100
	Neutral		83	17	0	100
	Agree		63	29	8	100
	Strongly agree		58	37	5	100

(continued)

Table 2 (continued)

Statements	Findings		Score category (%)			Total
			80–100	60–79	50–59	(%)
A9: Enables me to share knowledge gained concretely	Disagree	% within A9	100	0	0	100
	Neutral		100	0	0	100
	Agree		50	46	4	100
	Strongly agree		63	25	12	100
A10: Visual context makes learning more effective	Neutral	% within A10	75	25	0	100
	Agree		34	58	8	100
	Strongly agree		69	25	6	100

analysis indicating the percentage of respondents scoring the specified high, moderate, and low scores within the different levels of agreement to each statement in Section A. A chi-square test was also performed to examine whether or not a relationship exists between the field trip report scores obtained by students and the perception of the respondents toward field trips for gaining knowledge and lifelong learning experience by comparing all 10 statements in Section A with the scores obtained.

By using a p value of 0.05, the results indicated that there is no significant difference between the two variables: perception on knowledge acquisition and report-writing scores. Thus, this indicates there is no relationship between the two variables: report scores and perceptions. The discrepancies may be due to the small sample size in this study. However, the frequency of the agreement for every statement indicates that the respondents had an overall positive perception immaterial of the grades obtained be it a high pass or a moderate pass. This further explains the findings in the table indicating that higher disagreement does not result in lower scores (A4, A6, A8, and A9). A conclusion can be made that since the respondents' performance totals to a 100 % passing rate based on the reports, field trips are indeed effective teaching methods and report writing does assess their performance in an effective manner for it parallels the positive perception respondents have about field trips on the learning outcomes.

3.2 Section B: Soft Skills Development

Figure 3 shows the findings of the respondents' perceptions on the soft skills developed during field trips. Among the skills suggested to the respondents include Communication skills (B1), Team spirit (B2), Confidence (B3), Professionalism (B4), and Intellectual Discussion (B5).

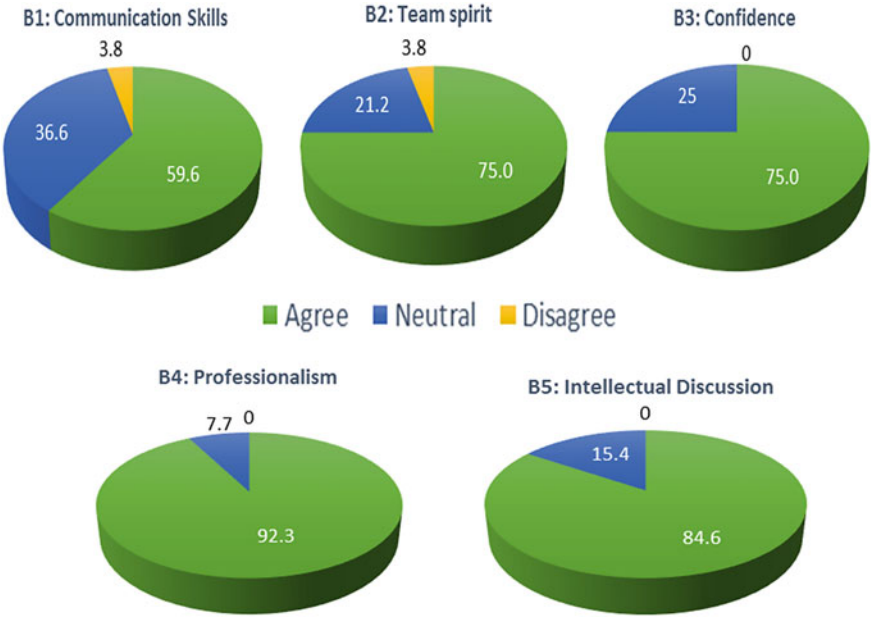


Fig. 3 Respondents' perceptions on development of soft skills

An analysis of the findings indicates that a high percentage of respondents' agreed that field trips help to develop the five skills suggested: 92.3 % agreed that that it develops professionalism, 84.6 % agreed that it develops intellectual engagement, 75 % agreed that it develops both confidence and team spirit and 59.6 % agreed that it develops communication skills. These findings were surprising as prior to the research being conducted, the researchers had a presumption that field trips would develop communication skills the most. As a result, the researchers randomly selected a few respondents and asked them the reasons for the lower percentage of agreement toward communication skills being developed through field trips. The respondents' responses were unanimous. They claimed that the duration of the field trips was too short and they were totally caught up with the experience of visualizing the processes and industrial practices in a new environment leaving them with limited time to discuss and communicate with one another. Furthermore, in most occasions there was more one-way communication where they needed to listen to the presenters from the respective organizations. This is supported by Rennie (2007) who reported that there are minimal opportunities for students to personally interact and connect to the experience. Development of professionalism was recorded as the skill with the highest response of agreement which can be attributed to the experience they had in the locations of the field trips which comprised of research and diagnostic centers in government and private organizations which practice high levels of professionalism in their operations. This had indirectly transferred to the students who realized the necessity of maintaining

such professionalism since they too may have a career in similar fields. According to Behrendt and Franklin (2014), increased understanding in knowledge promotes learning and higher order thinking skills which further increases confidence and intrinsic motivation. Seventy-five percent of respondents agreeing that field trips increased their confidence level is suggestive of the ability of students to understand the content of the module more clearly.

3.3 Section C: Suitability as Learning Environment

The survey also elicited responses from the respondents on their perceptions of the suitability of field trip locations as an effective learning environment. The findings are shown in Fig. 4.

The respondents in this study were taken to two locations comprising a diagnostic center located in Subang, Selangorand a national blood center located in Kuala Lumpur. Both these locations were selected based on the relevance to the modules taught. A list of suitable locations were identified and subsequently two highly recognized establishments were selected based on discussion with peers teaching similar modules in other universities. A total of 88.5 % of respondents agreed that the locations selected has motivated them to participate in more field trips while 86.6 % agreed that the field trips to the locations has made the module more fun and enjoyable. In terms of the locations having helped them to appreciate and value the field of study, keeping them intellectually engaged, and enabling them to gather information through the utilization of efficient technologies (smart devices), the percentage of respondents' agreement was the same, 84.6 %. The

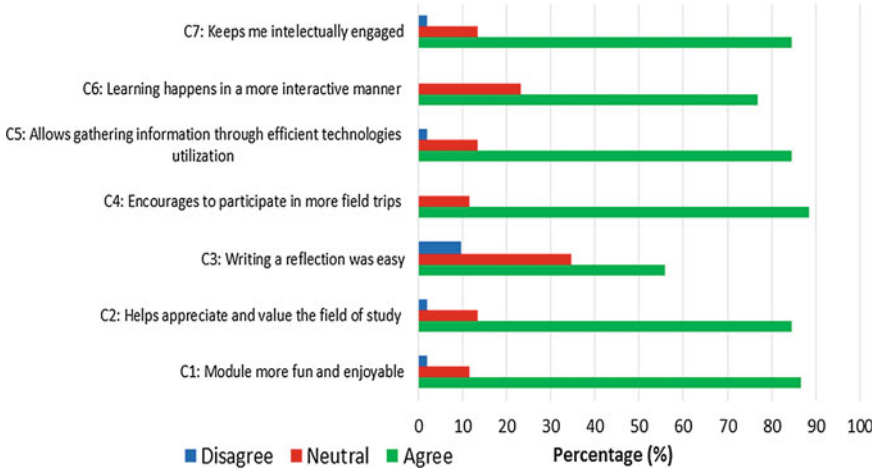


Fig. 4 Respondents' perceptions on learning environment

lowest percentage of agreement was seen in the respondents' perception toward writing a reflection on the field trips with only 16 respondents (55.8 %) agreeing that it was easy to do so. This finding, however, is contradictory to the analysis of the field trip report scores which indicate 32 respondents (62 %) obtained a high score of 80–100 %, 17 students (33 %) obtained 60–79 %, and 3 students (5 %) obtained 40–59 % totaling to 100 % passing rate.

3.4 Section D: Citizenship and Global Perspectives

In order to evaluate the effectiveness of field trips, the respondents were also asked to indicate their agreement or disagreement toward field trips' contributions toward citizenship and global perspectives. The findings are shown in Fig. 5. Majority of the respondents (94.2 %) agreed that the field trips enabled them to understand the related professional practice in the field of biomedical science and another 90.4 % agreed that the field trips increase knowledge on social responsibility and current issues. The lowest percentage of agreement (75 %) was observed in the statement that field trips help respondents to understand the ethical issues in the field of biomedical sciences. However, it must be noted that the percentage is still positive. Nabors et al. (2009) reported that when on a field trip, the venue of the trip is not the only element that affects students as they also gather knowledge and understanding about their surroundings and communities. The findings shown in Fig. 5 relate to the citizenship and global perspectives whereby students understand the society and environment they live in better.. Findings of this study are also in agreement with

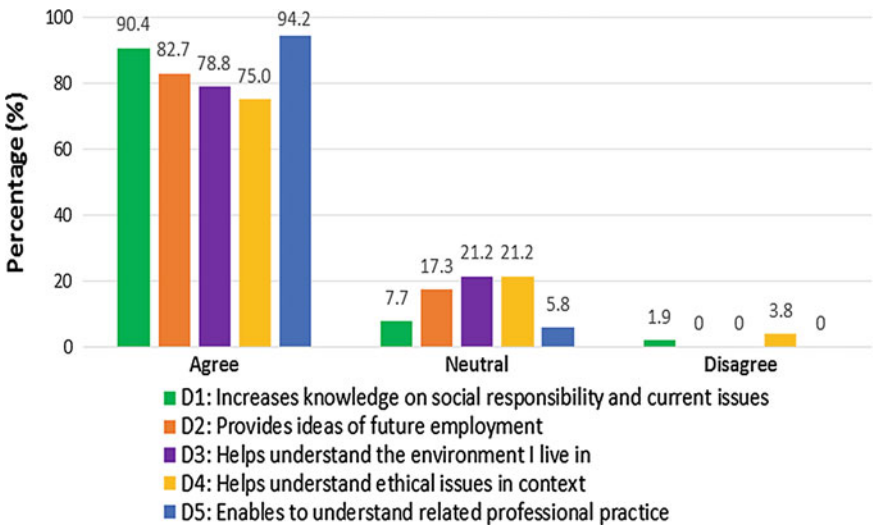


Fig. 5 Respondents' perceptions on citizenship and global perspectives

those of Hutson et al. (2011) who stated that field trips with a single focus would provide a potential impact to students' cognitive skills, knowledge, interests, and future career. This too supports the high percentage of agreement among the respondents in this study that not only does the trips create interest in learning, but also provides ideas for future employments in the similar field.

3.5 Section E: General Feedback and Satisfaction

In addition to the above specific findings, some general statements regarding respondents' levels of satisfaction with the field trip were also sought through the questionnaire. Some of the key statements relevant to the present study are discussed. In terms of suitability and relevance of places selected, all the respondents (100 %) agreed and 98.1 % agreed that the objectives of the field trips were clearly outlined. This finding is quite consistent with the percentage (96.2 %) agreeing that the objectives were achieved. 90.4 % respondents gave feedback that they would like to be involved in the location selection procedure. On the whole, the respondents gave positive responses about the staff they met at the respective field trips. In this study, report writing seems to be well suited to the selected modules as it not only required the students to provide specific descriptions of the processes seen but also reflect on the benefits and outcomes of the trip. However, in assessing knowledge gained through field trips, other methods of assessing students learning and perceptions' have also been documented in several literatures. For example, in a study conducted among students taking biology using pre- and post-test method also concluded that there was a significant increase in knowledge (Prokop et al. 2007).

4 Conclusion

The findings of this study indicate two strong conclusions. Firstly, it supports the effectiveness of field trips as a teaching and learning tool. Secondly, it has proven that report writing is an effective method of assessing the outcome of field trips as the majority of the students performed well in the report and this is supported by their positive perceptions of field trips. It can also be concluded that field trips provide an opportunity to motivate and engage students to appreciate and understand concepts taught in the classroom. However, it must be noted that there was a limitation in this study in terms of sample size. Although the sample size is rather small in this research which is attributed mainly to the nature of the program and the intakes, the data were still adequate for the analysis. The small number of students participating in the field trip also enabled the students to be more engaged during

the trips without distraction. The researchers recommend that in future, educationalists should incorporate field trips in the curriculum to reap the benefits in students' learning process to support the concept of teaching and learning within and beyond the classroom.

References

Association for Experiential Education (AEE). (2012). *What is experiential education?* Retrieved April 14, 2015, from http://www.aee.org/about/whatIsEE

Behrendt, M., & Franklin, T. (2014). A review of research on school field trips and their value in education. *International Journal of Environmental and Science Education, 9*(3), 235–245.

Broad, B. (2003). *What we really value: Beyond rubrics in teaching and assessing writing*. All USU Press Publications. Book 140. http://digitalcommons.usu.edu/usupress_pubs/140

Feller, I., & Stern, P. C. (2007). *A strategy for assessing science. Behavioral and social research on aging*. Washington: National Academies Press.

Gibbs, G. (1992). *Assessing more students*. Oxford: Oxford Centre for Staff Development.

Hornby, W., & Laing, D. (2003). Assessment survey report no 1: Efficiency and effectiveness in assessment. Retrieved from https://www4.rgu.ac.uk/files/EFFICIENCY%20AND%20EFFECTIVENESS%20IN%20ASSESSMENT.pdf

Hudak, P. (2003). Campus field exercises for introductory geoscience courses. *Journal of Geography, 102*(5), 220–225.

Hutson, T., Cooper, S., & Talbert, T. (2011). Describing connections between science content and future careers: Implementing texas curriculum for rural at risk high school students using purposefully-designed field trip. *Rural Educator, 31*, 37–47.

Kisiel, J. (2006). More than lions and tigers and bears-creating meaningful field trip lessons. *Science Activities, 43*(2), 7–10.

Knapp, D. (2002). Memorable experiences of a science field trip. *School Science and Mathematics, 100*(2), 65–72.

Kolb, D. (1983). *Experiential learning, experiences as the source of learning and development*. Englewood Cliffs, New Jersey: Prentice Hall.

Krepel, W. J., & Duvall, C. R. (1981). *Field trips: A guideline for planning and conducting educational experiences*. Washington, DC: National Science Teachers Association.

La Marca, N. (2011). The Likert scales: Advantages and disadvantages. Retrieved January 07, 2015 from Field research in organizational psychology. https://wordpress.com/2011/12/05/the-likert-scale-advantages-and-disadvantages/

Lei, S. A. (2010a). Assessment practices of advanced field ecology courses. *Education, 130*(3), 404–415.

Lei, S. A. (2010b). Field trips in college biology and ecology courses: Revisiting benefits and drawbacks. *Journal of Instructional Psychology, 37*(1), 42–48.

Martin, S., & Seevers, R. (2003). A field trip planning guide from early childhood classes. *Preventing School Failure, 47*(4), 111–180.

Mawdsley, R. D. (1999). Legal issues involving fieldtrips. *School Business Affairs, 65*(9), 28–31.

McKenzie, G. D., Utgard, R. O., & Lisowski, M. (1986, September/October). The importance of field trips as a geological example. *Journal of College Science Teaching*, 17–20.

Michie, M. (1998). Factors influencing secondary science teachers to organise and conduct field trips. *Australian Science Teacher's Journal, 44*, 43–50.

Mirka, G. D. (1970). *Factors which influence elementary teachers' use of out-of-doors*. Columbus: The Ohio State University.

Montgomery County Public Schools. (2006). http://www.mcps.k12.md.us. Accessed June 22, 2015.

Nabors, M. L., Edwards, L. C., & Murray, R. K. (2009). Making the case for field trips: What research tells us and what site coordinators have to say. *Education*, *129*(4).

Nadelson, L., & Jordan, J. (2012). Student attitudes toward and recall of outside day: An environmental science field trip. *Journal of Educational Research, 105*(3), 220–231.

National Research Council. (2009). *Learning science in informal environments: People, places, and pursuits*. Washington, DC: The National Academies Press.

Prokop, P., Tuncer, G., & Kvasnicak, R. (2007). Short-term effects of field programme on students' knowledge and attitude toward biology: A Slovak experience. *Journal of Science Education and Technology*, *16*(3).

Rennie, L. J. (2007). Learning outside of school. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education*. Mahwah, New Jersey: Erlbaum.

Scarce, R. (1997). Field trips as short term experiential education. *Teaching Sociology, 25*, 219–226.

Scribner-MacLean, M., & Kennedy, L. (2007). More than just a day away from school: Planning a great science field trip. *Science Scope, 30*(5), 57–60.

Tal, T., & Morag, O. (2009). Reflective practice as a means for preparing to teach outdoors in an ecological garden. *Journal of Science Teacher Education, 20*(3), 245–262.

Touhidur, R., & Spafford, H. (2009). Value of field trips for student learning in the biological sciences. Teaching and Learning Forum 2009. http://ctl.curtin.edu.au/events/conferences/tlf/tlf2009/refereed/rahman.html. Accessed 22 June 2015.

Conceptual Understanding: A Transitional Process from Novice to Expert User in Radio Frequency and Microwave Course

Md. Shahar Aftanasar and Norizah Mohamad

Abstract Assessing conceptual understanding helps depict competency and expertise in a course. This paper investigates the conceptual understanding and competency in radio frequency and microwave (RFnM) course. Current assessments were unreflective in acquiring in-depth conceptual understanding on microstrip. In this paper, a new assessment using recorded interview was added to measure conceptual understanding and improvements based on course duration, academic performance, and peer review. The new assessment was performed in week 7 during Part 1 and week 14 during Part 2. There were in total 17 participants from third-year electronic program. Scoring rubric was specifically developed to measure student's conceptual understanding observed from verbal screenshot using cloud application available online. Findings have shown that conceptual understanding takes time to develop and retaining meaningful information improved from multiple representations on underlying concept. Study also signaled that self-efficacy may possibly elevate novice learners to proficient or expert learner.

Keywords Conceptual understanding · Multi-instructional learning · Novice–expert learner · Technology-based feedback

1 Introduction

Conceptual understanding helps student in learning and emulates how experts use information and knowledge in different situation and settings. This can also be loosely defined as *understanding of principles governing a domain and interrela-*

Md.S. Aftanasar (✉) · N. Mohamad
School of Electrical and Electronic Engineering, Transkrian Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, P. Pinang, Malaysia
e-mail: afta@usm.my

N. Mohamad
e-mail: norizah@usm.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_16

tion between units of knowledge in a domain (Rittle-Johnson 2006). Current practice in the fields of engineering shows that development of problem-solving skill is enhanced through worked examples, lectures, and laboratory work. These computational skills, according to several scholars (Millians 2011; Fuchs et al. 2008), deal with competency and intelligence in executing standard process and procedures in given problem-solving activity. However, despite getting good results in summative examinations, students have little underlying concrete understanding in what they have learned. This is a concern for most engineering educators since complex engineering problem demands association, experience, and intuition, which are internal, and ability to consolidate information, facts, and concepts. This process is, as eloquently described by Montfort et al. (2009), is purely a top-down cognitive process and is related to experience of "usefulness" or "transferability" in applying the knowledge outside the context in which it was learned. In short, it is envisaged that there is a need to investigate whether the same is true for the delivery of RFnM course. And if so, what technical intervention can be made?

Currently in RFnM elective course at the School of Electrical and Electronic Engineering, USM, students were required to submit written laboratory report that incorporates fundamental theories, methods, and findings, in explaining the learned concept. It was found that elaboration in most concepts was usually presented as "dry" mathematical formula, and most explanations were excerpts from articles in the Internet, rather than coming up with their own personalize way of explanation. This triggers a question whether students really embody the meaning behind learned concept from the course.

In this paper, we have experimented on students' conceptual understanding based on their experience in an improved assessment method for RFnM course, taught in turn-based (sharing) teaching. Students were requested to participate in two assessments carried out in Part 1 (weeks 1–7) and Part 2 (weeks 8–14). Note that this study only reflects improvements that were made in Part 1. Some instructional change and online learning platform were implemented in Part 1 to support teaching and learning using Web-based learning platform, technology-based hands-on experience, verbal and written communication feedback, cooperative learning, and interactive lecture. The underlying research questions are as follows:

1. What is the student performance on conceptual understanding based on content, context, and connectivity of any given concept after teaching and learning experience from Part 1?
2. What improvement as a result from revised strategy from perspective of student's academic achievements, peer-review performance, and time factor, based from conceptual understanding assessment from Part 1?

2 Literature Review

Systematic learning, as being put by Piaget (1963), is a developing process in which people progress through subsequent stages. Each stage consists of new perspective with new info, methods, and processes. As learners ascend from novice to expert user of particular knowledge, their views and horizon become broader and flexible. In addition, Bruner (1960), from a constructivist perspective, describes learning as a complicated process where learners make progressive associations and constantly learn and unlearn, or readjusting, to understand the relationship between concepts. In essence, learning is an experience and processes where a learner continuously categorizes, structures, associates, refers, infers, applies, and makes improvisation (Venters et al. 2014; Venters and McNair 2010), and concludes based on the previously learned knowledge, experience, and perception (Biggs 1996) to aid learner construct complicated understanding.

In this paper, conceptual learning theoretical framework centers on three major engineering education researches, i.e., interactive learning with feedback, using technology in learning, and the use of representational fluency in developing experts. This follows an article written by Venters and McNair (2010) who propose attention to be given on learning approach as an intervention to improve conceptual knowledge in statics. Another paper based on statics course by Venters et al. (2014) proposes a pedagogy change in learning statics by adding "process problem" that requires student to write about their assignment, rather than simply using mathematical formula to work out the problem. The outcome from the pedagogy change measures students' reflection and metacognition. It classifies a student on how a content being expressed between purely on mathematical approach or conceptual approach.

In conceptual-based learning, one of the challenges in correct interpretation of any concepts is a deeply held belief on any particular prior information or knowledge, which interferes with a new one. This is known as misconception. Conceptual understanding and misconceptions can be improved and corrected, and then enhanced when students experience learning through various interactions and learning strategy, requiring pedagogy change. For example, an article by Krause et al. (2014) implements several pedagogies, namely interactive with hands-on activity (concept sketching), interactive with sorting activity (concept–context sort with no hands-on), interactive discussion only, and passive (lecture). Based on their findings, interactive with hands-on activity scored the highest in conceptual change based on Materials Concept Inventory (MCI).

There are major researches on the use of technology to improve learning experience. In a work by Chi (2005), computer simulation is coupled with theory-based curricular to aid students in "unobservable" dynamics at miniscule scale. With the aid of technology and online access, students actively construct their own meaning through interaction with content (Streveler et al. 2008). In a model implemented by Krause et al. (2014), Web-designed formative feedback using Web-enabled tools has been used to facilitate strategies, practices, and assessment.

Web-enabled resources were developed to aid instructor and students in improved teaching and learning environment. These tools and outcomes of the study show that multi-interactive pedagogy has potential in improving learning motivation and attitude in learning (Krause et al. 2014). Also, studies have confirmed that improved motivation and self-esteem explain why student improved in self-regulate their studies and performance (Nicol and MacFarlane-Dick 2006).

Improving conceptual understanding is done through multiple representation and connectivity between the concept in-focus and other related concepts. Lesh Transitional Model (Lesh and Doerr 2003) has been implemented in mathematical learning and posits elementary ideas which can be represented in several modes. The model suggests that understanding a concept is expressed in the ability to translate between modes. In providing different experiences on important concepts, Lesh Transitional model is widely implemented in investigating representation and representational fluency for understanding concepts. This model emphasizes on interrelation between (i) contextual, (ii) symbolic, (iii) language (spoken and written), (iv) pictorial, and (v) concrete representation. Understanding the concepts depends on the student's ability to translate between these components. With greater complexity in interconnection between established nodes or representation, students exhibits greater flexibility, meaning and embodiment of the knowledge Krause et al. (2014).

Verbal medium containing sign and symbol represents communication between a source and a receiver carrying communication-bearing information. According to Krause et al. (2014), in quantifying conceptual understanding, an inquiry technique was implemented in this research. Participants were asked to elaborate on a given topic, and data collected were analyzed in various properties: semanticity, generativity (or productivity), and displacement. Semantics are signals that convey meaning. Generativity is the capability to generate numbers of messages with limited number of symbol. Displacement conveys messages in relation to spatial position, e.g., time, space, and imagination.

3 Methodology

3.1 Research Design

For the assessment, the topic requires students to elaborate important concepts related to microstrip transmission line. Important concepts related to microstrip were taken from Pozar (2009), which serves as a primary book source of reference during the course. Hands-on practical implementation followed by lecture sessions, with additional supporting info and exercises, made available online. By week 7, students have been presented with several concepts that were categorized into: transmission line, electrical characteristics, and physical characteristics.

In identifying relevant category for the assessment, student responses from a mock (in week 5) interview was rated, coded, categorized, and compiled onto a rubric. The emerging category was based on the framework of novice–expert scale (Dreyfus and Dreyfus 1980) used for skill acquisition and definition of conceptual understanding put forth by National Assessment of Educational Progress (2003) specifically measuring mathematical abilities at school. The three emerging categories based on open coding and implemented in scoring rubric are as follows:

1. ***Contextual Proficiency***—spoken technical terms that are embedded in spoken statement of the concept based on learned experience from the learning outcome;
2. ***Experiential Fluidity***—the changes in the flow of conversation covering various aspects of information and concepts based on learned experience from the learning outcome;
3. ***Depth***—complexity of explanation with deep holistic understanding on the concept based on learned experience from the learning outcome.

3.2 Participants

Data used were compiled from third-year students taking EEE322—Introduction to RF and Microwave Engineering Course. This is an elective course under telecommunication group with initial total participants of 20 students. However, only 17 students successfully completed both interview sessions.

3.3 Assessment Instrument

A free Web-based tool, voicethread.com, was used to record the event. Each student was given three minutes: two minutes for preparation and one minute for recording. Consequently, each recorded session was transcribed and coded, and marks were given using a scoring rubric developed from emerging three categories. Prior to the 1st assessment, students were given a mock topic and some pointers for their preparation.

3.4 Data Collection and Procedure

Request for participation and consent was made at the beginning of the course. The duration was 14 weeks with a mock interview given in week 5. Assessments were conducted in week 7 and week 14. Student's academic performance was from cumulative grade point average, CGPA. Peer-review evaluation for each student

was conducted in week 7. Peer-review scoring rubric was used by the students to assess each member of their team and is confidential.

4 Data Analysis and Discussion

For research question 1, data from 17 students (N_{total} = 17, labeled RF01–RF17.) indicated most students have shown improvements from Assessment 1 to Assessment 2 based on increased in overall marks, as depicted in Fig. 1. The overall average marks rose from 57 to 73 % over the duration of weeks 7–14. Also in 2nd assessment, the result has revealed that 4 students have managed to gain 100 % full points. Excerpt from 3 students (RF07, RF03, and RF01) show decrease of 20–30 %, and one student reflects no changes. Figure 2 shows the difference between marks from Assessments 2 and 1 for all responses. As previously stated, most students show positive improvements from the overall result with 4 students attaining close to 40 % increase in marks, the largest leap by far. From the inscription, inference seems to show preference in relationship found during hands-on laboratory sessions, indicating similar findings by Krause et al. (2014).

For research question 2, each perspective is dealt, respectively. The relationship between result from performance (mark difference between Assessments 2 and 1) and participants' CGPA was measured. Pearson's correlation coefficient (r) reveals extremely weak association (r = +0.1) between Assessment 1 and CGPA. However, association between performance and CPGA shows strong relationship (r = +0.5), indicating higher chances that students improve conceptual understanding after continuously improving their knowledge with experience. The

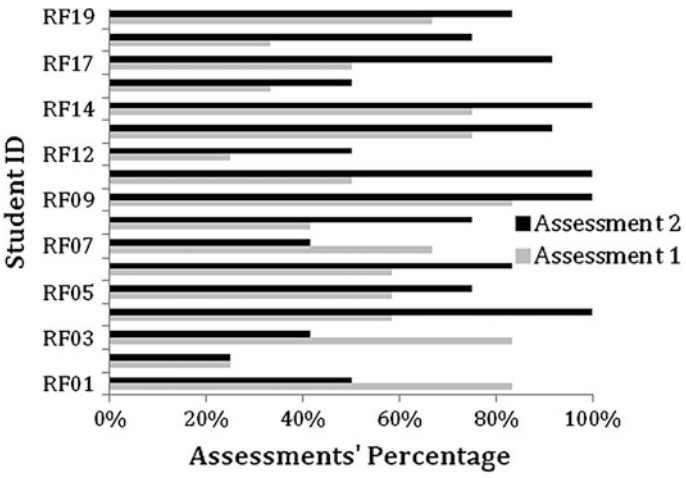


Fig. 1 Performance between Assessments 1 and 2 based on student's ID

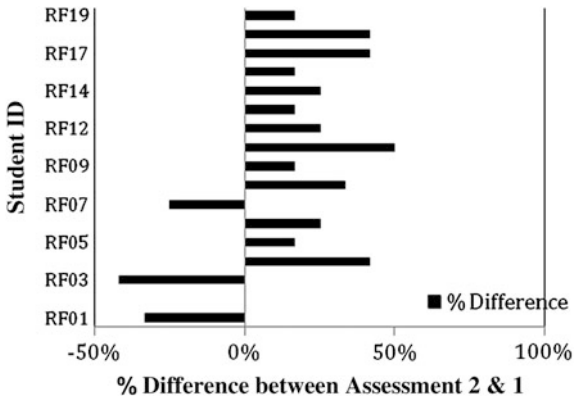


Fig. 2 Percentage difference in marks between Assessments 2 and 1

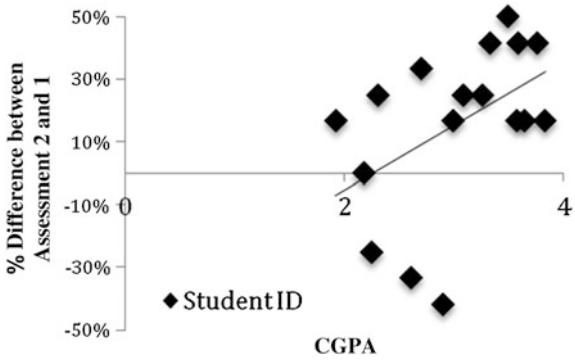


Fig. 3 Trendline showing strong significant relationship (r = +0.5) between CGPA and % difference between Assessments 2 and 1

trendline in Fig. 3 show most improvements having some positive degree of association with their CGPA. Evidently, students with CGPA 3.00++ ($N_{>3.00}$ = 10) show more significant improvement than others. Nevertheless, those in the range between CGPA 1.75 and 3.00 ($N_{<3.00}$ = 7) display varying results with only half of the candidate showing improvements or unchanged. When asked why this happen, students have not prepared themselves and waited for week 15 (study week) to review their learning materials. Also, they found difficulty and uncertainty in putting their prior experience into context.

Further investigation on correlation coefficient on each measured categories for conceptual understanding in Assessments 1 and 2, according to CGPA and peer review, is shown in Table 1. There seems to be no noteworthy relationship between most variable in Assessment 1. However, relationship between contextual proficiency for peer review indicates weak correlation (r = +3), and signifying student's degree of involvement in team may improve their ability to explain better. Overall marks from Assessment 2 seem to show positive significance (r = +0.5) in relation to CGPA. This is readily shown in Fig. 4. Experiential fluency and depth seem to

Table 1 Person correlation coefficient (r) performed for Assessments 1 and 2 based on measured criteria, in relation to CGPA and peer-review assessment

Categories in conceptual understanding	Assessment 1 (r)		Assessment 2 (r)	
	CGPA	Peer review	CGPA	Peer review
Contextual proficiency	0.0	+0.3	+0.5	0.1
Experiential fluency	−0.3	−0.1	+0.3	0.0
Depth	−0.1	0.0	+0.3	0.1

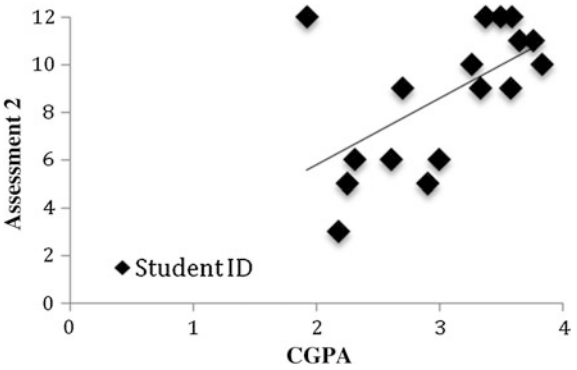


Fig. 4 Trendline showing strong significant relationship (r = +0.5) between CGPA and overall marks in Assessment 2

show weak significance and were based on their ability to relate, explore, and associate learned facts and knowledge to different perspective and implementation.

Peer-review results conversely show no significant relationship with all criteria. Teamwork attitude, which specifies student functionality, leadership, personality, and engagement during cooperative learning, may not be have utmost influence in improving conceptual understanding.

5 Conclusion

A new method of assessment was administered for RFnM course. From both interviews, preference by students to connect concepts related to microstrip from Part 1 may attribute to instructional variation and hands-on activity. This might indicate that conceptual understanding development may skew toward favorable and memorable teaching and learning strategy, particularly in relation to hands-on experience, followed by facilitation and support given in Part 1. In the second interview in Part 2, frequent inferences made to contents in Part 1, compared to Part 2 during the interview may have predominantly expresses students deeper appreciation, better retention and understanding. Additionally, improvement in all categories seems to indicate better connectivity as student elevates to expert user of the knowledge. Also, this was mainly found in student with good academic background.

Conceptual understanding has some relationship with students' individual academic achievement, in all aspect of assessment, i.e., breadth, depth, and lexicon. And conceptual understanding seems to have less impact against student's behavior and function in cooperative learning and teamwork activity, as seen from the peer-review assessment.

Future works includes improvement in feedback system either in-class or online that further enhances conceptual understanding and introduce concept mapping as a tool to improve conceptual understanding. This can provide avenue of research that will address response from different learning capabilities. In addition, duration for the interview should increase to provide better insight in all categories. Finally, future data will include final examination for cross-referencing in providing variation in analysis.

References

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*(3), 347–364.

Bruner, J. (1960). The Process of Education. Cambridge, MA: Harvard University Press.

Chi, M. (2005). Commonsense conception of emergent processes: Why some misconceptions are robust. *Journal of Learning Sciences, 14*(2), 161–199.

Dreyfus, S. E., & Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition* (No. ORC-80-2). California Univ Berkeley Operations Research Center.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Lambert, W., Stuebing, K., & Fletcher, J. M. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition. *Journal of Educational Psychology, 100*(1), 30–47.

Krause, S., Sean, M., Chan, C., Waters, C., Carberry, A., & Koretsky, M. (2014). Web-enabled formative feedback and learning resources for enhancing student attitude, achievement, and persistence. In *Frontiers in Education Conference (FIE)* (pp. 1–8). Madrid: IEEE.

Lesh, R. A., & Doerr, H. M. (2003). *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching*. Routledge.

Millians, M. (2011). Computational skills. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of child behavior and development* (p. 396). US: Springer.

Montfort, D., Brown, S., & Pollock, D. (2009). An investigation of students' conceptual understanding in related sophomore and graduate-level engineering and Mechanics courses. *Journal of Engineering Education, 98*(2), 111–129.

National Assessment of Educational Progress (NAEP). (2003). What does the NAEP mathematics assessment measure? Retrieved on May 3rd, 2015, from http://nces.ed.gov/nationsreportcard/mathematics/abilities.asp

Nicol, N. J., & MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218.

Piaget, J. (1963). *The origins of intelligence in children*. New York: Norton.

Pozar, D. M. (2009). *Microwave engineering* (4th edn.). London: Wiley.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development, 77*(1), 1–15.

Streveler, R. A., Litzinger, T. A., Miller, R. L., & Steif, P. S. (2008). learning conceptual knowledge in the engineering sciences: Overview and future research directions. *Journal of Engineering Education, 97*(3), 279–297.

Venters, C., & McNair, L. (2010). Learning statics: A cognitive Approach. In *ASEE Southeast Section Conference*, pp. 1–10.

Venters, C., McNair, L., & Parerti, M. (2014). Writing and conceptual knowledge in Statics: Does learning approach matter? In *IEEE Frontiers of Education Conference (FIE)*, pp. 1–8.

Deploying Right Assessment for Programme Outcomes (POs) in Enhancing Curriculum Structure Development

Abdul Rahman Mohd Yusoff and Mohammad Syuhaimi Ab-Rahman

Abstract This paper discusses the deployment of the right assessment methods and measurements for institutional programme outcomes (POs). Apart from applying the twelve (12) standard engineering student attainments, possible new attributes to be added with distinct niche are considered as advantage. By strengthening existing PEO pillars and added value for PO attainment, the resulting synergy can enhance university's mission and vision as well as fulfil stakeholders' needs. By designing and giving the correct interpretation of PO, it will enhance graduate values upon completion of their engineering degree programmes. Correct assessment methods must be implemented and followed by the continuous quality improvement (CQI) process. Selected performance indicators are used to reflect PO assessment reliability and effective method of measuring. Without addressing this correctly, the overall quality of graduates produced in engineering discipline will be questionable. By deploying the right assessment measurements and continuous quality improvement (CQI), academic curriculum structure is able to reach new height in becoming a new benchmark for other institutions to follow.

Keywords Curriculum structure · Programme outcomes (POs) · Graduate attainment · Continuous quality improvement (CQI)

1 Introduction

Institutions of higher learning (IHL) offering engineering degree programme are required to implement outcome-based education (OBE) in compliance with accreditation requirements. The institution should decide what domains to be

A.R.M. Yusoff (✉) · M.S. Ab-Rahman
Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
e-mail: abdrahman.mdyusoff@yahoo.com

M.S. Ab-Rahman
e-mail: syuhaimi1977@gmail.com

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_17

instilled in their engineering graduates. The programme offered to students can be customized, both in structures and in academic curricula in order to achieve those proficiencies. The OBE design process involves restructuring of curriculum, carrying out assessment as well as continuously monitoring the achievement result. The result must reflect the actual attainment level of student learning rather than just accumulation of credited courses. One of the most important activities for the university is the development of established curriculum in line with the national and international standards. The key component of curriculum strength lies on the achievement of programme objectives or students' learning outcomes. The illustration of an example of OBE system (Gamboa et al. 2013) clearly indicates that institutional vision and mission as well as programme educational objectives are closely related with student programme outcome (PO) attainment. The strength of this OBE system lies in the main supporting pillars of PEO. However, these pillars are further strengthened by programme outcomes (POs) and course outcomes (COs) that are continuously improved and upgraded within the system. The aim of this paper was to show the importance of getting right PO for the university to support this pillars (PEO) and institutional goals by providing right knowledge, skills and attributes for their graduates.

2 Review of Programme Outcomes

The process involved successfully delivering desired outcomes from Bloom's taxonomy domains (Bloom 1956). There are three main domains which are cognitive, affective and psychomotor attributes. Cognitive domain revolves around knowledge, comprehension and critical thinking order for any particular learning items covered. The understanding from what they were taught makes it important to know how well students digest learning process. On the other hand, the affective domain describes the way people react emotionally and their abilities to feel another living thing's pain or joy. The human aspect in dealing with subject matters significantly portrays how well a student grasps any given problem. Finally, the psychomotor domain describes the ability to physically manipulate a tool or instrument like a tool or equipment. This situation is similar to instil proper skills on student to make him excel in their field of studies. These set of characters are considered the minimum standard stipulated for engineering students under the Washington Accords (Hanrahan 2011). These programme outcomes should be evaluated throughout the educational delivery process, assessment and consequently continual quality improvement (CQI) stages. In order to ensure that the right domains are securely intact, the institution should focus on its programme and course level. At the end, the cycle of continuous improvement will transform and strengthen the institutional curriculum structure.

There are many ways and techniques how education can be done through delivery methods such as lecturing, tutoring and other classroom teaching and learning techniques. Each delivery method has its own characteristics in reflecting

the outcomes to assist students in meeting PO. Several assessment methods were applied to monitor the progress of student in attaining these outcomes. One of the most commonly used methods is applying formative and summative assessment techniques. Formative assessment represents a programme to build up the potential of the students in preparing them for their final performance or achievement. Meanwhile, summative assessment represents the actual achievement to be desired upon completion of the course output programme. In other words, the progress of the student can be monitored throughout their course of studies. An immediate action can be taken if flaws are detected in the process of bringing the students level to meet PO target. Similar to programme educational objective (PEO) statement, the programme outcomes (POs) should also be set in a **s**ystematic, **m**easurable, **a**chievable, **r**ealistic and in **t**imely manner. These s.m.a.r.t abbreviations ensure the strength of the university graduates produced able to fulfil industries interest and their current needs. POs can be considered as brick supporters for building's pillars (PEO) supporting academic curriculum structure with right knowledge, skills and attributes. The main attributes identified for POs include technical competency, soft skills competencies and some additional elements of engineering aspects on complexity, sustainability and safety requirements raised by the accreditation standard body. By applying a consolidation strategy, this will result in good curriculum structure base on proper measuring indicators. It was developed in a structured manner whereby each stage has a specific purpose such as strengthening knowledge base in technical, engineering, multidisciplinary and the soft skills. The outcomes have to undergo process establishment, assessment and implementation phases. Thus, the outcomes are reviewed and evaluated against the intended plan. These stringent target requirements triggered the institution and faculty to envisage what graduates should have achieved at the end of their study programme graduation.

3 Processes of Determining PO

There are several approaches proposed in developing PO statements. The structure used involves Bloom's taxonomy approach in distributing cognitive domains according to their learning level. The scheme used by UKM's Department of Electrical, Electronics and Systems Engineering (EESE) will be extended to other engineering departments as references. All the 6 levels of distribution serve as a guideline in formulating student assessment tool for the lecturers in their teaching course evaluation and PO assessment. The department also had set the distribution of cognitive domain from Year 1 until Year 4 as listed in Table 1.

The EESE Department has taken the initiative to distribute proportionally the cognitive domain throughout the 4-year degree programme with ratios as illustrated in Fig. 1. The evaluation for lower level knowledge and understanding can measure up to 30 % for Year 1. The cognitive domain for middle level is almost equally distributed in between Year 2 and Year 3. As engineering students gain more experience and master their learning stages, the ratio for upper level of critical

Table 1 Bloom's taxonomy on cognitive

Taxonomy	Cognitive stage	Illustration
Level 1	Knowledge	Observation and recall of information
Level 2	Comprehension	Demonstrative understanding of facts and ideas
Level 3	Application	Solve problems using required skills or knowledge
Level 4	Analysis	Distinguish between facts and inferences
Level 5	Synthesis	Relate knowledge from several areas
Level 6	Evaluation	Compare and discriminate between ideas

cognitive thinking increases. By Year 4, the critical synthesis and evaluation stage passed through 50 % level, reducing lower cognitive level to minimum percentage possible. As for the psychomotor and affective domains, the measurement can be evaluated using other non-formative tools such as rubrics, peer evaluation and satisfaction survey as this could be qualitative measurement in nature.

There are twelve (12) stated POs suggested as reference in the EAC Manual (2012) whereby graduating engineering students is expected to attain upon completing their degree as listed in Table 2.

It is the responsibility of each engineering department running the programmes to develop their own departmental PO statements. Each department definitely is looking for specific PO needs and specialization in their respected fields being in chemical, civil, electrical and mechanical or other engineering fields. Basically, PO1–PO5 represents more on technical skills utilizing cognitive and psychomotor domains. Meanwhile, PO6 and above represent more towards soft skills with cognitive and affective domains. In addition to these 12 POs, there are also others incumbent traits involving issues on complexity, sustainability and safety to further enhanced student learning outcomes. Selected POs must be measured for its effectiveness by applying some performance indicators for each detail of criteria set for each PO statements.

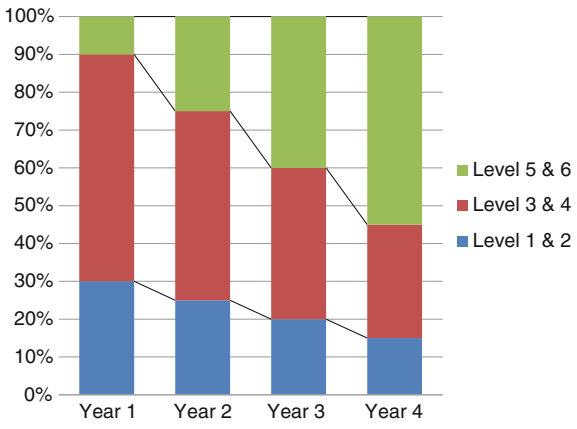


Fig. 1 Cognitive level of distribution for 4-year engineering programme

Table 2 Minimum programme outcomes as suggested in EAC Manual

Ref.	Programme outcomes	Description
PO1	Engineering knowledge	Apply knowledge of mathematics, science, engineering fundamentals and an engineering specialization to the solution of complex engineering problems (cognitive and psychomotor)
PO2	Problem analysis	Identify, formulate, research literature and analyse complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences and engineering sciences (cognitive and psychomotor)
PO3	Design/development of solutions	Design solutions for complex engineering problems and design systems, components or processes that meet specified needs with appropriate consideration for public health and safety, cultural, societal and environmental considerations (cognitive and psychomotor)
PO4	Investigation	Conduct investigation into complex problems using research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of information to provide valid conclusions (cognitive and psychomotor)
PO5	Modern tool usage	Create, select and apply appropriate techniques, resources and modern engineering and IT tools, including prediction and modelling, to complex engineering activities, with an understanding of the limitations (cognitive and psychomotor)
PO6	The engineer and society	Apply reasoning informed by contextual knowledge to assess societal, health, safety legal and cultural issues and the consequent responsibilities relevant to professional engineering practice (cognitive and affective)
PO7	Environment and sustainability	Understand the impact of professional engineering solutions in societal and environmental contexts and demonstrate knowledge of and need for sustainable development (cognitive and affective)
PO7	Ethics	Apply ethical principles and commit to professional ethics and responsibilities and norms of engineering practice (cognitive and affective)
PO9	Communication	Communicate effectively on complex engineering activities with the engineering community and with society at large, such as being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions (cognitive and psychomotor)
PO10	Individual and teamwork	Function effectively as an individual, and as a member or leader in diverse teams and in multidisciplinary settings (affective)
PO11	Lifelong learning	Recognize the need for, and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change (cognitive and affective)
PO12	Project management and finance	Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments (affective)

Source EAC Manual (2012)

4 PO Formation Process

The process of PO formation starts with the agreement which was reached after lengthy discussions with head of faculty and management approval. The programme outcomes should be in line with existing current programme curriculum and stakeholder's demands and cordially meet recognized accreditation standard. The final draft is meant for further discussion and to be delivered for reviewing within all interested parties including employers, alumni, IAP, students and parents. All the feedback received will then be analysed and fine-tuned to show the involvement of the stakeholders. The board of committee should decide on the importance and emergent needs for any changes. If the decision committee satisfied with the justified reasons, then new POs will be published, distributed and made known to all students through reachable instructional media and publication channels. The detailed illustration of the flow is shown in Fig. 2. The formation of PO is crucial as student attainments should follow the current needs of the stakeholders involved and make the process dynamic in nature. The inputs and improvement on POs throughout the reviewing stages are what makes programme outcomes more current and relevant.

5 PO Mapping

Most courses offered in the department can be mapped to their PO and standardized to all engineering degree programmes (see Fig. 3). This aims to simplify the process of monitoring plans and empowering students with the expected outcomes (which is also known as graduate attribute). Each PO will be defined where indicators will be put before identifying courses most suitable measure. Courses taught might have more than one PO during teaching and learning activities. The course assessment will be regularly reviewed to ensure that students reach the level of required domains. The faculty can hold special workshops to define each domain exist in the PO and explain its performance indicator in accordance with their respective niche. This is considered one of the most effective methods of measurement for each relevant course.

6 PO Measurements

The process of measuring PO relationship linkages is meant to comply with to the accreditation. To explain the concept, one programme outcomes sample for the Department of Electrical, Electronics and System Engineering (PO1) regarding engineering knowledge is chosen (see Table 3). The details and criteria needs for that particular PO1 are clearly elaborated. This is to show the significant of having PO details in evaluating the outcomes.

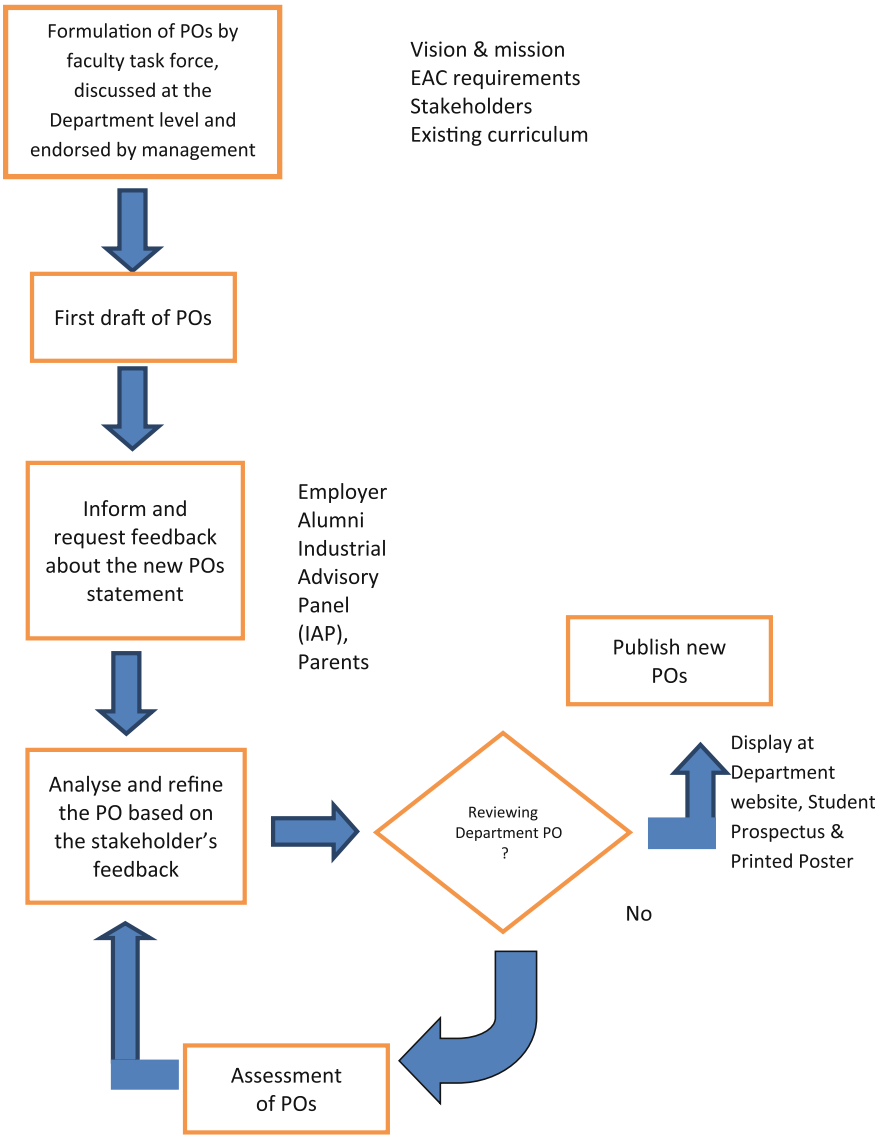


Fig. 2 Process of establishment and revision of the POs used in EESE, UKM

PO1 represents the first set of attributes for engineering graduates to have a good understanding of engineering knowledge. UKM EESE Department needs to guide graduates to achieve this first programme outcomes (PO1) by elaborating the expected content of outcomes related to engineering fundamental knowledge and the problem in complex engineering matters. PO1 is considered part of strong indicators in student's attainment. The process of attaining programme outcomes

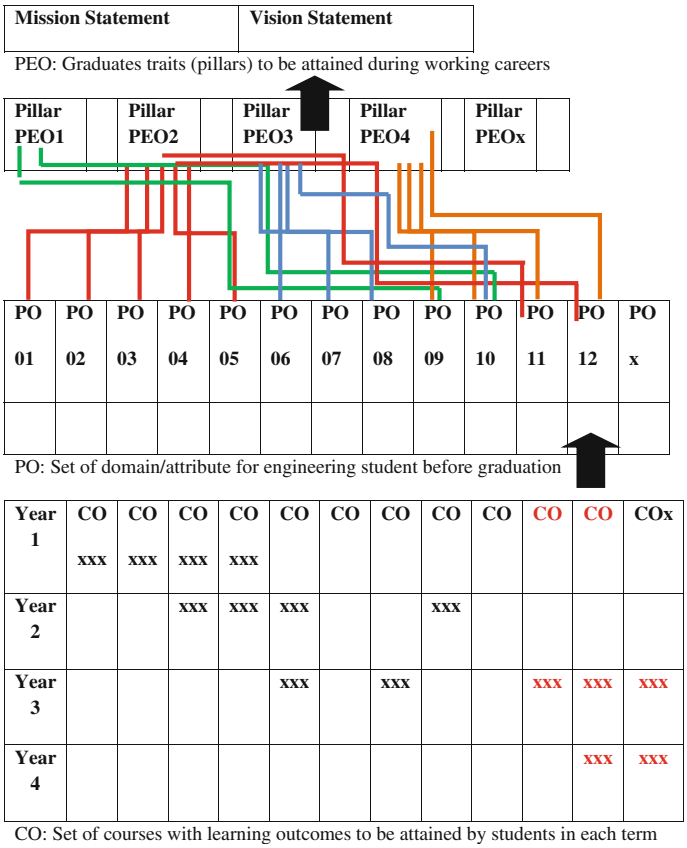


Fig. 3 Mapping and linkage of institution mission and vision to PEO–PO–CO

can be broken down by assigning what types of delivery methods and assessment tools used to measure the PO1 attainment result. Delivery method of teaching and learning such as lecturing, laboratory, problem-based learning (PBL) and group assignment projects was put in place. PO1 learning outcomes and achievement can be assessed through examinations, quizzes, presentation, project papers and other evaluating their soft skills competencies. Furthermore, the type of questions and answers greatly depends on the type of domain and level applied in their learning process. The higher the level, the more difficult types of questions are set. By using performance indicators in measuring student's achievement, the process will ensure the effectiveness of learning process and reflected the actual performance of PO. The performance indicators used can be in percentage (%) average score point or targeted CGPA level per cohort or class. The results directly reflect to programme educational objective (PEO) as well as the mission and vision of the institution. In most cases, many institutions are trying to match PO to their PEOs.

Table 3 Details of programme outcomes on teaching delivery, assessment, performance indicators and target achievement for PO1

PO1	Engineering knowledge	Apply knowledge of mathematics, science, engineering fundamentals and engineering specialization to the solution of complex engineering problems (for electrical electronics and system engineering)
Delivery methods	Teaching, Laboratory, Project design Problem-based learning (PBL)	Assessment methods: examination, quizzes, laboratory report, project, presentation
Bloom's/level Taxonomy C—cognitive P—psychomotor A—affective	Target setting	Performance indicators
PO1C1 PO1P1 PO1C3 PO1P4 PO1C4 PO1P5	100 % students must achieve average Score point of 2.5/5.0 or CGPA 2.0/4.0	1. Acquire fundamental knowledge of maths and science to solve engineering problem 2. Formulate analytical models using fundamental knowledge 3. Analyse complex engineering problems

Source UKM EESE Department Self-Assessment Report (SAR) 2013

PEO is linked to the programme outcomes (POs) by applying matrix relationship where desired attributes are mentioned clearly. Should there be any non-attainable desired attributes, the institution might consolidate in the next PO review and updated the PEOs. This also applies to the rest of PO2 till PO12 including those that are considered extra domains (POx) aligned to institutional objectives. The similar approaches are taken to ensure every engineering courses offered in the department are linked to their programme outcomes. The linkage is important to ensure desired elements at course outcome (CO) level can be measured to justified PO achievements by the end of each semester (Masni-Azian et al. 2014).

CO–PO mapping confirms the courses offered met the prescribe outcomes and considered part of PO attainment evidences. It is therefore important to monitor the students' progress throughout the 4-year engineering programme. The reason is to ensure that all stated department POs can be attained by each student before their graduation time. In order to improve student PO attainment, a newly developed performance criteria has been set to monitor whether changes to the current educational delivery and assessment methods are required (Amirulddin et al. 2009). In ensuring a more reflected results be obtained, course outcomes can be further break down into smaller performance indicators for measuring CO–PO achievement (Shanableh 2014). In another instances, selected courses such as final-year project (FYP) or integrated project are set as a benchmarking in measuring POs achievement as illustrated in (Arshad et al. 2012). These selected courses are used to ensure

students are capable of fulfilling most engineering desired attributes before completing their studies.

A structured curriculum makes it clear for students to achieve each programme outcomes by using performance indicators. By measuring this achievement target, the outcomes can determine the quality of academic programmes offered. The curriculum was designed in a structured way with specific purposes such as empowering knowledge of engineering fundamentals, core components, specialization and related engineering disciplines. It is transparently easy to analyse the achievement level by measuring, evaluating and assessing the PO for reflecting student's real attainment.

Generally, student PO attainment can be measured in two different methods either direct or indirect. Direct measurement for graduates and alumni can be done through surveys using indirect questions. The questionnaires are more details for respondent to elaborate. This gives valuable inputs, and the results are more subjective. Meanwhile, indirect measurement uses the mapping technique for PEO–PO–CO linkages. The matrix shows how PO attainment directly determined PEO achievement. Similarly, the CO achievement also reflects the contribution to PO attainment. This method is also beneficial in identifying PO domains that directly contribute to programme educational objectives (PEOs) achievement (Nordin et al. 2014). Both measurement methods are equally important in determining the actual achievement on alumni (direct method) and identifying any unattained domains by mapping them directly to PEO (indirect method). Both measurements complement each other and able to form a reliable measurement and improving overall educational programme through PEO achievements. PO assessment has similar ways to show its effectiveness by monitoring the performance of student's course learning outcomes. Course outcomes (COs) can be mapped to PO which contributes to PO achievement by applying some performance indicators. Others application includes using direct/indirect assessment techniques for measuring PO outcomes (Nordin et al. 2014) and in some rare cases solely applying CQI on PEO and PO (Abdullah et al. 2008). Furthermore, in many cases, the changes occurring in PO statements are derived mainly from external input sources from the industries and other interested parties (Mutalib et al. 2012).

7 PO Selection Process

PO can be derived from several sources either internally and externally. In developing PO, it is essentially important to know the university internal needs which aligned with its vision and mission. Since the autonomy and final choices lie on decision by management, the intended PEO pillars should be discussed by the internal committee in charge of fostering these needs. There are also other certification requirements to be fulfilled such as Malaysian Qualifications Agency (MQA) and International Organizational for Standardization (ISO) requirements. The numbers of input vary according to the numbers of targeted set by the

institution. By having internal committees, the objectives, direction and features of PEO can be determined more precisely. The management team set the niche for the institution after processing inputs from all involved stakeholders. In some instances, PEO appears to have domain that is not found in PO minimum attributes which is considered as benefit to the institution. The 'niche' can be added as extraordinary features that differentiate the institution with the new pillars presence. Therefore, it is the university's decision on what types of graduate's attainment the institution would like to offer in satisfying their stakeholders.

There are eight main engineering courses that are considered as best benchmarking for ensuring students able to reach a minimum standard set. Criteria's contents within those courses were used to cater the strong relationship of student attainment. These courses include engineering sciences and maths skills, engineering application courses, integrated project, professional engineering practices and ethics, laboratory works, industrial training, technical talks and visit industries, and final-year project. All these listed engineering courses cover most of the PO domains matching the minimum requirement for engineering students to graduate. As a guideline, PEO is considered strong pillars to enhance the knowledge, skill and traits while practising engineering profession. Meanwhile, accumulation of POs domain within students will prepare them with minimum requirements to step into future challenging engineering careers and practising.

The process of selecting PO pillars for students takes a proper strategic planning within their study phase starting from beginning of Year 1 until Year 4. Each stage needs to be structured accordingly in accumulating term of allocating cognitive education domains required to be achieved. For example in Year 1, the level of knowledge is to develop understanding more on basic fundamental of engineering level. As times passed, the level should be geared up towards achieving high level of cognitive learning level in Year 4. This is due to the maturity of students able to grasp higher order of thinking within their study scope. The PO development and selection process with its details structure content relationship is illustrated in Table 4.

8 Result and Discussion

There are a few factors to be considered in ensuring the right PO deployment is successfully applied. Following are some considerations to be given in established PO and selecting the right method and measurement in sustaining outcome-based education (OBE). These brief guidelines can be used to develop PO starting from defining the scope covered until the process of measuring PO effectiveness at student's level before entering engineering job market. This guideline can ensure the smoothness of designing and selecting PO in supporting OBE implementation. Followings are the gist of deploying the right assessment for programme outcomes (PO) for IHL considerations.

Table 4 PO development table and curriculum content selection in EESE

Needs of PEO establishment	PEO obtained	PO selection	Combined university own PO and 12 standard suggested POs	PO matrix and structure curriculum content	Determine studies priorities
Four main PEO for UKM Engineering Department involving PO8, 10, 11, POx • PEO 1 • PEO 2 • PEO 3 • PEO 4	List of additional PEO obtained from external and internal views (PEO unique for each university)	PO1 till PO12 POx (additional)	PO1—Engineering knowledge PO2—Problem analysis PO3—Design/dev of solution PO4—Investigation PO5—Modern tool usage PO6—The engineer and society PO7—Environment and sustainability PO8—Ethics and professionalism (Pillar PEO1) PO9—Communication PO10—Individual leadership/teamwork (Pillar PEO2) PO11—Lifelong learning (Pillar PEO3) PO12—Project management and finance POx—Malaysia aspiration (Pillar PEO4)	8 main courses each engineering programme offered 1. Engineering science, math skills 2. Engineering application 3. Integrated project 4. Professional eng. practice 5. Laboratory work 6. Industrial training 7. Exposure to engineering practice (talk and visit) 8. Final-year project	Mapping can be assessed based on the importance of PO (pillars and attributes)

- What is the institutional PO statement? Is there any linkage between the PEO and mission and vision (M&V) and/or strengthening the attributes with the course outcomes (COs)?
- How does the institution define and elaborate details for each of the 12 suggested the POs? (This includes the 'performance indicators' for measuring them.)
- How does the selection of PO involve the stakeholders in the process of formulation, assessment and continuous quality improvement (CQI)?
- What type of delivery model is used for PO? How to do the 'mapping' of courses be done in relation to PO?
- What PO measurement model is used? How do they make the measurement? Is measurement method use 'direct' or 'indirect'? What type of tools and measuring frequency is required?
- What PO assessment model for 'evaluation' is used? How to check the effectiveness and to confirm the student's PO achievement, especially for 'graduating students'?
- What improvements do PO contribute? How do institution make or use CQI 'results' assessment holistically and sustainably?
- Finally, how PO improvements contribute to achieve the PEO, mission and vision of the institution?

PO attainment for engineering students is seen as an important stage for preparing them into engineering careers. Gaining a sound background knowledge, skill and right attributes is what is required by students in completing the educational programme. Hence, by sharing UKM's EESE experiences, every institution of higher learning need not only manages to offer good programme but also needs to prepare students with competencies in both technical and soft skills. One of the best measurement methods used by employees to recruit fresh engineers is by looking through holistically not only to knowledge gain throughout their academic years, but also on their competencies skills acquired and commendable working attitudes.

9 Conclusion

What the paper intends to achieve is to reveal the important information on POs so it can be shared within similar engineering programmes offered. The process of selecting and establishing PO is a delicate one and requires several due considerations. Failure to define and elaborate the needs of stakeholders can have a backlash in translating engineering programme that meets its mission and vision. POs should be measured using criteria and performance indicators to reflect its effectiveness in meeting programme outcomes. There should be clear evidence that the objectives and outcome target are achievable. Should there be any non-performing criteria's set, it requires immediate action through the continuous quality improvement. By

rectifying the deficiencies, the programme can improve student achievements and improve their skills in fulfilling the PO. The importance of developing proper POs with the use of domains helps improve the selection of highly competitive graduates. This is a challenge to all institutions offering engineering programme to produce fresh engineering graduates with minimum 12 POs to qualify them as certified engineers. The final outcomes of producing graduates that meet shareholders' expectation are what is more important to be achieved especially in meeting current employability market. Therefore, it is necessary to consider carefully the suitable domains for engineering graduates in order to ensure they will be employable and will remain competitive in their professional careers.

References

Abdullah, S., Atiq, R., Rahmat, A. O. K., Zaharim, A., Muhamad, N., Deros, B., et al. (2008). Implementing continual review of programme educational objectives and outcomes for obe curriculum based on stakeholders' input. In *International Conference on Education and Educational Technology (EDU'08)* (pp. 218–223).

Amirulddin, U. A. U., Osman, M., & Hamid, F. A. (2009). Analysis of programme outcomes achievement for electrical engineering programmes in UNITEN. In *Proceedings of 2009 International Conference on Engineering Education, ICEED2009—Embracing New Challenges in Engineering Education*, pp. 148–152. doi:10.1109/ICEED.2009.5490596

Arshad, I., Razali, S. F. M., & Mohamed, Z. S. (2012). Programme outcomes assessment for civil and structural engineering courses at Universiti Kebangsaan Malaysia. In *Procedia Social and Behavioral Sciences*, *60*, 98–102. doi:10.1016/j.sbspro.2012.09.353

Bloom, B. S. (1956). *Taxonomy of educational objectives book 1: Cognitive domain*. *Evaluation* (Vol. 3). Retrieved from http://www.amazon.com/dp/0582280109

EAC. (2012). *Engineering programme accreditation manual*.

Gamboa, R., Namasivayam, S., Al-Atabi, M., & Singh, R. (2013). Quantitative measurement of students PO attainments for Taylor's University Engineering Programmes. *Procedia Social and Behavioral Sciences*, *103*, 753–762. doi:10.1016/j.sbspro.2013.10.396

Hanrahan, H. (2011, September). The Washington Accord past, present, future.

Masni-Azian, A., Rahimah, A. H., & Othman, M. S. (2014). Towards OBE: A case study of course outcome (CO) and programme outcome (PO) attainment for product design and development course. *IOSR Journal of Research & Method in Education, 4*(2), 55–61.

Mutalib, A. A., Maulud, K. N. A., Yusuff, N. I. M., Rahmat, R. A. A., & Chik, Z. (2012). The achievement of program outcomes from the perspective of CSED, UKM final year students: Part II-session 2006/2007. *Procedia Social and Behavioral Sciences*, *60*, 350–357. doi:10.1016/j.sbspro.2012.09.390

Nordin, R., Bakar, A. A. A., Zaki, W. M. D. W., Zulkifley, M. A., & Huddin, A. B. (2014). Relationship between direct and indirect assessment to improve the teaching and learning process for electrical engineering programmes. *Global Journal of Engineering Education, 16* (3), 104–109.

Shanableh, A. (2014). Alignment of course contents and student assessment with course and programme outcomes—A mathematical approach. *Engineering Education*, *9*(1), 48–61. doi:10.11120/ened.2014.00021

Assessment for Learning: Students' Perception on Peer Review in a Private University

Kim Lian Lee and Mohd Naim Norbaizura

Abstract The process of assessing students has long been placed in schools. This process in the classroom can take many forms. The most important feature is that it is seamless with teaching and learning and constantly informing the teaching and learning process. While it can generally be seen that the increased use of assessment leads to higher quality of learning, it is often claimed that peer evaluation is often ignored. A key argument is that peer students can assess their counterparts' work and providing their feedback to them, while higher education can build on this assessment to evaluate students' overall performance. This shift in focus, whereby students are seen as having a proactive rather than a reactive role in generating and using feedback, has profound implications for the way in which teachers organize assessments and support learning. Thus, this research looks at peer-review issues and how effective this process in evaluating students. By using the evidence in this research, it will provide insights to teachers and the students in higher education on how peer assessment can improve students' learning.

Keyword Peer assessment · Teaching and learning · Students' evaluation

1 Introduction

One of the most important roles for higher education is that it provides a foundation for lifetime of learning, including socialization process. Formal learning and informal learning have existed for centuries from the Chinese civil servant examinations for public office, to Aristotle's students' assessment for medieval guilds (Earl 2006). Since the beginning of industrialization took place, schools have become a place for social establishment for evaluating student's achievement.

K.L. Lee (✉) · M.N. Norbaizura
Taylors University, Subang Jaya, Malaysia
e-mail: KimLian.Lee@taylors.edu.my

M.N. Norbaizura
e-mail: norbaizura.mohdnaim@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_18

Universities need to prepare students to learn beyond the academy once they leave these universities. This is a daunting challenge. Recently, the role of higher education in lifelong learning has been documented through a number of researches. They include a focus on learning outcomes (Hussey and Smith 2003), the use of graduate attributes (Hager and Holland 2006), the promotion of key skills, the adoption of an agenda of employability (Dearing 1997), and the development of capability (Stephenson and Yorke 1998). However, these researches are insufficient. Sadler (1998), Hounsell (2003), and Yorke (2003) argued that there are inadequacies of formative assessment practices during the students' courses. There has also been substantial criticism of the role of summative assessment and its negative effects on student learning (Ecclestone 1999; Knight 2002; Knight and Yorke 2003).

On the other hand, Woodward (1998) disputes that assessment practices do have positive effects on learning, such as journal reflections and portfolios. There has been a considerable interest in the concept of constructive alignment as well (Biggs 2003). Constructive alignment is used to examine whether the components in the teaching system, particularly the methods used and the assessment task, are closely aligned to the learning activities assumed in the intended outcomes. This includes initiatives that look at the long-term consequences of university courses, which includes assessment, on subsequent learning in professional practice (Mentkowski et al. 2000). It was suggested that current assessment practices in higher education did not equip students well for a lifetime of learning and the assessment challenges they would face in the future.

This study looks at one particular aspect of assessment, namely assessments in which students play a role as assessors. The literature review focuses on forms of self-, peer, and co-assessment from the point of view of their applicability in higher education. The following research question is addressed:

What are the main findings from research on self-, peer, and co-assessment?

2 Literature Review

Boud and Falchikov (2006) pointed out that there are basically two assumptions on the purposes of assessment. The first is to provide a certification of achievement, which allows students to graduate with a validated record of their performance in the program. Certification is used by employers and educational institutions, typically to make judgements about acceptability of candidates for employment and further study. The second purpose of assessment is to facilitate learning. This is done through the various kinds of test or assignment. Students are able to effectively judge their own achievements and what they need to do to learn more effectively within the program. These two purposes have been associated with two sets of practices: summative and formative assessment, respectively.

However, there are issues as to whether summative and formative assessment practices are able to adequately address a wider set of needs, for instance equipping

students for a lifetime of learning. Though in principle they might be able to do this, there are sufficient problems; nevertheless, it is effective to establish an additional purpose of assessment (Boud and Falchikov 2006).

Assessment is seen to employ a profound influence on student learning: on what students focus their attention on, on how much they study, on their quality of engagement with learning tasks, and, through feedback, on their understanding and future learning (Gibbs and Simpson 2003). As students become ever more strategic, the way assessment influences learning becomes critical in determining student learning outcomes and performance.

One of the most encouraging contributions from educational research is understanding the role of assessment in the process of learning. Black and Wiliam (1998) synthesized over 250 studies linking assessment and learning and found that the intentional use of assessment in the classroom to promote learning improved student achievement. Crooks (1988) also did a meta-analysis which discovered that classroom assessment had both short- and long-term effects on learning. In the short term, classroom assessment could focus attention on important aspects of the subject, give students opportunities to practice skills and consolidate learning, and guide further instructional or learning activities.

In the medium and long term, assessment held the possibility of influencing students' motivation as learners and their perceptions of their capabilities, communicating, and reinforcing teaching goals, including performance criteria and desired standards of performance; influencing students' choice of and development of learning strategies, skills, and study patterns; and influencing students' subsequent choice of courses, activities, and careers.

Because of these factors, it has become obvious that assessment is a powerful tool for learning. If learning is the goal, major changes are required in the way that assessment is used in classrooms. Learning is now viewed as a process of constructing understanding by attempting to connect new information to what is already known so that ideas have some personal coherence.

Earl (2003) and Earl and Katz (2005) have expanded on the differentiation made by the Assessment Reform Group to describe three intertwined but distinct assessment purposes—assessment for learning, assessment as learning, and assessment of learning. Assessment for learning is designed to give teachers information that will allow them to modify the teaching and learning activities in which students are engaged, in order to differentiate and understand how individual students approach their learning. This assessment suggests that students are all learning in individual and idiosyncratic ways, while recognizing that there are predictable patterns and pathways that many students go through.

2.1 Benefits of Peer Assessment

Falchikov and Goldfinch (2000) pointed out that a program of assessment which incorporates an element of peer assessment is beneficial to learning. Specific

benefits cited include the following: increased student responsibility and autonomy; evaluative skill development; insight into assessment procedures and expectations for high quality work; students work harder with the knowledge that they will be assessed by their peers; potential for providing increased levels of feedback without increasing demands on tutors (Walker 2001); and encourages deep rather than surface learning (Brown et al. 1996).

Evidence also suggests that students often fail to fully understand or utilize assessment criteria, do not know what a good or bad piece of work looks like, are focused toward the awarded mark or grade, and, as such, fail to read, understand, or adequately process tutors' feedback or act upon it (Crème and Lea 1997; Ding 1998; Glover and Brown 2006; Hounsell 1987; Lea and Street 1998; Wotjas 1998). Therefore, this is another reason why peer assessment is important given its reputed benefits in terms of skill development and improved learning and performance on assessed work (Brown et al. 1996).

3 Research Methodology

The primary purpose of the paper is to analyze students' perception of the peer assessment process, reflecting on its benefits and suggesting modifications for its improvement. The participants were from students currently pursing different degree courses at Taylors University. A total of 115 responses were received for data analysis.

The study methodology consisted of the following steps and substeps:

1. Drafting a questionnaire on student's perception of peer assessment (see Appendix).
2. Design and implementation of peer assessment in each course. Within this aim, the corresponding activity was adapted to each particular course, but taking into account the following common criteria.
    - Before starting the questionnaire, the teacher explains to students the aims of the questionnaire and that the questionnaire is based on previous peer assessment done for a particular course subject.
    - Following the teacher's explanation, students answer the questionnaire based on the subject where peer assessment is evaluated. The students are given a maximum of 15 min to complete Sections A and B of the questionnaire.
3. Collection of results from questionnaire.
4. Analysis and discussion of results obtained from the different courses.

For many students, it was the first time that they had assessed their peers; therefore, it was important that the lecturer explained the peer assessment activity in detail before answering the questionnaire. Many authors have suggested that students need to understand learning objectives before evaluating their peers (Black and Wiliam 1998; Sadler 1989). It has also been shown that lecturers and students can have different perceptions of objectives of an activity as well as evaluation

criteria (Hounsell 1997; Norton 1990). As such, detailed explanations are important, so that the students might adopt and apply them in a reasonable way (Cestone et al. 2008; Falchikov and Goldfinch 2000; Lane 2007).

Specifically, the study is interested in the opinion of students with regard to their ability to participate in a peer-review process, the responsibility in which it entails, the level of preparation involved, the degree to which learning is achieved, and subjectivity and the demands of peer assessment in comparison with that of a lecturer. The questionnaire was developed with a 10 set of questions (Section B) to be answered on a Likert scale of four possible responses, ranging from '1—strongly disagree' to '4—strongly agree.' Two open questions were included in the questionnaires, requesting that students indicate two benefits and two limitations of peer assessment. For each subject, the student's responses to the multiple choice questions were analyzed quantitatively using statistical analysis, while open questions were analyzed qualitatively.

4 Results

There are 115 respondents of which 36 % are males and 64 % are females. These students are from different majors such as food science and nutrition (29 %), medicine (36 %), international business and marketing (18 %), accounting and finance (6 %), business and finance (2 %), finance and economics (5 %), and business administration (4 %). The findings are shown in Table 1.

Table 1 Respondents' characteristics

Respondents' characteristics	Classification	Frequency	Percentage (%)
Ethnic group	Malay	35	31
	Chinese	66	57
	India	6	5
	Others	8	7
Gender	Male	41	36
	Female	74	64
Age	50 and above	0	
	40–49	0	
	30–39	0	
	20–29	115	100
Program	Food Science and Nutrition	33	29
	Medicine	41	36
	International Business and Marketing	21	18
	Accounting and Finance	7	6
	Business and Finance	2	2
	Finance and Economics	6	5
	Business Administration	5	4

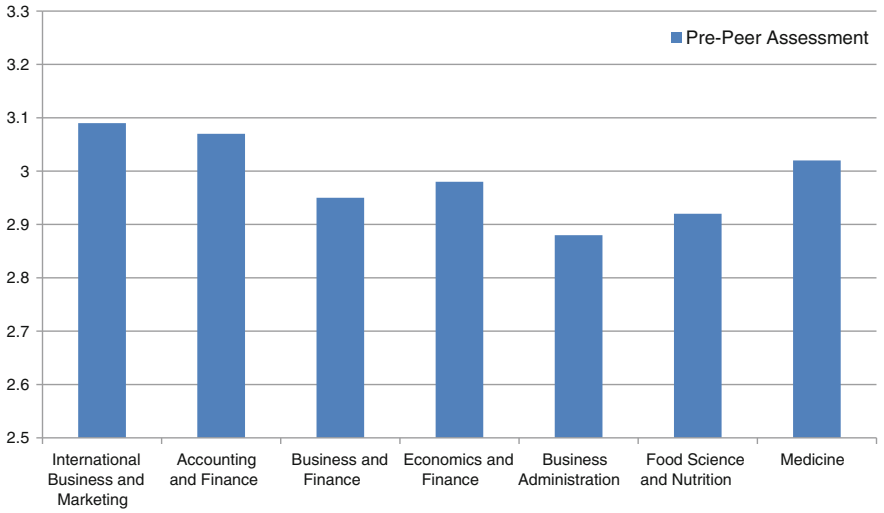


Fig. 1 Mean scores for pre-peer assessment questionnaires by program

Generally, the student's opinion of peer assessment was positive, as the mean average for all questions by program was above 2.0, and the overall average for all programs was 2.6 (on a Likert scale of 1–4). If we analyze the characteristics of the programs, some clear differences are observed (Fig. 1). Students enrolled into the Business Administration program are more reluctant to use peer assessment. However, students who are enrolled into the International Business and Marketing program show a better predisposition toward and interest in the benefits of peer assessment.

All respondents perceive to view peer assessment positively between programs. However, after further analysis, there are different elements of the process where variations are detected. The most significant of these are highlighted below.

4.1 Confidence in Personal Abilities

For most of the programs, there is a positive trend with regard to opinions on the training and skills needed to participate in a peer assessment process (Question 1). Generally speaking, students valued being trained to participate in the peer-assessment process. However, students felt that they did not have the capabilities of assessing their own peers because of their lack of knowledge in the subject. Some of the opinions of this attitude given in the open answers are as follows:

Peers may not be as knowledgeable as lecturers.
If my peers know less about that particular topic, there won't be a fair assessment.
Lack of peer skill.

4.2 Variable Responsibility

The perception of student responsibility in this assessment system varies according to the type of activity performed. Many students feel that if they were to be assessed by their peers, their peers will not be able to handle the responsibility. This is because peer assessment requires students to be fair and respondents felt that there will be a sense of biasness when students assess their peers. Students felt that their peers might not be subjective in their judgment. Many respondents feel that in order to safeguard friendships and certain bonds, their assessment will not be objective. Also, in order to ensure that the individual student get high assessment marks, the student in turn will assess his or her peers highly. In the end, the objectivity of peer assessment is questioned.

If we relate the answers to Question 4 regarding perception of responsibility to the rating of peer assessment in each program, the study concludes that the weight of the rating students award this form of assessment does not affect their perception of responsibility. That is, a higher percentage rating does not equate to a higher perception of responsibility.

4.3 Demanding Fellow Students and Variable Objectivity

Question regarding fellow students being demanding (Questions 10) and the subjectivity of their assessment (Question 9) score among the highest in the questionnaires, although there is a high diversity of opinion among students.

With regard to whether students are more subjective than the lecturer (Question 9), the results show different parameters of analysis: the type of task, area of knowledge, and level of knowledge among students. Students attribute a lack of objectivity on the part of their peers. Students perceive that when assessing, their peers do not adopt an attitude of complicity or competitiveness, but rather the dimension of friendship plays an important role. Student biased toward friends, and this generates a lack of trust in the system of peer assessment and the quality of corrections by peers, expressed in responses to the open questions. It is worth considering some quotes from students to illustrate these uncertainties:

Some peers are biased towards their group of friends (Medical students)
Could be bias between friends depending on how close the friendships are (Food Science and Nutrition student)
Assessing your peers might destroy the relationship between each other (International Business and Marketing student).

5 Discussion

Comparing the perceptions of students from different degree courses on peer-assessment provides some significant results that complement previous research findings. In general, although students perceive peer assessment in a positive light which is in line with studies from Cestone et al. (2008), Gielen et al. (2011), and Nicol and Macfarlane-Dick (2006), many students feel that they are inadequate to assess their peers. This is because of their lack of knowledge pertaining to the subjects. This is both seen for the humanities and science students who find it difficult to accept the responsibility of assessing their peers because in comparison with their lecturer, they perceive that their knowledge as inadequate. Despite being provided with the rubrics for peer assessment, students' perception is that they are inadequate and not fit to assess their peers.

Secondly, in order to maintain certain friendships and bonds, students find themselves not being able to assess their peers objectively. This view differs to that of the Western culture (Mizoguchi et al. 2006). In the Eastern context, the study finds that there is much hesitation among students to evaluate their peers. Students believe that because of their friendship, they are unable to remain objective. Even if they were to remain objective and score their peer low, they know that friendships will be broken and are not willing to do so. This is because students generally would want to maintain those friendships and bonds.

5.1 Future Research

Researchers can look at the perceptions of peer assessment by further examining the learning process. Future research can be done in assessing peer assessment by providing a set of questionnaire before and after the process. By administering the questionnaires before and after, the researcher is able to reflect further on the benefits of peer assessment and suggests modification for improvement. Also, the method of assessment can be differentiated in future research. Peer assessment can be done either through written assessment or oral assessment. As both written and oral activities differ in their assessment, it would be interesting to see how students perceive this process separately. Hanrahan and Isaacs (2001) comment that the skills of peer assessment are necessary for graduates in their careers and also personal life. This is because peer assessment will aid potential graduates to contribute constructively in collaborating with their peers. Besides that, the experience of peer assessment will also provide other important skills such as them being subjected to criticism by others of the same position and rank and to help them with the responsibility of judging the quality of a peer fairly and objectively.

Appendix

A. YOUR PERSONAL INFORMATION

We need some basic information about yourself and your company to enable meaningful analysis of the results.

1. Your gender:

☐ Male

☐ Female

2. Your ethnic group:

☐ Malay

☐ Chinese

☐ Indian

☐ Others, please specify: _____

3. What is the highest level of formal education you have completed and its discipline?

	LEVEL	**DISCIPLINE**
☐	Primary and secondary education	_____
☐	Diploma or Bachelors degree	_____
☐	Masters degree	_____
☐	Professional qualification (please specify)	_____
☐	Others; please specify	_____

4. Please state your age.

☐ 20 – 29 years ☐ 30 – 39 years

☐ 40 – 49 years ☐ 50 and above

5. Which school do you belong to _____

6. What programme are you currently studying _____

B. PRE-PEER ASSESSMENT

Initial questionnaire used to determine student's perceptions before participating in the peer assessment process.

1 = Strongly disagree; 2 = Disagree; 3 = Agree; 4 = Strongly agree

1.	I have the necessary skills to participate in a peer assessment process	1	2	3	4
2.	My lecturer(s) clearly explained the procedure for effective peer assessment	1	2	3	4
3.	My lecturer(s) made the tools and instruments available to me to perform effective peer assessment	1	2	3	4
4.	Peer assessment means a lot of responsibility for the student	1	2	3	4
5.	Peer assessment will make me prepare my work better	1	2	3	4
6.	Peer assessment will force me to look for more and broader information on the contents of the module or activity	1	2	3	4
7.	Peer assessment will allow me to detect my own mistakes and learn from them	1	2	3	4
8.	Peer assessment will allow me to view learning critically and constructively	1	2	3	4
9.	I think my peers will be more subjective in their assessment (not following predetermined and representative criteria for the activity being assessed) than my lecturer(s)	1	2	3	4
10.	My peers will not be as demanding as my lecturer(s) in their assessment	1	2	3	4

Name two aspects you consider might be positive about this peer assessment experience

1. _____

2. _____

Name two difficulties or obstacles you might find in peer assessment

1. _____

2. _____

References

Biggs, J. (2003). *Teaching for quality learning at university* (2nd ed), Society for Research into Higher Education & the Open University Press, Buckingham.

Black, P., & Wiliam, D. (1998). Assessment in education: Principles. *Policy & Practice, 5*(1), 7–68.

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education, 31*(4), 399–413.

Brown, R., Pressley, M., Van Meter, P., & Schuder, T. (1996). A quasi-experimental validation of transactional strategies instruction with low-achieving second-grade readers. *Journal of Educational Psychology, 88*, 18–37.

Cestone, C. M., Levine, R. E., & Lane, D. R. (2008). Peer assessment and evaluation in team-based learning. *New Directions for Teaching and Learning, 2008*(116), 69–78.

Crème, P., & Lea, M. (1997). *Writing at university*. Buckingham: OUP.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*, 438–481.

Dearing, R. (1997). *Higher education in the learning society: The report of the National Committee of Inquiry into Higher Education*. Norwich, Her Majesty's Stationery Office, UK.

Ding, L. (1998). *Revisiting assessment and learning: Implications of students' perspectives on assessment feedback*. Paper Presented at the Scottish Educational Research Association Annual Conference, University of Dundee, Dundee, September 25–26, 1998.

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Experts in Assessment series. Thousand Oaks, California: Corwin Press, inc.

Earl, L. (2006). Assessment—A powerful lever for learning. *Brock Education, 16*(1), 1–15.

Earl, L., & Katz, S. (2005).*Rethinking classroom assessment with purpose in mind*. Winnipeg, Manitoba: Western Northern Canadian Protocoll.

Ecclestone, K. (1999). Empowering or ensnaring? The implications of outcome-based assessment in higher education. *Higher Education Quarterly, 53*(1), 29–48.

Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta analysis comparing peer and teacher marks. *Review of Educational Research, 70*(3), 287–323.

Gibbs, G., & Simpson, C. (2003). Measuring the response of students to assessment: The Assessment Experience Questionnaire. 11th Improving Student Learning Symposium, UK.

Gielen, S., Dochy, F., & Onghena, P. (2011). An inventory of peer assessment diversity. *Assessment & Evaluation in Higher Education, 36*(2), 137–155.

Glover, C., & Brown, E. (2006). Written feedback for students: Too much, too detailed or too incomprehensible to be effective? *Bioscience Education E-journal, 7.*

Hager, P., & Holland, S. (2006). *Graduate attributes, learning, and employability*. Dordrecht: Springer.

Hanrahan, S. J., & Isaacs, G. (2001). Assessing self and peer assessment: The student's views. *Higher Education Research & Development, 20*(1), 53–59.

Hounsell, D. (1987). Essay writing and the quality of feedback. In J. T. E. Richardson, M. W. Eysenck, & D. Warren-Piper (Eds.), *Student learning: Research in education and cognitive psychology*. Milton Keynes: SRHE/Open University.

Hounsell, D. (1997). Contrasting conceptions of essay-writing. In: F. Marton, D. Hounsell, & N. Entwistle (Eds), *The Experience of Learning* Edinburgh: Scottish Academic Press.

Hounsell, D. (2003). Student feedback, learning and development. In M. Slowey & D. Watson (Eds.), *Higher education and the life course* (pp. 67–78). Buckingham Society for Research into Higher Education & Open University Press.

Hussey, T., & Smith, P. (2003). The uses of learning outcomes. *Teaching in Higher Education, 8*(3), 357–368.

Knight, P. T. (2002). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education, 27*(3), 275–286.

Knight, P. T., & Yorke, M. (2003). Assessment, learning and employability. Maidenhead: Society for Research into Education & Open University Press.

Lane, D. (2007).*Engineering feedback: A student-developed approach to the assessment of peer evaluation in civil engineering*. Chicago: National Communication Association.

Lea, M., & Street, B. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education, 23*(2), 57–172.

Mentkowski, M., et al. (2000). *Learning that lasts. Integrating learning, development and performance in college and beyond*. San Francisco, CA: Jossey-Bass Publishers.

Mizoguchi, R., Dillenbourg, P., & Zhu, Z. (2006). *Learning by effective utilization of technologies: Facilitating intercultural understanding*. Amsterdam, Netherlands: IOS Press.

Nicol, D, J. & Macfarlane-Dick. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–216.

Norton, L. S. (1990). Essay writing: What really counts? *Higher Education, 20*(4), 411–442.

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education, 5*(1), 77–84.

Sadler, D.R. (1989). Formative assessment and the design of instructional systems.*Instructional Science, 18*, 119–144.

Stephenson, J., & Yorke, M. (1998). *Capability and quality in higher education*. London: Kogan Page.

Walker, A. (2001). British psychology students' perceptions of group work and peer assessment. *Psychology Learning and Teaching, 1*(1), 28–36.

Woodward, H. (1998). Reflective journals and portfolios: Learning through assessment. *Assessment & Evaluation in Higher Education, 23*(4), 415–423.

Wotjas, O. (1998). Feedback? No, just give us the answers. *Times Higher Education Supplement, 25*.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*(4), 477–501.

Positive Impacts of Service Learning on Students' Personal Outcome and Social Outcome

Pei Leng Kelly Tee and Chandra Sakaran Kalidas

Abstract Service learning has received considerable attention in the field of education lately, especially among universities across Malaysia. Service learning is an experiential learning approach to teaching and learning that has a set of diverse learning outcomes. Students found that service learning has provided them a meaningful academically community service experience which enhance their personal growth in their undergraduate study. Over the years, there have been quite a number of studies which reported on the numerous benefits of service learning (Astin et al. in How service learning affects students. Higher Education Research Institute, UCLA, 2000; Singleton in J Gerontological Soc Work 50(1):105–118, 2007; Eyler et al. in At a glance: What we know about the effects of service learning on college students, faculty, institutions and communities, 1993–2000, 2001; Eyler and Giles in Where's the learning in service learning? Jossey-Bass Publishers, San Francisco, 1999). Therefore, this paper aims to explore the positive impacts service learning has on students' personal outcome and social outcome among our undergraduates. The study employs a qualitative study which utilizes students' reflective writing as the main data source and fieldwork observation. The key findings of this paper suggest that service learning helped enhance the process of self-discovery among students which in turn developed their overall personal and social growth.

Keywords Service learning · Positive impact · Self-discovery · Social outcome

P.L.K. Tee (✉) · C.S. Kalidas
Taylor's University, Subang Jaya, Malaysia
e-mail: peileng.tee@taylors.edu.my

C.S. Kalidas
e-mail: chandrasakaran.khalid@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_19

1 Introduction

Service learning is a unique and powerful teaching and learning pedagogy that integrates community service and classroom learning. This new form of experiential learning has become widespread everywhere, including in Malaysia, where this educational strategy has gained recognition in the last five years, both in schools and institutions of higher learning.

Service learning has been widely practiced in the USA for more than fifteen years (Learn and Serve America 2010). According to Skinner and Chapman (1999), in 1999, 32 % of all public schools and 50 % of all high schools practiced service learning as part of their learning methodology and in 2009; the number has increased significantly to 1190 schools in USA (Campus Compact 2009). In Malaysia, as mentioned earlier, there has been significant development in the use of service learning, especially in institutions of higher learning, either in the form of co-curricular activities, outreach programmes or community services due to its status as a compulsory module for undergraduates.

Over the years, there have been quite a number of studies which reported on many benefits to those who participate in this experiential pedagogy (Astin et al. 2000; Singleton 2007; Eyler et al. 2001; Eyler and Giles 1999). These studies also show that students gain the maximum benefit when service learning links or connects the curricular concepts to the context of real-life situations. Several research studies suggest that service learning programmes do have positive influences on the educational institutions conducting the programmes, on the community receiving the services and on the student participants who may benefit academically, socially or personally (Billig 2009; Conway et al. 2009).

This paper aims to explore the positive impacts service learning has on students' personal outcome and social outcome among undergraduates of a private institution of higher learning. The question that guided this study is as follows:

1. What are the positive impacts of service learning?
   (a) What personal outcomes did the students experience?
   (b) What social outcomes did the students experience?

2 Literature Review

2.1 Service Learning Defined

There have been various terms used to define service learning; from community service projects to field studies to internship programmes and the need to distinguish between co-curricular and curricular service learning.

According to Denby (2008), curricular service learning integrates service learning with academic and programme curricula with clear programme goals and

Table 1 A service and learning typology

Service LEARNING	Learning goals primary; service outcomes secondary
SERVICE learning	Service outcomes primary; learning goals secondary
service learning	Service and learning goals completely separate
SERVICE LEARNING	Service and learning goals of equal weight and each enhances the other for all participants

Sigmon (1994)

outcomes for the students. The service learning projects are carried out by the institution in collaboration with community partners to meet the community's needs and to enhance students' understanding of the programme content. An integral part of this type of service learning is the reflective component that is designed to help students to form the transformative link between their service, the curriculum, the community and the head, the heart and the hand. Co-curricular service learning, on the other hand, is part of the students' extracurricular activity. It is not directly connected to students' coursework or official programme plan. Students may get involved in working with the community and gaining some experience and understanding of the society around them, but it remains separate from the academic curricula (Denby 2008). This paper will focus on curricular service learning.

Sigmon (1994) had provided a more precise definition of service learning through a typology that compares different activities that integrates service and learning. This typology includes the view that service learning occurs when there is a balance between learning goals and service outcomes. This typology is useful in establishing criteria to differentiate and to provide a basis for diverse service-oriented experiential academic programmes (see Table 1).

Prentice and Robinson (2007) define service learning as an academic programme that integrates community service with academic instruction with emphasis on critical and reflective thinking and personal and civic responsibility. This view is shared by Hussey and Smith (2002), who suggest that service learning provides students with authentic experiences that can be related to course content so that students obtain knowledge of facts, rules, procedures and learned skills and abilities. Bringle and Hatcher (1995) pointed out that service learning is an academic experience in which students learn and develop through service activity to help meet a community's need and to reflect on the service to gain further experience and knowledge of course content and an enhanced sense of personal and social responsibility.

All these definitions suggest that service learning is an experiential pedagogy that has positive influence on the community receiving the services, on the educational institution and on the student participant, which also the main focus of this paper. Several researches point to positive influences on five areas: personal development, attitudes towards education institution, academic achievement, civic engagement and social skills (Billig 2009; Conway et al. 2009). This paper aims to prove that service learning has positive impacts on students' personal and social outcomes.

2.2 Personal Outcomes

Research has shown that service learning is associated with positive personal outcomes, such as personal and interpersonal development, acquisition of leadership skills and self-discipline (Billig 2009; Conway et al. 2009). A research conducted by Shalabi (2014) shows significant contribution of service learning in the area of personal development. According to Shalabi, this finding confirms Dugan's findings (2006) that service learning provides practical opportunities to develop students' leadership and professional communication skills as well as enhancing their collaboration skills. A study by Moely et al. (2002), conducted on service learners and non-service learners of a university, shows service learners have higher scores for social justice attitudes, leadership skills and problem solving skills as compared to non-service learners.

An explanatory study by Simons and Cleary (2006) suggests that service learning contributes to students' personal development in the area of perceived personality and self-knowledge and this study is consistent with the findings by Eyler and Giles (1999) that suggest service learners have a better perception of oneself and have greater self-knowledge. Reports suggest that students who participate in service learning have shown increased self-esteem and self-efficacy and fewer behavioural problems as compared to non-service learners (Shaffer 1993; Switzer et al. 1995).

Several other studies found that students acknowledged service learning has improved their interpersonal skills. Students noted that through service engagement, they had developed their ability to work well with others and appreciate cultures of different community. Students agreed that this programme allowed them to connect well with the community and take ownership of the problems faced by the community they serve (Gallini and Moely 2003; Reinke 2003). Simons and Cleary (2006) noted in their study that half the students who participated in service learning activities registered for another service learning course the following semester and this truly shows the high regard and the benefit students derive from service learning.

A study by Yamauchi et al. (2006) discovered that service learning helped students and youth explore career options. Yamauchi noted that students who had participated as service learners had a stronger set of job- and career-related skills and aspirations, such as knowledge on planning activities, job interview skills, and preparation of career plans compared to non-service learners. This study is consistent with Furco's study (2003), which found strong statistically noteworthy differences on formulation of career plans and finding a job that is personally satisfying or beneficial to others. These studies show students and youth believe service learning had provided them with confidence, knowledge and that cutting edge to be better prepared for a competitive job market compared to the non-service learner groups.

2.3 Social Outcomes

Several research studies show that service learning increases students' social awareness and understanding of their community's needs, increases understanding of social and cultural diversity and reduces ethnocentrism and stereotypical beliefs (Borden 2007; Eyler and Giles 1999). Borden (2007) further explained that students, who had completed a semester of service learning within a cultural group of different ethnicity, exhibited lower levels of ethnocentrism as compared to at the beginning of the programme. This change in attitude is possible due to the students' constant engagement with the diverse culture during the few months of the service programme. Borden had examined the reflective essays written by the student to find indications and clues as to the role service learning played in changing the students' attitude. Borden's finding is further emphasized by Dunlap (2000), who point out that service learning helps to integrate students with communities that are culturally different to them and this leads to a more meaningful inter-group interactions and relationship.

Service learning provides students the opportunity to study relationship among members of the community and to understand that all humans are the same and should not be labelled or stereotyped by their ethnicity, colour or socioeconomic status. Some institutions use service learning to create awareness among their students on how stereotypes dehumanize people and how stereotyping people can have negative impact on individuals and society (Keller et al. 2003). Keller, Nelson and Wick further reiterates that there comes a change of attitude in students as they acquire in-depth knowledge of the issues and able to look at others in a multidimensional terms. Prentice and Robinson (2007) point out that service learning allows students to recognize their lack of understanding their own biases and to make an effort to acquire accurate information to be culturally competent. Prentice and Robinson affirmed that service learning has allayed students' fear of the unknown and provided opportunities to develop empathy and compassion for people who are unlike them.

A study by Astin et al. (2000) suggests that service learning increased students' awareness of the world and of their personal values as they increased engagement in the classroom experience. The findings suggest that students developed a heightened sense of social values that allow them to interact and relate their knowledge and understanding to the real-world situations. The development of these skills and knowledge is believed to have made these students to have greater confidence in their communication and social skills which helped to have a better appreciation and understanding of the "real world".

3 Methodology

3.1 Case Study

This study employs a case study design which investigates a phenomenon in a natural setting within a bounded system (Merriam 1998; Yin 2003). In this study, the researchers wanted to gain insights on service learning as a platform to discover the positive impacts it has on students' personal outcome and social outcome. This is further supported by Merriam (1998) that the use of case study helps to gain an in-depth understanding of the situation and the meaning it provides.

The present study was conducted in a community service course at a private university in Selangor, Malaysia. The course was a compulsory subject offered to undergraduate business students and the duration of the course was one semester which lasted for 14 weeks. The participants of this study were a mix of 51 local and international students. The students were a mixed-gender between the ages of 20–21 years. In terms of language, for some of the students, English is their first language while for the others English is their second or third language. The students were in their second year of their studies majoring in Finance & Economics and Banking and Finance degrees.

The participants provided two-to-one coaching and tutorial assistance in English and mathematics subjects to secondary two school students (mentees) in two secondary schools in Sentul. The students also acted as big brother/sister (mentor) to help their mentees break out of the poverty cycle by inspiring them to complete their secondary education, and helping them to develop the necessary literacy and numeracy skills needed for work and life. Additionally, the course provided opportunities for the mentors to use their present and newly acquired skills in real-life situations and learned the importance of contributing to the larger community of which they are a part of.

3.2 Data Collection

The data for this study were obtained from two qualitative research tools: (1) reflective journals and (2) fieldwork observation. These two tools are important in this study as they provided detailed information on the positive impacts service learning has on students' personal outcome and social outcome among our undergraduates.

3.2.1 Reflective Journals

The main data source for this study was obtained from the students' reflective journal. The reflective journals are important in this study as it provided detailed information on the service learning impacts and experience of the students. Consent

was obtained from the students, and they were informed that the reflective report is part of a research project. The students were asked to write five journal entries of about 500 words each regarding their service learning experience and the reflective reports were collected in week 10, a week after they completed the programme with their mentees.

In the reflective journal, students were asked to describe what they did, rate how much they enjoyed it, explain if and in what ways it benefited them, what they would like to change, memorable experiences and any connections they saw to the community service course. This is to gain insight on the students' service learning experience and to make meaning of their personal likes and dislikes, and personal development which took place in them.

3.2.2 Fieldwork Observation

This study also employed fieldwork observation which took place at the two schools in Sentul. The observations allowed the researchers to explore the interactions between the mentors and their mentees during the coaching sessions and to substantiate the emerging themes found from the students' reflective journals. The researcher's goal was to maintain the role of a passive participant observer; the researcher sought to observe from the margins of the classrooms while keeping researcher participation and interference to a minimum (Spradley 1980, pp. 59–60). There were times where the researcher was forced to be an active participant observer given the needs of the people among whom the service took place. At times, observations followed a macro-approach of the entire classroom which included observation of multiple mentors and mentees at once. Other times, the researcher moved around the classroom to focus closely on a particular student or smaller group of students. Field notes were taken which included a general description of the activity and interaction and verbal and non-verbal interaction.

3.3 Data Analysis

The qualitative data were organized using an EXCEL database and analysis followed a process of analytic induction to generate themes and codes. Analysis was ongoing throughout the research and began during the initial review and transcription of field notes. The researchers made both high- and low-level inferences about the data as it was being collected. This helped the researchers to "cycle back and forth between thinking about the existing data and generating strategies for collecting new, and often better, data" (Miles and Huberman 1994, p. 50). First, the coding categories for occurring themes were identified through the reading of the reflective journals and the field notes. The reflective journals and field notes were read through individually to develop a system of categorization. In order to develop an appropriate categorization, it took several rounds of individual categorization

followed by intensive discussions with another lecturer until a consensus on an appropriate categorization model was reached. The data were analysed based on the students experience in a service learning module. In the final analysis and report, general themes and findings are discussed in the results section and will be supported with evidence from the database. In sum, this study seeks to capture the students' experiences and to document the developmental process they go through over the service learning course.

4 Findings and Discussions

The findings indicated that the students experienced positive impacts from their service learning experience. The positive impacts can be grouped into two main categories which are students' personal outcome and social outcome. In personal outcome, the students indicated that it has helped them develop and or/practice several intra- and interpersonal qualities which included improvements in social and conversational skills, patience, teamwork and self-esteem. Additionally for social outcome, the students experienced an increase in social responsibility and citizenship skills and their commitment to service. Each of this will be explored in further detail.

4.1 Personal Outcomes

Under personal outcome, the students developed several intra- and interpersonal qualities which enhanced their communication skill, ability to work well with others, sense of being patient and overall sense of self-esteem.

The first key finding that was mentioned repeatedly by the students was the service learning experience helped them improved their communication skill. One student mentioned that, "*After being involved in this module, I felt easier interacting with others and I also know how to communicate with different people as each and every one has different personalities*". Additionally, another student added that, "*Communication skill is so important as we need to know how to speak courteously and effectively. This module trained me to be able to communicate with my mentee through the use of body language, oral communication and most importantly to listen to what the others are saying*". These experiences helped students to develop conversational and social skills through their service learning experience.

The second outcome that the students had acquired is the ability to work well with others. One student mentioned that, "*...I have learned that teamwork is very important in making everything succeed. Not only does teamwork benefit the whole project and team, it also teaches a lesson in each individual. Firstly, with teamwork when planning the games and activities, each task that needs to be completed can*

be accomplished faster and more efficiently compared to tackling any problems as an individual". This clearly shows that the students are able to work as a team to come up with weekly activities for their mentees and this shows that teamwork is crucial for team success. This finding concurs with Eyler and Giles (1999) which found in their study that 40 % of their survey respondents indicated that learning to work with people was one of the important lessons they took from their service learning experience.

The students had also learnt the importance of teamwork as a student stated that, "*Moreover, generating ideas together and as a team may create better methods and solutions to problems. The learning opportunities are endless as each individual is unique in their own way and have their talent, working with different people would allow opportunity to challenge each other's ideas, contributing to a successful completion of any given task*". This supported the notion on cooperative learning by Peterson and Miller (2004) which mentioned students engaged in cooperative learning activities were more likely to be thinking about something on task, were more engaged in the activity, were more likely to perceive the activity they were engaged in as important and were more likely to perceive the learning as challenging or requiring a high skill level.

Additionally, service learning helped students develop the art of being patient. Students had to be patient working with their mentees due to language barrier. A student quoted that, "*For me it was great opportunity to spend time with the mentees and learn how to be mentor as well as a role model. It has improved my sense of patience as it takes a great deal of understanding and persistence to help my mentee understand a simple grammar rule*". A few students also highlighted that service learning helped them to practice patience. One student explained that, "*I was killing myself as my mentee couldn't understand what I was explaining to her as her English is weak and I had to be very patient with her and finally, I used my very worse Malay language with the help of Google Translate to go through the lesson of the day. This has really tested my tolerance level and it showed me that it takes a great deal of patience to make someone understand something*". As a result, service learning does provide opportunities for students to practise and develop their sense of patience.

The service learning experience also taught the students to appreciate life more fully as one student stated that, "*I became aware of how fortunate and blessed I am. There are many people who don't have the necessities of life. This experience opened my eyes and made me see that homelessness, lack of education is a reality for some people*". As students worked and compared themselves to their mentees, they indicated feeling a sense of gratitude and appreciation for their own life. The students' descriptions in reference to gratitude reflect a well-known social psychology concept known as "relative deprivation theory" (Meyers 2010, pp. 526–527). According to this theory, one's personal happiness is relative not only to their individual past experiences but also to their relative comparisons with others in terms of wealth, status, health or general quality of life; it often promotes an individual's sense of life satisfaction or gratitude.

The final personal outcome reported by the students was it boosted their self-esteem. The service learning experience increased the students' self-confidence as they had a feel-good experience and made them feel good every-time they worked alongside their mentees. One student said, "*It was amazing how greatly others appreciated my time and effort as it gave me a sense of doing the right thing and a good use of my time*". Another student stated, "*In this programme, I did something which I've never done and experienced before as who would have thought that I'm able to help an individual and impact his life. This experience really encouraged me that even the smallest help is able to go a long way*". The students ability to see themselves as role models with the ability to employ their time and talents to make a difference in the lives of others and in their community at large and this was made possible through service learning. This supports the findings by Shaffer (1993) and Switzer et al. (1995) that service learning had increased students' self-esteem and self-efficacy.

4.2 Social Outcomes

In general, the students mentioned that participating in the service learning programme increased in their social responsibility and citizenship skills and their commitment to service as part of the social outcome.

After being involved in the service learning programme, the students felt an increase in their social responsibility and citizenship skills. The student found that being involved in the service learning programme has made them aware of the community and their surroundings. One student mentioned that, "*This module has enabled me to learn something that I cannot be found in books and especially not in a normal business degree. It prompted me to look at life from a different perspective and taught me to be selfish and ignorant about everything that is happening around me*". Another student echoed by saying, "*Honestly, to me community service is a way of giving back to the less fortunate people around us in order to make their life and future better. There is an old saying that goes, 'It is better to give than to receive' and I just realised how powerful that saying was after I have done five weeks of community service*". Additionally, a student stated, "*It is humbling and rewarding to serve others as volunteering gives me a sense of peace and it enforces my belief that we survive together as a community and not as an individual. It reminds me that we need each other more than we think and that without the kindness of others, there would be no hope*". This supports the study done Astin et al. (2000) that service learning increased students' awareness of the world and of their personal values as they increased engagement in the classroom experience.

The students also experienced positive effect on commitment to service as a result of their service learning experience. One student reiterated his experience by stating, "*I plan to participate in more community service projects as I somehow made a difference in my community after I've done five weeks of community service. It would be very fulfilling to know that I changed someone's life even in an*

insignificant way". This supports the study by Simons and Cleary (2006) who found that students have high regards from their service learning experience and registered for another service learning course in the following semester. Additionally, another student mentioned, "*I have learnt that it is vital to help each other and not to mention the community at large. Therefore, I have realized that we can always start with simple acts to bring change in our community as every time I accomplish one simple act like volunteering at a soup kitchen, it makes me feel good that I am part of the team*". The service learning experience also helped the students to recognize their potential to change the world and how it empowers them to make a difference in their community. This substantiates the findings from Eyler et al. (1997) which indicated that service learning programs "do appear to have an impact on students attitudes, values, skills and the way they think about social issues even over the relatively brief period of a semester" (p. 13). They found that participating in service learning increased students' belief that they can be personally effective in the community.

5 Conclusions

In conclusion, the findings from this study clearly indicated that service learning helped enhance the process of self-discovery among students which in turn developed their overall personal and social growth. Therefore, it is noteworthy to highlight the fact that the students themselves perceived a connection between service learning and their own personal growth and development. This self-perception is critical to students' own identify and affects their ability to see themselves as role models with valuable self-worth, independence, self-efficacy, and confidence. Specifically, students who completed this service learning module felt strongly that their experiences were beneficial to themselves and helped them to develop the following attributes: communication skills, teamwork, patience, gratitude, overall self-esteem, sense of social responsibility and citizenship skills and commitment to service. This concurs the findings by Shalabi (2014) and Dugan (2006) that service learning provides practical opportunities to develop students' leadership and professional communication skills as well as enhancing their collaboration skills.

The study also showed that those who participated in service learning were more likely to believe that improving social justice should be a priority for society. Those individuals were able to see things from the perspective of others and be open to new ideas. The students in this study expected that they would have a stronger commitment to serving their community as a result of their service learning experience. They believed they would have a greater awareness of social issues and would understand and appreciate people with differing backgrounds. They anticipated their service learning experience would empower them to make a difference in their community and help them recognize their potential to change the world. This confirms a study by Keller et al. (2003) that there comes a change of attitude in

students as they acquire in-depth knowledge of the issues and able to look at others in a multidimensional terms. In short, service learning offers the opportunity to create what George Kuh (1996) describes as a "seamless learning environment":

> The word seamless suggests that what was once believed to be separate, distinct parts (e.g. in-class and out-of-class, academic and non-academic; curricular and cocurricular, or on-campus and off-campus experiences) are now of one piece, bound together so as to appear whole or continuous. In seamless learning environments, students are encouraged to take advantage of learning resources that exist both inside and outside of the classroom… students are asked to use their life experiences to make meaning of material introduced in classes…. (p. 136)

References

Astin, A., Vogelgesang, L., Ikeda, E., & Yee, J. (2000). *How service learning affects students*. UCLA: Higher Education Research Institute.

Billig, S. H. (2009). Does quality really matter: Testing the new K–12 service learning standards for quality practice. In B. E. Moely, S. H. Billig & B. A. Holland (Eds.), *Advances in service learning research. Creating our identities in service learning and community engagement* (Vol. 9, pp. 131–158). Greenwich, CT: Information Age.

Borden, A. W. (2007). The impact of service learning on ethnocentrism in an intercultural communication course. *Journal of Experiential Education*, *30*(2), 171–183 (Fall 2007).

Bringle, R. G., & Hatcher, J. A. (1995). A service learning curriculum for faculty. *Michigan Journal of Community Service Learning*, *2*, 112–122.

Campus Compact. (2009). *2008 service statistics: Highlights and trends of Campus Compact's annual membership survey*. Boston, MA: Campus Compact. Retrieved from http://www.compact.org/wp-content/uploads/2009/10/2008-statistics1.pdf

Conway, J. M., Amel, E. L., & Gerwien, D. P. (2009). Teaching and learning in the social context: A meta-analysis of service learning's effects on academic, personal, social, and citizenship outcomes. *Teaching of Psychology, 36*, 233–245.

Denby, R. (2008). *Impact of service learning on student's sense of civic responsibility masters in education. Faculty of education*. The University of Western Ontario: London.

Dugan, J. P. (2006). Involvement and leadership: A descriptive analysis of socially responsible leadership. *Journal of College Student Development, 47*, 335–343.

Dunlap, M. R. (2000). *Reaching out to children and families: Students model effective community service*. Lanham, MD: Rowman & Littlefield.

Eyler, J., & Giles, G. E, Jr. (1999). *Where's the learning in service learning?* San Francisco: Jossey-Bass Publishers.

Eyler, J., Giles, D. E, Jr, & Braxton, J. (1997). The impact of service learning on college students. *Michigan Journal of Community Service Learning, 4*, 5–15.

Eyler, J., Giles, D., Stenson, C., & Gray, C. (2001). *At a glance: What we know about the effects of service learning on college students, faculty, institutions and communities*, *1993–2000* (3rd ed.). Retrieved from: http://servicelearning.org/filemanager/download/aag.pdf

Furco, A. (2003). *Service learning: A balanced approach to experiential education. In Introduction to service learning toolkit: Readings and resources for faculty* (2nd ed., pp. 11–14). Providence, RI: Campus Compact.

Gallini, S. M., & Moely, B. E. (2003). Service learning and engagement, academic challenge and retention. *Michigan Journal of Community Service Learning, 10*(1), 1–14.

Hussey, T., & Smith, P. (2002). The trouble with learning outcomes. *Active Learning in Higher Education, 3*(3), 220–233.

Keller, J., Nelson, S., & Wick, R. (2003). Care ethics, service learning and social change. *Michigan Journal of Community Service Learning, 10*(1), 39–50.

Kuh, G. D. (1996). Guiding principles for creating seamless learning environments for undergraduates. *Journal of College Student Development, 37* (2), 135–149.

Learn and Serve America. (2010). *What is service learning?* Retrieved from http://www.servicelearning.org/what-service-learning

Merriam, S. B. (1998). *Qualitative research and case study applications in education* (Rev ed.). San Francisco, CA: Jossey-Bass.

Meyers, D. (2010). *Psychology* (9th ed.). New York: Worth Publishers.

Miles, M., & Huberman, A. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage Publications.

Moely, B. E., Mercer, S. H., Ilustre, V., Miron, D., & McFarland, M. (2002). Psychometric properties and correlates of the civic attitudes and skills questionnaire (CASQ): A measure of student's attitudes related to service learning. *Michigan Journal of Community Service Learning, 8*(2), 15–26.

Peterson, S. E., & Miller, J. A. (2004). Comparing the quality of students' experiences during cooperative learning and large-group instruction. *The Journal of Educational Research, 97*(3), 123–133.

Prentice, M., & Robinson, G. (2007). *Linking service learning and civic engagement in community college students*. Washington, DC: American Association of Community Colleges.

Reinke, S. J. (2003). Making a difference: does service learning promote civic engagement in MPA students? *Journal of Public Affairs Education*, *9*(2), 129–37.

Shaffer, B. (1993) *Service Learning: An Academic Methodology. Stanford*, CA: Stanford University Department of Education cited in (Bhaerman, R., Cordell, K., & Gomez, B. (1998). *The role of service learning in educational reform.* Raleigh, NC: National Society for Experiential Education and Needham, MA: Simon and Shuster, Inc.).

Shalabi, N. M. (2014). Advancing intercultural understanding and personal development outcomes through service learning: insights from an international student. *Journal of Community Engagement and Scholarship, 7*(2). Retrieved from http://jces.ua.edu/advancing-intercultural-understanding-and-personal-development-outcomes-through-service-learning-insights-from-an-international-student/

Sigmon, R. (1994). *Serving to learn, learning to serve*. Linking Service with Learning: Council for Independent Colleges Report.

Simons, L., & Cleary, B. (2006). The influence of service learning on students' personal and social development. *College Teaching, 54*(4), 307–319.

Singleton, Judy L. (2007). Taking the class to the community with service learning. *Journal of Gerontological Social Work, 50*(1), 105–118.

Skinner, R., & Chapman, C. (1999). *Service learning and community service in K–12 public schools*. Retrieved from http://nces.ed.gov/surveys/frss/publications/1999043/index.asp

Spradley, J. (1980). *Participant observation*. Thousand Oaks, CA: Sage Publications.

Switzer, G., Simmons, R., Dew, M., Regalski, J., & Wang, C. (1995). The effect of a school-based helper program on adolescent self-image, attitudes, and behavior. *Journal of Early Adolescence, 15*, 429–455.

Yamauchi, L., Billig, S. H., Meyer, S., & Hofschire, L. (2006). Student outcomes associated with service learning in a culturally relevant high school program. *Journal of Prevention & Intervention in the Community, 32*(1), 149–164.

Yin, R. K. (2003). *Case Study Research: Design and Methods* (3rd ed.). Thousand Oaks, CA: Sage Publication.

Assessment Dialogues: Perceptions of Feedback and Feed-Forward

Tim Griffin, Lisa Armitage, Paul Parker and Sky Hugman

Abstract Carless (Stud Higher Educ 31(2):219–233, 2006) recommended 'assessment dialogues' to reduce tutors' and students' different perceptions of feedback. Written feedback and feed-forward on assessments are a form of dialogue. While this communication is important, so too are assessment dialogues between subject coordinators, tutors and assessment experts. A case study of a large-enrolment, first-year subject is a vehicle for the assessment dialogues. The assessment design is described along with an innovation in which students reflected on their academic writing for their first written assessment and, for their second, reflected on the feedback and feed-forward they received on their first. The first assessment dialogue involved the subject coordinator, an academic learning advisor and students and centres on designing and explicating the syllabus, particularly the assessments. The second assessment dialogue was between the academic learning advisor, tutors and the subject coordinator, focusing on the development and refinement of a feedback/feed-forward comment bank for content and academic writing, which was derived from the analyses of past feedback and feed-forward. The third and final dialogue was between tutors (markers) and students and an analysis of students' interpretations of feedback and feed-forward. It is concluded that there are challenges to the provision of effective feedback and feed-forward for student learning, but dialogues informed by data and theory are effective for gaining traction. In the current context, the bigger picture is to align assessment design, feedback and feed-forward to the first-year transition agenda.

T. Griffin (✉) · L. Armitage · P. Parker · S. Hugman
School of Social Sciences and Psychology, Western Sydney University, Sydney, Australia
e-mail: t.griffin@westernsydney.edu.au

L. Armitage
e-mail: l.armitage@westernsydney.edu.au

P. Parker
e-mail: p.parker@westernsydney.edu.au

S. Hugman
e-mail: s.hugman@westernsydney.edu.au

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_20

Keywords Assessment design · Assessment dialogues · Effective feedback and feed-forward · Tutor professional development

1 Introduction

'Greater Western Sydney', with a population of just over two million, is the fastest growing and most culturally diverse region of Australia. Western Sydney University has six campuses dispersed across the region, and our 45,000 students reflect its cultural diversity. Sixty per cent of students are the first generation in their family to attend university. The widening participation agenda at many Australian universities lifts the imperative to get assessments right, to assess *for* learning (rather than *of* learning) and to ensure that students are afforded opportunities to develop their capacity for academic writing (Gill 2015).

Ensuring best and consistent practice is particularly important for first-year students (Burnett and Larmar 2011; Kift 2009). Tinto (1993) states that first-year students' initial experience at university can impact significantly on their long-term engagement and retention in higher education. It is important to be explicit about expectations, provide opportunities for practice and feedback, establish effective study skills to become independent learners and engage in conversations about what it means to be a learner in a particular discipline (Kift 2009). Given the aforementioned widening participation in higher education in Australia, it is not surprising that sessional staff are increasingly employed to teach, especially in large first-year classes (Bell et al. 2010).

Those initial experiences of university can be a perplexing, not least because students have to navigate the assessment requirements across all of their subjects that determine whether they become 'successful' students. Some students come to university with a love of learning; others acquire it through their studies. Many students, due to competing demands of work and family (Universities Australia 2013), are instrumental learners in that assessments drive their learning. Assessments can be made more tractable through whole of curriculum design, including establishing consistent language and conventions and scaffolding of assessments (Gill 2015). While such strategies are deployed in the case study subject, the emphasis here is on how 'assessment dialogues' are facilitated to pilot students through their written assessment tasks.

Carless (2006) studied barriers that students face in interpreting and applying assessment language such as standards and criteria and proposed mitigation through more active dialogues with tutors. Carless considers that written feedback in assignments is a primary form of dialogue but that translation of tutors' comments and how they may be applied to future work also present difficulties for students. At issue are the unstable and tacit dimensions of academic language which can vary across disciplines, individual subjects of study and also across individual tutors in a team (Wingate 2006). Thus, while assessment dialogues between tutors and students

are essential to reach shared understanding, such shared understanding among the teaching team is also required for effective feedback and feed-forward.

We have expanded the notion of assessment dialogues to the teaching team, which includes the following: the subject coordinator (who is responsible for the learning and assessment design); the senior tutor (who assists with the administration of the subject and moderates the online student discussion forums); and the tutors (markers) and an academic learning advisor (who is on hand for assessment design and teaching delivery input). In this paper, a series of integrated dialogues are between the following: (1) the subject coordinator, the academic learning advisor and students using various modes and materials; (2) the subject coordinator, academic learning advisor and tutors though the development of feedback and feed-forward comments for online marking; and (3) tutors and students, concentrating on students' perceptions of feedback and feed-forward. These dialogues will be discussed in turn.

2 Case Study Subject

The case study introductory psychology subject enrols 1200 first-year students and is delivered across three of the university's six campuses. Assessments are purpose-designed for first-year students, and the philosophy is to provide incentives and encouragement to promote early successes and give students every chance of passing. Assessment design also takes into account the practical constraints of the size of the cohort and the available budget and the fact that all the tutors are sessional and paid by the hour, which means that they have limited availability to consult with students and limited time to spend on each written assignment to provide considered and strategic feedback and feed-forward.

Students attend four two-hour tutorials during the semester with the large cohort divided into two strands so that half the students are in tutorials at a time. Tutors therefore teach over a fortnight, so only half the number of tutors is required than for subjects of a similar size. This means that there is a stable and cohesive teaching team, and support and professional development of small group teaching and assessment skills are easier to transact. While there are significant affordances of the tutorial schedule, it also comes with a cost in that the volume of marking for each tutor is practically double other subjects, a factor that has implications for providing considered, strategic feedback and feed-forward.

2.1 Written Assessments

The first written assessment is a 'Literature Report', which is scaffolded and structured in two main ways: there are prescribed headings with guidance about what should be written under each heading, and there is a small set of prescribed,

accessible readings. The aim is for students to understand those readings and extract the most salient content to construct their assignments. There is a small weight (10 %) for 'referencing and readability' in the marking rubric, and the rest of the marks are allocated to content under the prescribed headings and a 'reflection' (see Sect. 2.2). The second written assignment is a structured 'Learning Journal' based on tutorial learning activities. This assessment asks students to relate the concepts and theories they engage with in tutorials to their own 'social world'. There is a small weighting for 'referencing and readability' (5 %) with emphases placed on the 'concepts and evidence' and 'relevance to my social world'.

As students are not in tutorials every week, an online tutor provides support, wrapped around assessments, through asynchronous discussion forums, assisting them to think through concepts and assessments and connecting them to relevant resources to facilitate completion of assessments (see Griffin et al. 2009 for a description of the online tutor).

2.2 *Feedback and Feed-Forward: Reflections on Learning*

The subject is highly rated by students, peers and learning advisors, and as part of the school's first-year retention strategy, it was shifted in 2015 from second (spring) to first (autumn) semester. This was an opportunity to further develop the assessment dialogues with tutors and students to encourage engagement with writing academic assessments as a process, rather than just content (Carless 2006).

Feedback and feed-forward are variously conceptualized. For Koen et al. (2012), feed-forward closes the gap between areas of deficiency and expected performance. Strategic feed-forward requires an holistic review of a student's writing to identify two or three areas for improvement, rather than the atomized feedback added to a rubric (although the rubric of itself is important feedback). Areas for improvement will depend on the student's current performance. For example, feed-forward for a student who has not fully answered the assessment questions or who has problems with English expression at the sentence level will require a different approach than for a student who has achieved a high mark against the criteria and standards. Examples of academic writing feedback from the comment bank that are calibrated across a range of capability levels are provided in Sect. 4.1.

The current assessment innovation is the opportunity for students to reflect on their written work. To encourage participation, marks were awarded in the rubric for providing a 'thoughtful' reflection. The first reflection is part of the Literature Report, the first written assessment students attempt, and described in the Learning Guide (syllabus) as: 'Reflection: Write a sentence (or two) about what you have learned about academic writing and, in particular, this type of literature report'. Having received feedback and feed-forward on their Literature Report, the reflection for the Learning Journal was as follows: 'Reflection (1) What did I learn from completing the Literature Report that I am now applying to writing my Learning Journal? (2) How am I applying the feedback I got on the Literature Report to the

Learning Journal?' The scaffolding of the reflections was aimed to encourage students to consider their feedback and follow up feed-forward suggestions, and also to allow markers an insight into how their guidance was interpreted and understood.

3 Assessment Dialogue 1: Between the Subject Coordinator, Academic Learning Advisor and Students to Improve the Learning Guide

The Learning Guide (syllabus) is the first dialogue with students, and it is important to get this right as all other practices hang off this document. The Learning Guide has been iteratively refined, particularly related to communication of assessment requirements, through reviews by internal and external learning advisors and a student panel. As part of the innovation reported here, the school's academic learning advisor assisted with placing the two written assessments in their respective genres and providing explicit explanations and adjunct resources for these. Detail was provided on different writing approaches relevant for the more formal Literature Report and on scholarly reflective writing for the Learning Journal. In consultation with the subject coordinator, the academic learning advisor also designed and adjusted corresponding learning activities in the tutorial workbooks.

The Learning Guide and the associated marking rubrics are not simply launched to students. Learning activities during tutorials are designed to engage students with the assessment tasks as described in the Learning Guide and the rubrics against which they will be marked. The online tutor is available for just-in-time support for assessments and to assist students to navigate and apply the Learning Guide and rubrics. These strands of assessment dialogues feed into student support for learning as well as assessment adjustment and design.

4 Assessment Dialogue 2: Between the Academic Learning Advisor, Tutors and the Subject Coordinator to Construct a New Comment Bank

The purpose of this assessment dialogue between the academic learning advisor, tutors and subject coordinator was to develop a feedback and feed-forward comment bank for online marking to promote consistency and clarity across markers and to make high volume marking more manageable. The effective feedback/feed-forward strategy was designed to complement the redesign of assessments to include student reflections. The objectives were to support tutors to provide effective feedback and feed-forward by collegial development of a custom-designed

comment bank for online marking of students' first piece of written work, the Literature Report, and providing professional development for tutors in best practices and critical discussion of the new comment bank.

Comment banks had been previously developed and used to align with the marking rubric in the subject, and there was scope for markers to create their own comments. The academic learning advisor reviewed the comments used by tutors to mark the Literature Report in 2014 along with the assessment materials and then analysed students' requests for remarks and the responses of the senior tutor (who works closely with and supports the subject coordinator) provided further elaboration of the criteria and standards. This analysis highlighted where feedback comments were not clear to the students.

Students who requested remarks had to engage with the marking rubric and make a case that they deserved more marks aligned with the criteria and standards. While there were relatively few of these, students reported confusion with vague comments. This analysis showed that students were correct and comments were vague at times, sometimes hard to understand, and did not give helpful steps/tips to improve work. The review of the comments used by tutors also showed misalignment of rubric standards with well-meaning comments (e.g. 'great' or 'well done' that do not align with standards from 'unsatisfactory' through to 'outstanding').

The academic learning advisor and senior tutor drafted a new comment bank informed by these analyses. The subject coordinator and two other experienced tutors provided input as the comment bank went through several development iterations. These dialogues combined disciplinary and content knowledge, writing development and educational theory knowledge, tutor insights into their students' written work and students' interpretation of feedback (via the remark requests).

The comment bank was constructed to model feed-forward and feedback on writing and content aligned with the marking criteria and standards. It was discussed and applied by the whole teaching team at a moderation meeting chaired by the subject coordinator, which included a presentation by the academic learning advisor on the best practice principles used to frame the analysis and findings.

The tutors, who have been encouraged to contribute to the continuous development of the subject (see Griffin and Burnett 2014), were engaged in constructive dialogue on the comment bank. Some had developed their own comments, which they were motivated to maintain and saw as more personal and individualized. They agreed that the content comments could be a reference point, rather than a prescription. Tutors were positive about the writing domains comment bank and agreed to use the text from the bank, but to edit to suit their own style of marking. Tutor feedback informed the next iteration of the comment bank, and 'stem' responses were developed, which were specific enough to give students the steps to improve, but which tutors could add their own stamp should they wish to. An analysis of the process of developing a comment bank shows that tutors are concerned with developing positive, personalized feedback.

4.1 Examples from the Comment Bank

The comment bank provided feedback and feed-forward examples that emphasized as follows: language that is congruent with the assessment instructions and marking rubric, especially task verbs; concrete ways a student can improve; a dialogue between tutor and student ('*You should not*' becomes '*A valuable tip I can give you is*....'); comments and marks that are commensurate with students' experience levels; and focus on different ways students can express their learning.

One assessment criterion for the Literature Report asked students to provide a 'summary of the main points …'. The academic learning advisor found there were cases of feedback that were not congruent with this criterion and where students were marked down for not having a 'developed discussion' or 'discussing in an analytical way' or for not making comparisons and contrasts. A common but vague feed-forward comment was 'needs development' or 'needs improvement'. This is not strategic, targeted feed-forward as it does not specify what needs improvement, nor tells the student how to go about improvement.

Through a team discussion of the analysis, it was apparent that some tutors expect quite high levels of sophistication in first-year student writing in comparison with the demands of the assessment outlined in the learning materials. This expectation implies that students will apprehend the tacit dimensions of assessment instructions and 'good' academic writing. High-level expectations are indicated in the allocated marks, against the standards, for the first written assignment; only two per cent of students achieved a high distinction (i.e. 85+ %).

Requests for remarks from the previous semester were also illustrative of a misconnection between the tutor's comments (and the criteria and standards) and the students' interpretation of the feedback. We know from student communications that those who formally request a remark underrepresent the number of students who are dissatisfied with their mark. This example below shows a mismatch between feedback and the assessment standards motivated by the tutor wanting to be positive and encouraging.

> For the introduction the marker gave me excellent feedback stating "great" "well done". I was told it was a very well developed definition of the levels of analysis. The marker also stated "great introduction"… However according to the rubric cube for my introduction I received 17% stating it was **Satisfactory**. So it does not coincide with the feedback, as she did not give me any negative feedback at all in the comments regarding the introduction.

Another student reacted strongly to a vague comment that provided no guidance regarding improvement.

> In this section I too have summarised the various effects of video game violence and television with a comparison between the two, to receive a comment such as 'this needs to be further developed'. Well thanks, that was a very developed 'constructive' comment. I absolutely covered all the necessary information required here … and seemed to have gotten a lower mark than the column that states 'missed a key point e.g. subtle effects'. THIS MAKES NO SENSE?

The student below made a case for a different standard based on the examples provided.

> …My marker advised I had some issue's with the example's I used in the content … however I did use all the examples we went through in class, and I was told in the feedback tutorial I had really good example's. For example … According to the literature I thought it was correct along with […'s] feedback. I feel for this I at least deserve … the **good** category.

The senior tutor responded to each of the remark requests in a considered, positive manner addressing the student's concerns and arguments and clarifying either why they got their original mark, or agreeing to additional marks.

The new comment bank consisted of 'content' comments related explicitly to the marking criteria and 'writing skills' comments developed by the academic learning advisor. The writing skills comments were under six headings with links to academic writing resources: referencing; basic grammar; summarizing; register/formal tone; structure–sentence level; and structure–paragraph level. An example from the each of the first three categories is in Table 1.

Further analyses of how the tutors used the comment bank will be conducted. The dialogue itself involving the academic learning advisor, tutors and the subject coordinator was constructive, and tutors were much more attuned to the traps of vagueness and non-alignment of comments with the marking criteria and standards and how students react to these. This was also part of ongoing professional

Table 1 Examples of academic writing comments

Category	Comment bank item
Improving paraphrasing	Paraphrasing means writing about another author's ideas in your own words and also providing the referencing details of where those ideas come from. This is important to avoid copying (plagiarism) but also because your reader is interested in your understanding of what the author is saying. So, it is not enough to change some words around, and it is about writing what YOU think the meaning of the text is
	There are some excellent resources for referencing and paraphrasing on the 'SSAP Academic Literacy Support—Students' vUWS[a] site. You will find this on your list of sites when you log into vUWS
Basic grammar	**Best practice advice to markers**: overwhelming the student with pages of grammar corrections is not effective for students or your time. It is enough to micro-edit one sentence or paragraph and then advise the student to look through the rest of their text for where they could apply the remedies, or point them to a resource. Major problems with grammar and expression can be broached generally in the final overall comment and resources linked there
Summarizing	Summarizing complex academic texts takes practice, but is a key skill so worth putting some time into. The trick is to make judgements about what are the main ideas and what are details that are not necessary for the 'big picture'. The Unistep Guide has excellent resources for practicing as summarized in Chap. 9

[a]vUWS is the institutional name for the Black Board Learning Management System

development for the tutors and subject coordinator on working with students on their academic writing, not just the discipline content, as illustrated by the comments presented in Table 1.

5 Assessment Dialogue 3: Between Tutors and Students Through Written Reflections and Feedback/Feed-Forward Comments

For Carless (2006), the most important dialogue is between tutors and students via feedback on assessments. Likewise, Jonason (2012) advocates for an active dialogic model of feedback. One indirect indicator of feedback effectiveness is student satisfaction. A major graduate survey in Australia (Scott cited in Gill 2015) concluded that assessment (standards, marking, expectations, management and feedback) was a recurrent 'hot spot' and consistently received a high ratio of 'needs improvement' to 'best aspect'. The most relevant institutional 'Student Feedback on Unit (Subject)' (SFU) items directly related to assessment are as follows: [Assessment Feedback]—*I was able to learn from feedback I received in this unit*; and [Assessment Guidelines]—*There were clear guidelines for all assessment tasks in this unit*. The feedback question is consistently the lowest scoring item on the SFU. The case study subject does well on this item compared to the university and the school, indicating that students appreciate the tutors' efforts to provide considered, personalized feedback (2015 results: subject, 4.4/5; university, 3.9/5; school 4.0/5). A search for 'feedback' in the 2015 SFU qualitative comments yielded both 'best aspects' and 'needs improvement' as set out in Table 2.

Table 2 Student feedback on unit: qualitative comments on feedback

Best aspects	Needs improvement
The feedback provided for assessments and the guidance given for each in the learning guide	Feedback on assessments could be more in depth to tell us how what we did wrong can be improved for next time
The feedback on the Literature Report was good	Better overall feedback on assessments to clearly state how improvements can be made …
… the academic staff were very friendly, encouraging and gave excellent feedback	The harsh comments on the literature report. I found that my comments/feedback wasn't very encouraging—told me everything I did wrong and nothing I did right
The learning guide is very detailed which is extremely helpful! Also teacher was an excellent tutor and gave amazing feedback on assessments	My tutors feedback on assignments was not helpful at all as she did not tell me how to improve … I was left in the dark about how to improve
Feedback received from assessments/assignments were informative and helped with my learning	I feel sometimes the feedback is not quite clear by their understandings rather than my understanding of this unit
Great feedback on assessments. Assessments were explained and guided in a very productive way	

Although the ratio of 'best aspects' to 'needs improvement' was close to 2:1, there were 6:5 related to feedback. It is interesting that the 'needs improvement' comments show that students want to know how to improve, suggesting that even those who were dissatisfied had a sophisticated understanding of feed-forward.

A preliminary analysis focused on students' perceptions of their feedback as expressed in their reflections. A sample of fifty sets (about 5 %) of reflections on the Literature Report and Learning Journal were collected: ten at each standard (fail, pass, credit, distinction and high distinction) and across all markers. The purpose of the analysis was to identify emerging themes of what students wrote about in their reflection dialogues. Data analysis proceeded by deriving coding domains from the reflections and specific coding categories through an initial sweep of the data in Microsoft Excel and further analysed using NVivo10. The number of mentions against each coding category was recorded. Some of the findings are discussed next.

There was a high frequency of meta-language about academic writing across all grade groups (with more nuance at higher levels). For example,

> I learnt to … understand the criteria of academic writing. It improves my sentence structure, paraphrasing and paragraphing. (F)

> While writing this literature report, I have learned how to write topic sentences, before this I did not know how powerful a topic sentence could be. I also learned how to reference in text, prior to this I basically knew nothing about in-text referencing. (C)

> I have learnt that it demonstrates more knowledge on your behalf to the marker/reader if you paraphrase rather than using direct quotes because this shows that you have thoroughly understood the concept that you have just read (HD)

This is significant for first-year students because developing a vocabulary for writing equips them to improve over time, rather than simply address the assessment task at hand. There is also evidence of genre awareness (Clark and Hernandez 2011), which means that students are well placed to further develop their academic writing by recognizing differing structures and purposes of academic text types. Genre awareness mirrors the genre placement of their assessments in the Learning Guide, indicating that the inclusion of student reflections, along with comprehensive feedback and feed-forward related to academic writing, had a positive impact.

For the Literature Report, the most prevalent code categories aligned with its design features. Sticking to the 'structure' and keeping to the word count 'concision' had the most mentions and is important academic writing skills to develop. Related comments clearly indicate that the scaffolded task design influenced the way students approached their writing.

> Academic writing is different to normal because of the way it is set out. With this report it worked well because it laid out the steps before joining them together. (F)

> I have really expanded on my skills on how to properly use academic texts and implement them, as well as having a clear format and structure. The literature report style of this assignment has taught me this. (D)

The technical aspects of referencing, as opposed to commenting on integrating evidence in the text (a higher order skill), were more prevalent in the marker feedback on both assessments. In the sample of Learning Journal reflections, there were 46 technical aspects of referencing (in-text or reference list) cited by students compared to only 3 mentions by students in their first Literature Report reflection, written before receiving feedback. This perhaps indicates marker preoccupation with the mechanics of referencing rather than with integrating evidence from sources. The number of mentions (15) of 'follow instructions' in student's reflections on marker feedback may relate to markers' assumptions about the transparency of written/assessment language. That is that if students are diligent and read the learning materials and instructions, they will somehow understand and be able to apply them. This is a common overestimation of the process and is proposed in the literature as major barrier to student learning (cf. Haggis 2006; Lea and Street 1998; Lillis 2001; O'Donovan et al. 2004).

The gathered reflections allowed an analysis of individual students' progression from the first to second written assessment. Here is one such example where the student focused on content for the first assessment reflection and academic writing for the second. This is the sort of development we are keen to promote for all students.

> In writing this literature review, I expanded my knowledge on how to critically analyse the topic of the Levels of Analysis to Understand the Influence of Violent Media on Aggression. It allowed me to compare, summarise, classify and evaluate the different ideas and research within the prescribed readings. (P)

> Through the feedback on my literature report I became aware of many grammatical errors that I was making, mainly through using contractions. This will allow me to be more professional in my writing style and improve how I review my work…Throughout my literature report, I did not in text reference correctly and thus got marked down for referencing, I aim to improve on this throughout my learning journal to better my grade on referencing. (D)

6 Conclusions

Knight and Yorke note that 'although many teachers give a lot of feedback on specifics, it is general feedback that has the greater power to stimulate learning' (cited in Carless 2006, p. 225). It is relatively easy to do corrective marking, picking up errors at the sentence level and referencing mistakes. Within a discipline, it is also relatively straightforward to comment on content. Although students need this feedback, it is feedback and feed-forward for learning that requires a more holistic review of a student's work and strategically targets where and how they can enhance their scholarly writing. Working against this is that online marking promotes volume of comments and markers in their enthusiasm can overwhelm students with feedback.

Providing effective feedback and feed-forward takes knowledge about academic writing, a motivation to inculcate it in students and skills in assessment and feedback and feed-forward. In our experience, tutors (who are all sessional) are willing partners in assessment dialogues, but are challenged by the volume of marking and diversity in the quality of students' written work, given the widening participation agenda. The authors, who are in turn, the subject coordinator, two academic learning advisors and the senior tutor, are working to improve the quality and consistency of feedback and feed-forward in the subject described herein. One academic learning advisor is also working at the school level to embed academic literacy within other subjects, which includes subject coordinator and tutor professional development, and runs workshops on this topic for tutors. We are steadily gaining traction as demonstrated, for example, in the student evaluations of this subject and the various levels of ongoing dialogues.

The bigger picture is to align assessment design and feedback and feed-forward to the first-year transition agenda. We are confident that a significant number of students are benefitting, in different ways, from the range of assessment scaffolds and supports and that students' written reflection on their learning is an important way of collecting their dialogues as a form of 'voice' in teaching team discussions. Our next practical step will be to ask tutors what they learned from student reflections on their feedback to begin a new cycle of assessment dialogues.

References

Bell, A., Maladenovic, R., & Segara, R. (2010). Supporting the reflective practice of tutors: What do tutors reflect on? *Teaching in Higher Education, 15*(1), 57–70.

Burnett, L., & Larmar, S. A. (2011). Improving the first year through an institution-wide approach: The role of First Year Advisors (FYAs). *International Journal of the First Year in Higher Education, 2*(1), 21–35.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education, 31* (2), 219–233.

Clark, I., & Hernandez, A. (2011). Genre awareness, academic argument, and transferability. *Writing Across the Curriculum Journal, 22*, 65.

Gill, B. (2015). Talking about the elephant in the room: Improving fundamental assessment practices. *Student Success*, 6(2).

Griffin, T., & Burnett, L. (2014). An online community of practice to enhance small group learning and teaching. In T. Bastiaens (Ed.), *Proceedings of e-learn: World conference on e-learning in corporate, government, healthcare, and higher education 2014* (pp. 729–735). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE). Retrieved August 19, 2015 from http://www.editlib.org/p/148806

Griffin, T., Gilchrist, A., & Thomson, R. (2009). The online tutor: Good practice in a large-enrolment, blended learning unit. In *Same places, different spaces. Proceedings of the Australasian Society for Computers in Learning in Tertiary Education* (*ASCILITE*), Auckland, December, 2009.

Haggis, T. (2006). Pedagogies for diversity: retaining critical challenge amidst fears of "dumbing down". *Studies in Higher Education, 31*(5), 521–535.

Jonason, A. (2012). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education, 14*(1), 63–117.

Kift, S. (2009). *Articulating a transition curriculum to scaffold and to enhance the first year student experience in Australian higher education.* Final Report for ALTC Senior Fellowship Program. ALTC: Strawberry Hills, NSW.

Koen, M., Bitzer, E. M., & Beets, P. A. D. (2012). Feedback or feed-forward? A case study in one higher education classroom. *Journal of Social Science, 32*(2), 231–242.

Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education, 23*(2), 157–172.

Lillis, T. M. (2001). *Student writing: Access, regulation, desire*. New York: Routledge.

O'Donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education, 9*(3), 325–335.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd Ed). University of Chicago Press: USA.

Universities Australia. (2013). *University students finances in 2012: A study of the financial circumstances of domestic and international students in Australia's universities.* Centre for the Study of Higher Education, University of Melbourne. Retrieved from: http://www.universitiesaustralia.edu.au/resources/272/1622

Wingate, U. (2006). Doing away with 'study skills'. *Teaching in Higher Education, 11*(4), 457–469.

Part V
Other Issues in Assessment for Learning

Title: A Meta-Analysis of the Effects of Instructional Tasks on L2 Pragmatics Comprehension and Production

Nour El Imane Badjadi

Abstract One of the vital aspects of instructed L2 pragmatics is the effect of instruction on comprehension and production. As such, research is needed to investigate the effects of different instructional designs on L2 pragmatics development. This paper reviews experimental and quasi-experimental studies to examine the differentiated effects of second language (SL/L2) pragmatics' instructional tasks in relation to comprehension and production outcome measures. After an exhaustive and systematic search, 24 studies met the inclusion criteria and were analyzed to extract data involving a total of 1508 participants. The participants comprised of learners receiving instructional interventions on several L2 pragmatics forms and functions. Across different contexts, proficiency levels, methodological nuances, and target pragmatic constructions, instructional tasks were classified into meta-pragmatic discussion and input-based tasks, with or without the provision of feedback, while outcome measures were classified into comprehension, structured production, or free production measures. The results reveal that, in accordance with instructional tasks, comprehension and production mean effect sizes diverge from small to large. In addition, interventions which actively engaged learners in meta-pragmatic discussion or provided recasts for production in input-based tasks produced larger effects. The meta-analysis yields significant insights into the role of knowledge representation and processing instruction in the acquisition of L2 pragmatics. To ensure even acquisitional pragmatics development, it is important to continue to investigate the roles of feedback and meta-pragmatic knowledge. Further implications for enhancing comprehension and production aspects of L2 pragmatics instruction and development are discussed, and avenues for further research are highlighted.

Keywords Instructed L2 pragmatics/L2 pragmatic competence · Comprehension/production tasks · Skill acquisition theory (SAT) · Adaptive control of thought (ACT) model

N.E.I. Badjadi (✉)
Central China Normal University, Wuhan, China
e-mail: nourelimane78@gmail.com

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_21

1 Introduction

To date, research has shed light on various perspectives of L2 pragmatics instruction. One reason for this enduring interest is the deep roots of pragmatic competence in the social conventions of the SL and coupled with its problematic nature, especially if developed under the input scanty conditions of the formal context (e.g., Kasper and Dahl 1991; Kasper and Schmidt 1996; Kasper and Rose 2001, 2002; Rose and Kasper 2001; Rose 1994, 1997, 2005; Kasper 2001a, b, 2009; Bardovi-Harlig 1999, 2001, 2013). However, after the field has overcome the controversy regarding the amenability of pragmatic competence to instruction, there has been a mounting need for research aimed at synthesizing the findings of research studies that focus on its development in L2 classrooms. Heading beyond the teachability of pragmatic competence, this study is intended to contribute to the ongoing discussion on the issues framing its instruction and development.

As a growing area within the field of SLA, acquisitional pragmatics was the focal point of several narrative reviews and a meta-analysis. For instance, Rose (2005), in a review article on instructed L2 pragmatics, identified three main trends of research in the field: the usefulness of instruction in comparison with exposure, what pragmatic features are teachable in a classroom context, and whether different instructional methods bring forth different results. In his discussion of the first two axioms of research respectively, Rose pointed to the influence of outcome measures and the complexity of learning targets in relationship to the sufficiency of instruction provided in terms of length and intensiveness. Though the third point of discussion was concluded with denoting inconclusive evidence provided by the reviewed studies, it fairly headed in the direction of supporting Schmidt's noticing hypothesis and, particularly, the benefits of explicit meta-pragmatic rule presentation and meta-pragmatic discussion, a finding which clearly points to the essential role of particular instructional tasks.

Likewise, Jeon and Kaya (2006)'s meta-analysis reviewed the literature for studies that examined the effectiveness of instruction in promoting L2 pragmatics. The study shed light on axial issues concerning second language pragmatics instruction. Jeon and Kaya's results provided evidence for the usefulness of instructional tasks derived from SLA theory for the development of L2 pragmatics. In addition, their findings were analogous to Rose's (2005) review in that instruction is advantageous over non-instruction and that explicit instruction is likely to be more effective than implicit instruction. The researchers further conclude that the type of outcome measure may influence the assessed impact of instruction. In specific terms, studies which used both natural and data-elicited outcome measures tend to have more robust effect sizes. They also found that long-term instruction probably leads to more evident instructional effects.

In the same line of thought, this meta-analysis reviews interventional studies that have investigated L2 pragmatics instruction in formal settings in order to examine the effectiveness of instructional tasks on the development of L2 pragmatics comprehension and production. Meanwhile, Jeon and Kaya's study reviewed

studies published until 2003 and searched the need for teaching L2 pragmatics, the methods of teaching it, the measures of assessing it, and the adequate duration of instruction, and the present study builds on and expands beyond it by synthesizing findings across studies published between 2003 and 2013 and by examining whether different instructional tasks result in dissimilar degrees of effectiveness and whether the development of SL pragmatic comprehension and production is variedly affected by instructional tasks.

This research is also relevant because it addresses the controversial findings yielded by previous synthetic reviews on the effects of pragmatics instruction. Reviews on L2 pragmatics (Kasper 2001a, b; Rose 2005) and the meta-analysis (Jeon and Kaya 2006) suggest that explicit instruction is generally more effective than implicit instruction. However, Rose (2005) noted that the findings of comparative studies were inclusive. Similarly, Jeon and Kaya (2006) noted that their results regarding the effectiveness of meta-pragmatic discussion should be interpreted with caution, given the small size of empirical studies included in their meta-analysis. A research gap that has not been addressed in previous reviews is a task-based coding and categorizing of primary studies. Given the controversy, the synergy between explicit and implicit knowledge in reality, and the impracticality of the explicit/implicit dichotomous distinction in practice, a task-based categorization of L2 pragmatics instruction would be more fruitful for theory, research, and practice.

This paper presents an initiative to address this gap through a discussion of the issues associated with designing and implementing pragmatic tasks to raise SL learners' comprehension and production of the socio-pragmatic and pragmalinguistic aspects of the SL. Besides, a feature characterizing the current trends in second language (SL) teaching is engaging learners in a variety of learning tasks (Skehan 1996, 1998; Bygate et al. 2001; Crookes and Long 1992; Swain and Lapkin 2001; Ellis 2000, 2003; Nunan 2004; Shehadeh 2005). One of the fundamental issues in instructed L2 acquisition is learning tasks' design and its effects on L2 pragmatics development (Bardovi-Harlig 2013). As such, this study is important since it not only heads beyond the dichotomous explicit/implicit distinction but also adopts a task-based classification of instructional interventions.

Moreover, there is considerable evidence suggesting that the observed effects of instruction are influenced by outcome measures (Rose 2005; Jeon and Kaya 2006; Safont-Jordà and Portolés-Falomir 2013). Therefore, this study is important because it examines whether instructional treatments are likely to differentially affect L2 pragmatics comprehension and production. Highlighting the need for cognition-oriented L2 pragmatics research, Kasper (2001a) notes "… there is a continued need for information processing theory. In their focus on the "higher" cognitive functions, sociocultural studies of pragmatic development have not had much to say about attention, memory, representation, restructuring, transfer, and speech processing in comprehension and production" (p. 525). Since the cognitive aspect is one of the prevailing perspectives on L2 pragmatics development, this meta-analysis would contribute to a deeper understanding of the relative effects of different instructional tasks on SL learners' comprehension and production development.

2 The Present Study

This study is aimed to examine the task design factors which influence the observed effects of instruction as well as to take a further step by examining the influence of outcome measures from the cognitive perspective of L2 pragmatics comprehension and production. In accordance with the purposes behind carrying out the present synthesis, this paper focuses on two dimensions: instructional tasks (meta-pragmatic discussion, input-based tasks, and feedback) and outcome measures (comprehension tasks, structured production tasks, or free production tasks). Research questions are, therefore, formulated as follows:

1. Do different instructional tasks variedly affect the effectiveness of L2 pragmatics instruction?
2. Are there any differences in the effects of instructional tasks on L2 pragmatics across comprehension and production outcome measures?

The research questions are addressed through two main steps. The first phase involved scheming and scanning primary studies for relevant information, and then categorizing them in accordance with their theoretical perspectives, the instructional procedures employed to operationalize it, and the data-eliciting measure(s) used to assess the effect of treatment. The second phase was carried out through summarizing and meta-analyzing the research traits and results of primary studies.

3 Method

3.1 Search Strategies

The present meta-analysis is based on exhaustive electronic and manual searches to locate full-text papers on L2 pragmatics instruction published between 2003 and 2013.

The electronic search covered the databases: EBSCOhost which is a comprehensive database encompassing ERIC and PhycINFO, two of the most commonly used databases in applied linguistics (Oswald and Plonsky 2010), and LLBA. The databases were searched using Boolean operators of the keywords: second (foreign) language, pragmatics instruction, interlanguage pragmatics development, teaching methods, pragmatics performance measures, and instructed SL learning/acquisition. This full-scale search was followed by a focused search using narrower terms closely related to the purposes of the present meta-analysis such as explicit, implicit, meta-pragmatic, input enhancement, consciousness raising, awareness, noticing, discourse completion tasks, and role play tasks.

The manual search covered over 10 of the widely cited journals in the field of applied linguistics and second language acquisition. The searched journals are Applied Linguistics, Annual Review of Applied Linguistics, Foreign Language

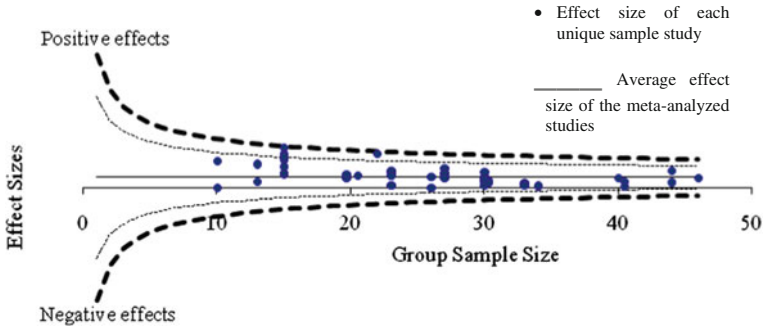


Fig. 1 A funnel graph illustrating effect sizes plotted against group sample size

Annals, International journal of Applied Linguistics, Language Awareness, Language Learning, Second Language Research, Studies in Second Language Acquisition, System, Journal of Pragmatics, and Intercultural Pragmatics journal. Data collection is also involved using Google Scholar to search foreword citations of Kasper and Dahl (1991), Kasper and Schmidt (1996), Bardovi-Harlig (1999, 2001), and Kasper and Rose (2001, 2002).

One of the challenges and choices a meta-analyst has to deal with is whether or not to include fugitive literature. The "file-drawer" problem is associated with the assumption that publications tend to accept studies yielding statistically significant results (Rosenthal 1994; Smith 1980). Because of the availability issue, that is, the difficulty to track and retrieve unpublished literature, meta-analysts often opt for excluding all unpublished studies from their synthesis (eg., Norris and Ortega 2000), as it is the case in the present meta-analysis. Other reasons for this decision is the difficulty of including a representative sample of unpublished literature and the fact that examining the study quality dimension, published versus unpublished studies, is not a focal point in this study. Nevertheless, since the problem of publication bias is well recognized in the literature of meta-analysis, forest and funnel plots can be used to examine publication bias (Rothstein et al. 2006). See Fig. 1.

3.2 Criteria for Inclusion

Drawing upon the research questions addressed in the present meta-analysis, the following inclusion/exclusion criteria are taken into consideration when filtering the initially collected studies:

1. The study adopts an experimental or quasi-experimental methodological design that allows for identification of instructional effects.
2. The dependent (learning target) variable is a language feature having a pragmatic communicative dimension (e.g., requests, refusals…) rather than a purely linguistic grammatical function (e.g., verb tense)

3. The measure of instructional effects is an assessment procedure acknowledged in ILP data-eliciting practice (PAJT, DCT, etc)
4. Participants are learners of a second or foreign language in the classroom context, i.e., second/foreign language is taken to mean a non-native language that is learned in a context where another language is the mother tongue (e.g., English for Malaysians, SL, English for Algerians, FL).
5. Tasks constituting the instructional treatment are adequately described in terms of pedagogic features congruous with the fields of L2 pragmatics, in particular, and SLA, in general (input-based activities, meta-pragmatic discussion, etc).
6. The study is written in English and published between 2003 and 2013.

Contrariwise, reasons for excluding studies are as follows:

1. It is an experimental design study which adopts a qualitative approach to the evaluation of learning (e.g., Takahashi 2005; Byon 2006; Usó-Juan 2013)
2. It is an ethnographic or observational study on acquiring a SL pragmatic feature or a review paper, or focuses on L2 pragmatics instructors' education and teacher preparation programs, or L2 pragmatics teaching curricula and materials (e.g., Vellenga 2011)
3. If two studies are conducted with the same participants and examine the effectiveness of the same features of instruction, only one study, where more thorough descriptive data are provided, is included. The following studies are, therefore, excluded: Fukuya and Martinez-Flor (2008); Takimoto 2006b, 2008a, b; Alcón-Soler and Guzmán-Pitarch (2013).

3.3 Coding

After the studies have been collected ($N = 24$), a coding scheme was designed to sum up the features of individual studies. Decisions regarding this step have been taken after reviewing relevant literature on instructed SLA and interlanguage pragmatics (Norris 2012; Norris and Ortega 2006, 2007, 2010; Ortega 2010; Alcón-Soler 2008; Oswald and Plonsky 2010), as well as meta-analysis-focused explanatory works (Borenstein et al. 2011; Cumming 2012; Cooper 2009; Lipsey and Wilson 2000). The coding of study features involved a reiterative process which encompassed two main stages. First, the initial coding of six representative studies (25 % of the studies identified) by two researchers includes the author. After each study was independently coded, the initial set of variables was discussed, disagreement was resolved, and a coding was developed (Table 1). The coding scheme was then used by the two researchers as a data collection instrument for all the studies included. Inter-rater reliability was then calculated for two variables:

Table 1 The coding scheme used to collect data from primary study reports

(A) Publication features				
Author				
Year				
Title				
Journal				
(B) Relevance features				
Research questions				
Target features				
(C) Participants and Context				
First language				
Target language				
Context				
Age				
Length of study				
Educational status				
Proficiency level				
Proficiency measure				
Setting laboratory or classroom:				
(D) Study design				
Total sample size				
Control group	Yes/no	Size:		
Pretest	Yes/no			
Experimental group(s)	Number of groups:	Size:	Explicit:	Implicit:
Posttest	Yes/no			
Delayed posttest	Yes/no			
(E) Substantive features				
Type of instruction	Explicit:		Implicit:	
Features of treatment	Meta-pragmatic rule explanation:	Consciousness-raising tasks:	Feedback as correction:	Feedback as question recasts:
Length of treatment				
Outcome measure(s)				
Reliability of outcome measure				
p value				
Means and standard deviations				
Effect sizes				

features of instruction (tasks and procedures) and outcome measures.[1] A detailed record of the included studies in terms of learner characteristics, study design, and instructional tasks is available in Appendices 1, 2, and 3.

3.3.1 Coding for Instructional Tasks

The first set of features to be coded are those delineating the type of instructional treatments through judging whether or not they involved meta-pragmatic discussion and input-based tasks, with or without feedback. In the present meta-analysis, coding the types of instructional tasks is based on the continuum nature of L2 instruction in SLA theory. On the one hand, meta-pragmatic discussion focuses on explaining the usage and use of pragmalinguistic and socio-pragmatic rules as acknowledged by the SL speech community. On the other hand, input-based tasks, such as form comparison, structured input, and input enhancement, are the communicative activities designed to raise learners' awareness of the target forms and functions embedded in input and to direct their attention to target forms and functions, often through exposure to SL socio-pragmatic and pragmalinguistic norms in context and comprehension questions; however, no meta-pragmatic information is provided. Inter-rater reliability computed through Cohen's kappa k for the type of instructional tasks was 0.89.

3.3.2 Coding for Outcome Measures

According to Kasper and Dahl (1991), comprehension is mainly measured through rating and multiple choice tasks, while production encompasses discourse completion, role plays, and authentic discourse tasks, respectively, in a continuum. Similarly, Norris and Ortega (2000) classified outcome measures into four categories, namely metalinguistic judgment, selected response, constrained constructed response, and free-constructed response. Likewise, in the present meta-analysis, outcome measures are classified according to the abilities they require participants to demonstrate to comprehension, structured production, or free production. In order to study instructional tasks' effects on L2 pragmatics comprehension and production, unique sample studies are categorized into comprehension, structured production, and free production outcome measures. First, comprehension outcome measures include multiple choice and pragmatic acceptability judgment tests since both tap on learners' ability to comprehend certain statements or communicative situations. Measures of structured production, represented by discourse completion tests (DCTs), require learners to move a step further by productively demonstrating their ability to understand and respond to a communicative scenario. Lastly, free production requires students to produce discourse in response to slightly

[1]The studies with an asterisk (*) in the references are the studies coded for this meta-analysis.

constrained prompts (for a detailed discussion of L2 pragmatics outcome measures, see Kasper and Dahl 1991; Bardovi-Harlig 2013). Cohen's kappa k for outcome measure was 0.92.

4 Analysis

In cases of multiple samples experimental studies contributing to a single study, each unique sample was considered as an independent experiment. Thus, overall, 50 unique sample studies are extracted from the 24 primary study reports.

4.1 Developing Contrast Categories

The quantitative results obtained from the included studies were aggregated through calculating Cohen's d for each unique sample study. To calculate contrasts for Cohen's d, unique samples are coded as either control or experimental groups. Cohen's d contrasts can be summarized as follows:

1. Contrasting each experimental group with the control group on the immediate posttest if the study reports data on one or more treatment groups and a true control group.
2. Contrasting posttest and pretest data for each experimental group if a study does not involve control group and reports pretest and posttest values on a dependent variable.

4.2 Calculating and Combining Effect Sizes

An inclusive strategy was adopted for calculating and combining effect sizes. First, for any study that employed different instructional treatments, each experimental condition was considered a unique sample study with independent effect sizes. This procedure is claimed "to provide the most representative picture" for the effectiveness of instructed SLA (Norris and Ortega 2000, p. 448). Effect sizes are also calculated for the identified unique sample studies on each of the outcome measures in order to examine the observed difference in instructional treatment effects elicited through different outcome measures. The rationale behind this decision is that there is evidence that outcome measures are likely to be a powerful moderating variable when assessing different aspects of instructional effectiveness(Safont-Jordà and Portolés-Falomir 2013; Martínez-Flor 2013, Fukuya and Martínez-Flor 2008; Jeon and Kaya 2006; Billmyer and Varghese 2000). Below is the summary of the method followed for effect size calculation and combination:

1. For studies contrasting the effectiveness of pragmatic discussion and input-based tasks, each experimental group is considered as a unique sample study with independent effect size(Fordyce 2013; Bu 2012; Nguyena et al. 2012; Hernández 2011; Nipaspong and Chinokul 2010; Ghobadi and Fahim 2009; Takimoto 2009; Martínez-Flor and Alcón-Soler 2007; Koike and Pearson 2005; Alcón-Soler 2005; Martinez-Flor and Fukuya 2005)
2. For studies examining the effectiveness of one type of instruction, whether meta-pragmatic discussion or input-based tasks, one effect size is calculated (Farahian et al. 2012; Jernigan 2012; Li 2012; Narita 2012; Chen 2011; Takimoto 2011; Alcón-Soler and Guzmán-Pitarch 2010; Eslami and Eslami-Rasekh 2008)
3. For studies comparing the effectiveness of two versions or subtypes of the same instructional type, an average effect size is calculated (Barekat and Mehri 2013; Takimoto 2012; Takimoto 2009; Takimoto 2006a)
4. For studies examining whether L2 pragmatics instruction is equally effective for participants with different proficiency levels (Chen 2011), or under different degrees of intensity (Li 2012), the groups are considered as one unique sample study and an average effect size is calculated.

Overall, effect size of 0.20 is considered to signify a small effect, of 0.50 signifies a medium effect, and of 0.80 indicates a large effect size (Cohen 1988). In addition to the average effect size, confidence intervals are computed to determine whether or not the null hypothesis can be rejected at $p < 0.05$ level, probability level. Confidence intervals indicate the robustness of the averaged effect sizes since higher significance is associated with narrower confidence intervals and vice versa. Hence, 95 % confidence intervals are calculated around mean effect sizes using a conservative random effect approach.

5 Results

5.1 Outliers

One of the threats to the reliability of the conclusions drawn from a meta-analysis is the presence of outliers. The latter are represented in substantial values which are likely to bias the averaged results. In the present synthesis, the mean effect size of L2 pragmatics instruction was calculated with and without outliers. The detection of outliers was performed using the combinations of Microsoft Excel functions (Dodge and Stinson 2010). Effect size values that fall more than two standard deviations apart from the mean are considered as outliers. With the inclusion of outliers, the overall effect of instruction on immediate posttests was 2.01; without outliers, the average effect size turned to be 1.63. To achieve higher reliability, outliers were excluded from the meta-analysis. The identified outliers include Farahian et al. (2010)'s study ($d = 10.33$) and the findings based on the

comprehension outcome measure in Alcón-Soler (2005)'s study (experimental group1, $d = 6.02$; experimental group2, $d = 7.28$). Note that unique sample studies rather than whole study reports were excluded from the quantitative analysis so that thorough and more accurate data are available for the meta-analysis. Effect sizes of unique sample studies are provided in Appendix 4.

5.2 Publication Bias

Given that meta-analysis involves summarizing quantitative evidence from multiple studies, publication bias can be a threat to meta-analytic results. In the absence of publication bias, a wide distribution of effect sizes coincides with lower sample sizes and narrows down as sample size increases approximately drawing a shape of a funnel. On the contrary, in the presence of publication bias, no studies would fall in the proximity of effect sizes of zero where nonsignificant findings are likely to be located. To examine bias in the studies being synthesized, unique sample studies ($n = 47$), after the exclusion of outliers, have been plotted against the average sample size of study reports; the results are illustrated in Fig. 1.

As displayed in the graph above, effect sizes intensively cluster around the mean effect size of ILP instruction, $d = 1.63$, which is considered to have large magnitude and suggests over one standard deviation of difference between the experimental and control groups, which implies that most studies reported positive effects of treatment regardless of sample size. More importantly, the figure reveals that a number of effect sizes are located in the vicinage of null or negative area of the funnel graph indicating that research studies reporting nonsignificant effects have been included in the meta-analysis and, thereby, the absence of publication bias.

5.3 The Differential Effectiveness of Instructional Tasks

On the whole, the results reveal that the two types of instructional tasks resulted in strong effects of instruction. The standardized effect sizes of both types of instruction indicate huge effect of instruction (meta-pragmatic discussion/MPD, $d = 1.96$; input-based tasks/IBTs, $d = 1.01$). In order to explore the factors contributing to the overall effectiveness of L2 pragmatics instruction, features of instructional tasks reported in the studies being meta-analyzed have been coded and categorized. By so doing, subgroup instructional treatments adopting either meta-pragmatic discussion or input-based tasks are formed depending on the presence or absence of meta-pragmatic discussion/explicit rule explanation (MPD) and input-based tasks (IBTs), with or without feedback in both cases. The rationale behind this subgroup analysis is that two features outstandingly characterized meta-pragmatic discussion in the study reports: meta-pragmatic discussion (MPD) and correction feedback (F),

Table 2 Average effect sizes of the effectiveness of meta-pragmatic discussion and input-based tasks, with and without feedback

Features of instruction	Meta-pragmatic discussion +/− feedback		Input-based tasks +/− feedback	
	MPD + F (k = 7)	MPD − F (k = 15)	IBT + F (k = 10)	IBT – F (k = 8)
Average d	2.03	1.86	1.14	0.84
SD	1.38	1.1	0.72	0.52

MTD meta-pragmatic discussion; *IBTs* input-based tasks; *F* feedback; *k* number of unique sample studies

while recast feedback (F) is more often used with implicit groups. The categories and their mean effect sizes are illustrated in Table 2.[2]

Table 2 shows the average effect sizes of meta-pragmatic discussion and input-based tasks. On the one hand, learning tasks which bring together meta-pragmatic discussion and feedback (MPD + F: d = 2.03, SD = 1.38) seem to be more constructive than meta-pragmatic discussion without feedback. (MPD − F: d = 1.86, SD = 1.1). Interestingly, however, feedback does make observable difference in instructional treatments using input-based tasks (IBT + F: d = 1.14, SD = 0.84; IBT − F: d = 0.84, SD = 0.52). Further discussion on the roles of meta-pragmatic knowledge, input-based communicative activities, and feedback will be discussed in the next section.

5.4 *Effectiveness of Instruction Across Comprehension and Production Outcome Measures*

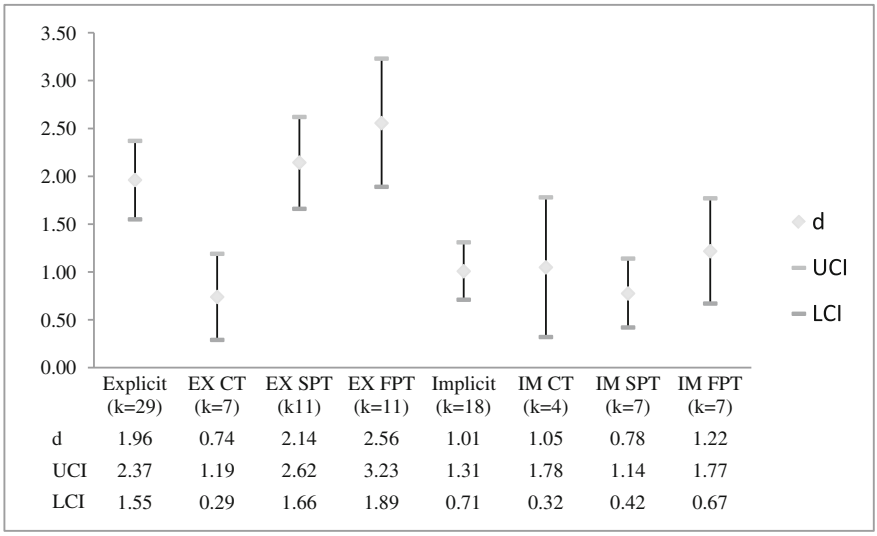
Unique sample studies for each of instructional treatments using meta-pragmatic discussion or input-based tasks were coded according to the outcome measures used (Appendix 4).

To investigate the effect of instruction on different aspects of language learning and use, average effect sizes are calculated for the two types of tasks. Interestingly, 50 % of the study reports used a combination of outcome measures (Appendix 3). First, comprehension outcome measures reported include pragmatic acceptability judgment tasks (PAJs) and multiple choice tests (MCTs). Second, tasks designed to elicit structured production are mainly discourse completion tests (DCTs), which were administered in the study reports in the open-ended written form. Noticeably, DCTs were most commonly used for measuring L2 pragmatics performance.

[2]The study reports Takimoto (2006, 2009, 2012) and Farahian et al. (2012) have been excluded from this subgroup meta-analysis because of reporting one mean effect of instruction for the MTD ± F experimental groups.

Finally, free production measures were the most varied and involved an array of oral and written activities such as written production in two genres (Fordyce 2013), oral discourse production (Narita 2013; Hernández 2011; Alcón-Soler and Guzmán-Pitarch 2010), oral peer feedback (Nguyen et al. 2012), e-mails and phone calls (Martinez-Flor and Fukuya 2005), and, definitely, the traditional role play tasks. An important implication of this variety is the growing integration of L2 pragmatics with other aspects within the field of SLA, such as genre-based, cooperative, and English for specific purpose approaches. The present meta-analysis falls in this same line of thought by examining how meta-pragmatic discussion and input-based approaches to instruction contribute to the development of acquisitional pragmatics' comprehension and production. The results come in full circle in the illustrative Fig. 2.

As shown in Fig. 2, meta-pragmatic discussion is overall more effective than input-based tasks with an average effect size difference of 0.95. The results also demonstrate that meta-pragmatic discussion's effect as measured through free production overtakes the strongest effect size (MPD × FP: $d = 2.56$). However, input-based tasks' effect size is more significant than meta-pragmatic discussion's effect size on comprehension tasks (IBT × C: $d = 1.05$; EXCT: $d = 0.74$). The results are further discussed in the next section.



	Explicit (k=29)	EX CT (k=7)	EX SPT (k11)	EX FPT (k=11)	Implicit (k=18)	IM CT (k=4)	IM SPT (k=7)	IM FPT (k=7)
d	1.96	0.74	2.14	2.56	1.01	1.05	0.78	1.22
UCI	2.37	1.19	2.62	3.23	1.31	1.78	1.14	1.77
LCI	1.55	0.29	1.66	1.89	0.71	0.32	0.42	0.67

Fig. 2 Average effect sizes and confidence intervals of explicit and input-based tasks across outcome measures. MPD: meta-pragmatic discussion; MPD × C: meta-pragmatic discussion effect on comprehension (C); MPD × SP: meta-pragmatic discussion effect on structured production (SP); MPD × FP: meta-pragmatic discussion effect on free production (FP); IBT: input-based tasks; IBT × C: input-based tasks effect on comprehension (C); IBT × SP: input-based tasks effect on structured production (SP); IBT × FP: input-based tasks effect on free production (FP); *k*: number of unique sample studies; *d*: average effect size; UCI: upper confidence interval; LCI: lower confidence interval

6 Discussion

The results presented in the previous section are discussed with reference to the two research questions dealt within this study: (1) *Do different instructional tasks variedly affect the effectiveness L2 pragmatics instruction*? (2) *Are there any differences in the effects of instructional tasks on L2 pragmatics across comprehension and production outcome measures*? The first research question focused on the factors contributing to the effectiveness of L2 pragmatics learning tasks and was addressed through the subgroup analysis centered on the features characterizing the instructional designs of each of the primary studies (Appendix 3).

Overall, the findings show that meta-pragmatic discussion, input-based communicative activities, and the type of feedback provided affect L2 pragmatics development from multiple but related perspectives. In addition, it is worth noting the findings with regard to the importance of meta-pragmatic discussion are in line with earlier studies on the factors that are likely to facilitate SL and L2 pragmatics development (Ellis 2005; Bardovi-Harlig and Griffin 2005; Hassall 2008; Kondo 2008). The findings also provide evidence for the worth of recast feedback (Lyster 2004; Russell and Spada 2006).

To start with, the finding that meta-pragmatic discussion was more advantageous for structured production and free production is consistent with several studies in SLA research (e.g., Norris and Ortega 2000; Ranta and Lyster 2007), and with the claims of skill acquisition theory (DeKeyser 1997, 1998, 2001, 2003, 2007a). On one side, meta-pragmatic discussion's effect as measured through free production overtakes the strongest effect size. This finding highlights the pivotal role the process of proceduralization plays in reinforcing L2 pragmatics development. Proceduralization explains the interplay between declarative and procedural knowledge representations. Lyster and Sato (2013, p. 72) define the two types of knowledge representation stating: "Declarative knowledge is static information such as historical or geographical facts encoded in memory. Procedural knowledge entails knowing how to do things including the ability to apply rule-based knowledge to cognitive as well as motor operations." In other words, while declarative knowledge is constructed through the internalization of explicit information about SL pragmatic features and their functions, procedural knowledge refers to the ability to draw on explicit knowledge representation when using SL pragmatic norms in practice. Proceduralization is a concept introduced in Anderson's adaptive control of thought (ACT) model (Anderson 1993, 1996; Anderson and Fincham 1994), which claims that knowledge is first acquired in an explicit form, and due to subsequent practice, it transforms to practical procedures employed when executing relevant cognitive operations, and then, the explicit rules become no longer accessible. Besides receiving explicit declarative information, the majority of participants in meta-pragmatic discussion treatments engaged in explicit consciousness-raising tasks which are likely to trigger and stimulate the process of proceduralization.

Similarly, meta-pragmatic discussion's effect size is stronger on free production than on structured production. A reasonable clarification arises from the effect of proceduralization through practice (DeKeyser 2007b) coupled with the assumption that implicit knowledge exists as procedural facts that can be easily and rapidly accessed in unplanned language use, whereas explicit knowledge exists as declarative facts that can only be accessed through the application of attentional processes (Ellis et al. 2009). Thus, on the one hand, in structured production tasks, participants rely on retrieving the SL forms and their norms of use in accordance with the interaction situations to be completed in open-ended written discourse completion tests. However, when performing free production tests, participants capitalize on their procedural knowledge due to the effect of practice and the proceduralization of explicit rules. According to Ellis (2003): Production, then, may constitute the mechanism that connects the learner's dual systems, enabling movement to occur from the memory-based to the rule-based system and vice versa. If this interpretation is correct, learners may not be so reliant on input as has been generally assumed in SLA. They may be able to utilize their own internal resources, via using them in production, to both construct and complexify their interlanguage (p. 115). On the other hand, the less strong effect sizes achieved through input-based treatments on both of structured and free production tasks could be explained by the very nature of production just mentioned in the quote above and by the absence of the initial explicit phase when SL pragmatic norms are first internalized, discussed, and proceduralized. This argument is supported by the claim that comprehension draws on learners' ability to transform input into intake, while production requires deep processing (VanPatten 2002a, b; Leow 2007).

Surprisingly, input-based tasks' effect size is more significant than meta-pragmatic discussion's effects on comprehension. One explanation to participants in treatments that involved input-based tasks outperforming the groups that received meta-pragmatic discussion can be derived from the assumptions of skill acquisition theory (SAT) due to its emphasis on practice and feedback (DeKeyser 2007a; Leow 2007; Leeman 2007). An assertion central to SAT is that "in order to promote continued L2 growth in classroom settings, opportunities for contextualized practice are needed to complement input-driven approaches designed to trigger noticing and awareness of SL features" (Lyster and Sato 2013, p. 71). In other words, participants in input-based treatments are presented with tasks which involve input enhancement, comprehension-oriented questions, and structured input activities, and hence had ample opportunities for practice with regard to comprehension. At this point, it can be suggested that the missing string for meta-pragmatic discussion groups' less efficient performance on comprehension tests resides in the lack of practice. In parallelism, the same treatment groups' effect sizes are larger on production tests which were less practiced in input-based treatments. However, although input-based tasks were more successful in promoting L2 pragmatics comprehension, they seem to be less effective in enhancing learners' performance on production which, at least initially, requires clear representation of explicit knowledge so that what to say, to whom, and how could be competently retrieved.

Furthermore, feedback is likely to strengthen the effects of meta-pragmatic discussion (Table 2), especially in relation to the nature of feedback provided. This last assumption could be valid since all the studies reported using corrective feedback. Unlike recasts, feedback as error correction seems to fall short of enabling learners work out solutions for the flaws in their pragmatic performance. Besides, when presenting corrections, learners' attention is not directed to noticing the gaps in their production but rather to solving a communication "breakdown" while carrying out an interactional task (Schmidt 1994; Robinson 1995), and therefore, its presence moderately affected the effectiveness of meta-pragmatic discussion.

Interestingly, however, feedback makes a clear difference with input-based tasks (IBT + F, $d = 1.14$, SD = 0.84; IBT − F, $d = 0.84$, SD = 0.52). All the input-based task studies which included feedback provision reported combining their implicit treatments with recast feedback, except for Hernández (2011) where participants received correction for their responses. A long-standing assumption in SLA research is that recasts are usually designed and implemented in a way that they help learners notice gaps in their pragmatics production reinforcing, as a result, the transformation of input into intake ready for retrieval in upcoming performance tasks (Philp 2003; Iwashita 2003; Leeman 2003). Obviously, this is not to say that corrective feedback (mere provision of the correct form) is useless but to suggest that L2 learners notably benefit from opportunities to modify or restructure their output/production provided through interactional feedback (Nicholas et al. 2001; Mackey 2006; Mackey and Goo 2007; Li 2010). Although instruction in both conditions of input-based tasks with and without feedback is beneficial, the general finding is that input-based tasks are most effective when accompanied by recast feedback. This finding reinforces the notion of knowledge representation discussed earlier. The interpretation suggested for the usefulness of meta-pragmatic discussion over input-based tasks probably lies in the lucid explicit knowledge constructed due to providing explicit information and discussion regarding SL pragmatic features and their norms of use. Alternatively, recast feedback is likely to boost the effectiveness of input-based tasks through helping develop sophisticated implicit knowledge representations and eventually assimilate features of input into intake.

7 Implications

The main findings of this study can also be discussed from the perspective of theoretical and practical relevance for second language learning and instruction. Findings from the subgroup analyses point toward the differential effects of L2 pragmatics instruction in terms of instructional features as well as comprehension and production aspects of SL development.

The findings point at differential effects of explicit and input-based tasks on L2 pragmatics comprehension and production. The evidence derived from this

synthesis study further places the variability in L2 pragmatics performance across comprehension and production data-eliciting measures within the cognitive framework of skill acquisition theory. One practical contribution of the present meta-analysis is to inform interventional studies focused on investigating the enhancement of both comprehension and production aspects of pragmatic competence in instructed SL learning contexts (Bardovi-Harlig and Griffin 2005; Rose 2005; Ellis 2005).

Moreover, the present synthetic research brings instructed L2 pragmatics closer to SLA-rooted research issues, especially from the perspective of comprehension and production theoretical models, skill acquisition theory in this case. Practically, this meta-analysis also carries important implications for SL teachers. The results from the combined groups' effect sizes across comparative treatment conditions emphasize the importance of integrated instruction in L2 pragmatics comprehension and production. That is, future models of instructed L2 pragmatics would benefit from taking into consideration the shortcut to faster development through integrating meta-pragmatic discussion, input-based practice, and both types of feedback, correction and recast.

Lastly, it is noteworthy that meta-pragmatic discussion's effect sizes display wider confidence interval than the input-based tasks. This finding indicates that participants receiving input-based tasks seem to be approximately equally alert to input features, whereas participants in explicit conditions show dissimilar rates of understanding and applying the rules that have been overtly stated during instruction. A relevant assumption is that individual differences such as memory capacity, i.e., the ability retain and retrieve explicit information, learning style, analytic versus global for instance, seem to moderate the usefulness of instructed L2 pragmatics. These findings provide evidence for the interaction between learning conditions and individual differences. Therefore, further research may be needed to augment the consistency of results for the process and role of proceduralization in L2 pragmatics development.

Taken together, the findings reveal that both types of ILP intervention resulted in strong effects of instruction. The advantage of both types of instruction is probably not surprising since it is widely acknowledged that though all languages involve performing requests and apologies for instance, learners may not be aware of the corresponding forms and functions in the SL and/or the context and norms regulating their proper use. Besides, classroom discourse provides very limited exemplars of speech acts and lesser speech act sets leading to shallow L2 pragmatics awareness. Similarly, Kasper and Roever (2005) suggest that instruction aimed at promoting L2 learners' pragmatic competence plays the role of input generator through encouraging exposure to authentic language use, along with providing ample opportunities for learning through noticing, practice, and production.

8 Conclusions, Limitations, and Suggestions for Further Research

From a meta-analytic perspective, this article has attempted to provide a synthesis of the ways in which instruction could be tailored to promote L2 pragmatics development. As the research questions imply, the contribution of this study revolves around the role of instructional tasks' design in developing SL pragmatic comprehension and production as measured through three types of outcome measures.

Taking a cognitive theory of SLA standpoint, this meta-analysis has aimed to arrive at a deepened understanding regarding the effectiveness of two task-based approaches to instructed L2 pragmatics, namely meta-pragmatic discussion and input-based tasks, at the levels of comprehension and production under constrained and non-constrained task demands. By and large, the standardized effect sizes of both types of instructional tasks indicate huge effects of instruction in both conditions (meta-pragmatic discussion, $d = 1.96$; input-based tasks, $d = 1.01$). The meta-analysis yields significant insights into the role of knowledge representation and processing instruction in the development of L2 pragmatics. Besides, the discussion of the meta-analytic results emphasizes the roles of meta-pragmatic discussion and input-based tasks in promoting L2 pragmatics comprehension and production and sheds light on the effect of feedback.

The same as any research work, this meta-analysis has a number of limitations. One limitation is that most of the synthesized studies have been conducted in the foreign language context; the foci were on the inside classroom instructional techniques and measurement procedures. With the increasing interest of the role that context plays in the development of acquisitional pragmatics in second language contexts, further research relevant to this issue is needed, especially with regard to second language learning in the study abroad context. A further challenge to any meta-analytic work is that "It is not possible, in case of multiple component treatments, to evaluate the contribution of each component to the outcome" (Kasper 2009, p. 269). However, though it is true that estimating how effective a technique within an amalgam of instructional procedures is impossible, the present meta-analysis draws on the premise that some treatment features are necessary if pragmatic instruction is to be effective. Finally, as indicated by the large confidence intervals (Fig. 1), meta-pragmatic discussion seems to be more effective with some learners and less useful to others. This raises questions as to whether proficiency, psychological factors, and exposure to input mediate the effectiveness of ILP instruction and whether an input-based warm up prior to meta-pragmatic discussion would result in more robust results regarding the usefulness of instructed L2 pragmatics in foreign language contexts.

Summing up, by examining the effects of learning tasks, this study was aimed at providing insights into the development of L2 pragmatics comprehension and production in instructed contexts. Through the present synthesis, it is hoped that the findings would provide important insights into how instruction could be designed to promote L2 pragmatics development and inspire future research endeavors.

Appendix 1: Learner Characteristics

Study	FL	TL	Age	Mean	Educational level	Proficiency level	Length of study	Mean
Barekat and Mehri (2013)	Persian	English	16–20	18	/	Intermediate	/	
Fordyce (2013)	Japanese	English	18–22	20	University	Intermediate	/	
Farahian et al. (2012)	Persian	English	19–25	22	University	Intermediate	/	
Bu (2012)	Chinese	English	20–21	20.5	University	Intermediate	/	
Jernigan (2012)	Heterogeneous	English	/		/	Intermediate	/	
Li (2012)	Heterogeneous	Chinese	/		/	Intermediate	/	
Narita (2012)	Japanese	English	18–24	21	University	Intermediate	3 years	3
Nguyena et al. (2012)	Vietnamese	English	/		/	Intermediate	6–9 years	7.5
Takimoto (2012)	Japanese	English	20	20	University	intermediate	/	
Chen (2011)	Chinese	English	/		University	intermediate	/	
Dastjerdi and Farshid (2011)	Persian	English	23	23	University	Intermediate	Over 4 years	4
Gu (2011)	Chinese	English	22–27	23.5	University	Advanced	10–15 years	12.5
Hernández (2011)	English	Spanish	19	19	University	Advanced	/	
Takimoto (2011)	Japanese	English	20	20	University	Intermediate	8 years	8
Nipaspong and Chinokul (2010)	Thai	English	17–19	18	University	Intermediate	10 years	10
Alcón-Soler and Guzmán-Pitarch (2010)	Spanish	English	18–30	24	University	Intermediate	/	
Ghobadi and Fahim (2009)	Persian	English	14–18	16	/	Intermediate	/	
Takimoto (2009)	Japanese	English	18–40	29	/	Intermediate	5–22	13.5
Eslami and Eslami-Rasekh (2008)	Persian	English	23–30	26.5	University	Advanced	9 years	9
Martínez-Flor and Alcón-Soler (2007)	Spanish	English	19–20	19.5	University	Intermediate	6–10	8
Takimoto (2006)	Japanese	English	18–40	29	/	Intermediate	5–25	15
Koike and Pearson (2005)	English	Spanish	/		University	Intermediate	1.5	1.5
Alcón-Soler (2005)	Spanish	English	17–18	17.5	Secondary school	Intermediate	7–10	8.5
Fukuya and Martínez-Flor (2005)	Spanish	English	19–25	22	University	Intermediate	Over 5 years	5

"/" refers to missing data

Appendix 2: Sample Size and Study Design Across Studies

Study	Total sample size	Experimental group(s)	Control group	Pretest	Posttest	Delayed posttest
Barekat and Mehri (2013)	45	2(15)	15	+	+	+
Fordyce (2013)	81	2(44 + 37)	/	+	+	+
Farahian et al. (2012)	64	1(32)	32	+	+	−
Bu (2012)	90	2(30)	30	+	+	−
Jernigan (2012)	34	1(34)	/	+	+	−
Li (2012)	30	1(20)	10	+	+	+
Narita (2012)	41	1(22)	19	+	+	+
Nguyena et al. (2012)	69	2(19 + 28)	22	+	+	+
Takimoto (2012)	45	2(15)	15	+	+	+
Chen (2011)	40	1(40)	/	+	+	+
Dastjerdi and Farshid (2011)	38	2(38)	/	+	+	−
Gu (2011)	40	2(20)	/	+	+	−
Hernández (2011)	91	2(36 + 30)	25	+	+	+
Takimoto (2011)	59	2(41)	18	+	+	+
Nipaspong and Chinokul (2010)	39	2(13)	13	+	+	−
Alcón-Soler and Guzmán-Pitarch (2010)	92	1(92)	/	+	+	+
Ghobadi and Fahim (2009)	60	2(30)	/	+	+	−
Takimoto (2009)	60	3(15)	15	+	+	+
Eslami and Eslami-Rasekh (2008)	52	1(25)	27	+	+	−
Martinez-Flor and Alcón-Soler (2007)	81	2(25 + 24)	32	+	+	−
Takimoto (2006)	45	30	15	+	+	+
Koike and Pearson (2005)	99	2(38 + 29)	32	+	+	−
Alcón-Soler (2005)	132	2(44)	44	+	+	−
Fukuya and Martínez-Flor (2005)	81	2(25 + 24)	32	+	+	−

Appendix 3: Instructional Design

Study	Target features	Instructional treatment	Length of treatment (h)	Outcome measure
Barekat and Mehri (2013)	Requests	Consciousness-raising CR/CR + feedback versus no instruction	4	WDCT
Fordyce (2013)	Epistemic stance	Explicit versus implicit	Approximately 3	W production
Farahian et al. (2012)	Refusals	Explicit versus no instruction	3	WDCT
Bu (2012)	Suggestions	Explicit/implicit/no instruction	18	DCT
Jernigan (2012)	Speech acts	Implicit	/	PAJT, DCT
Li (2012)	Requests	Implicit	1	PAJT + DCT
Narita (2012)	Hear-say evidential markers	Implicit	2	MCT + PAJT + Oral Discourse Production
Nguyena et al. (2012)	Constructive criticism	Explicit/implicit/no instruction	7	WDCT, Oral Role Play and Peer Feedback
Takimoto (2012)	Request downgraders	Explicit/implicit/no instruction	Approximately 2	PAJT, WDCT
Chen (2011)	Compliments	Explicit	8	WDCT
Dastjerdi and Farshid (2011)	Compliments	Explicit versus implicit	3	WDCT
Gu (2011)	Requests	Explicit versus implicit	2	WDCT + RP
Hernández (2011)	Discourse markers	Explicit/implicit/no instruction	Approximately 2	ODPT
Takimoto (2011)	Request downgraders	Implicit versus no instruction	Approximately 2	PAJT + WDCT
Nipaspong and Chinokul (2010)	Pragmatic awareness	Explicit/implicit/no instruction	15	MCT
Alcón-Soler and Guzmán-Pitarch (2010)	Refusals to requests	Explicit/no instruction	12	ODP
Ghobadi and Fahim (2009)	Thanking	Explicit versus implicit	9	WDCT + RP
Takimoto (2009)	Request downgraders	Explicit/implicit/no instruction	Approximately 3	PAJT+ WDCT +RP
Eslami and Eslami-Rasekh (2008)	Requesting and apologizing	Explicit/no instruction	7	PAJT + WDCT
Martínez-Flor and Alcón-Soler (2007)	Suggestions	Explicit/implicit/no instruction	12	PAJT

(continued)

(continued)

Study	Target features	Instructional treatment	Length of treatment (h)	Outcome measure
Takimoto (2006)	Requests	Explicit/implicit/no instruction	Approximately 3	WDCT + RP
Koike and Pearson (2005)	Suggestions	Explicit/implicit/no instruction	1	MCT + WDCT
Alcón-Soler (2005)	Requests	Explicit versus implicit	12	Comprehension and production tasks
Fukuya and Martínez-Flor (2005)	Suggestions	Explicit/implicit/no instruction	12	Production tasks

Appendix 4: Effect Size Values for Explicit and Implicit Unique Sample Studies

I. Explicit

Study	Population	Effect size *d*
Structured production tasks (SPT)		
Farahian et al. (2012)	64	10.335
Bu (2012)	90	1.919
Li (2012)	30	1.615
Nguyena et al. (2012)	69	2.166
Chen (2011)	40	2.203
Ghobadi and Fahim (2009)	60	2.770
Eslami and Eslami-Rasekh (2008)	52	1.667
Koike and Pearson (2005)	99	0.767
Barekat and Mehri (2013)	45	3.587
Takimoto (2012)	45	1.314
Takimoto (2009)	60	3.031
Takimoto (2006)	45	2.540
Average d = 2.14	LCI = 1.66	UCI = 2.62
Comprehension tasks (CT)		
Li (2012)	30	−0.007
Narita (2012)	41	1.486
Nipaspong and Chinokul (2010)	39	0.428
Eslami and Eslami-Rasekh (2008)	52	0.003
Koike and Pearson (2005)	99	0.78
Alcón-Soler (2005)	132	6.018

(continued)

(continued)

Study	Population	Effect size *d*
Takimoto (2012)	45	1.218
Takimoto (2009)	60	1.265
Average *d* = 0.74	LCI = 0.29	UCI = 1.19
Production tasks (PT)		
Fordyce (2013)	81	1.488
Alcón-Soler (2005)	132	4.566
Fukuya and Martínez-Flor (2005)	81	3.076
Narita (2012)	41	4.495
Nguyena et al. (2012)	69	2.366
Alcón-Soler and Guzmán-Pitarch (2010)	91	2.544
Ghobadi and Fahim (2009)	60	1.420
Fukuya and Martínez-Flor (2005)	81	2.234
Takimoto (2009)	60	2.778
Takimoto (2006)	45/3	1.889
Hermandez 2011	91	1.260
Average *d* = 2.56	LCI = 1.89	UCI = 2..56
Overall effect of explicit treatments, *d* = 1.96	LCI = 1.55	UCI = 2.37

II. **Implicit**

Study	Population	Effect size d
Structured production tasks (SPT)		
Bu 2012	90	1.157
Jernigan (2012)	34	0.281
Nguyena et al. (2012)	69	0.295
Ghobadi and Fahim (2009)	60	1.410
Nguyena et al. (2012)	69	0.296
Koike and Pearson (2005)	99	0.817
Takimoto (2011)	59	1.170
Average *d* = 0.78	LCI = 0.42	UCI = 1.14
Comprehension tasks (CT)		
Jernigan (2012)	34	0.434
Nipaspong and Chinokul (2010)	39	1.839
Koike and Pearson (2005)	99	0.387
Alcón-Soler (2005)	132	7.279
Takimoto (2011)	59	1.533
Average *d* = 1.05	LCI = 0.32	UCI = 1.78

(continued)

(continued)

Study	Population	Effect size d
Production tasks (PT)		
Fordyce (2013)	81	0.295
Alcón-Soler (2005)	132	1.386
Fukuya and Martínez-Flor (2005)	81	2.391
Nguyena et al. (2012)	69	1.673
Ghobadi and Fahim (2009)	60	0.482
Fukuya and Martínez-Flor (2005)	81	1.483
Hernández (2011)	91	0.81
Average d = 1.22	LCI = 0.67	UCI = 1.77
Overall effect of implicit treatments, d = 1.01	LCI = 0.71	UCI = 1.31

References

*Alcón-Soler, E. (2005). Does instruction work for learning pragmatics in the EFL context? *System, 33*(3), 417–435.

Alcón-Soler, E. (2008). Investigating pragmatic language learning in foreign language classrooms. *IRAL-International Review of Applied Linguistics in Language Teaching, 46*(3), 173–195.

*Alcón-Soler, E., & Guzmán-Pitarch, J. R. (2010). The effect of instruction on learners' pragmatic awareness: A focus on refusals. *International Journal of English Studies*, *10*(1), 65–80.

Alcón-Soler, E., & Guzmán-Pitarch, J. R. (2013). The effect of instruction on learners' use and negotiation of refusals. *Utrecht Studies in Language & Communication, 25*, 41–63.

Anderson, J. R. (1993). Problem solving and learning. *American Psychologist*, *48*(1), 35.

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist, 51*(4), 355.

Anderson, J. R., & Fincham, J. M. (1994). Acquisition of procedural skills from examples. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 20*(6), 1322.

Bardovi-Harlig, K. (1999). The interlanguage of interlanguage pragmatics: A research agenda for acquisitional pragmatics. *Language Learning, 49*, 677–713.

Bardovi-Harlig, K. (2001). Pragmatics and second language acquisition. In R. Kaplan (Ed.), *The handbook of applied linguistics* (pp. 182–192). Oxford, UK: Oxford University Press.

Bardovi-Harlig, K. (2013). Developing L2 pragmatics. *Language Learning, 63*(1), 68–86.

Bardovi-Harlig, K., & Griffin, R. (2005). L2 pragmatic awareness: Evidence from the ESL classroom. *System, 33*(3), 401–415.

*Barekat, B., &Mehri, M. (2013). Investigating the effect of meta-linguistic feedback in L2 pragmatic instruction. *International Journal of Linguistics*, *5*(2), 197–208.

Billmyer, K., & Varghese, M. (2000). Investigating instrument-based pragmatic variability: Effects of enhancing discourse completion tests. *Applied Linguistics, 21*(4), 517–552.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. Wiley.com.

*Bu, J. (2012). A Study of the effects of explicit and implicit teachings on developing Chinese EFL learners' pragmatic competence. *International Journal of Language Studies*, *6*(3), 57–80.

Bygate, M., Shehan, P., & Swain, M. (2001). *Researching pedagogic tasks: Second language learning teaching and testing*. UK: Pearson Education Limited.

Byon, A. S. (2006). Developing KFL students' pragmatic awareness of Korean speech acts: The use of discourse completion tasks. *Language Awareness, 15*(4), 244–263.

*Chen, Y. S. (2011). The effect of explicit teaching of American compliment exchanges to Chinese learners of english. *English Teaching & Learning*, *35*(4), 1–42.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Psychology Press.

Cooper, H. (2009). *Research synthesis and meta-analysis: A step-by-step approach* (Vol. 2). Sage.

Crookes, G., & Long, M. (1992). Three approaches to task-based design. *TESOL Quarterly, 26*, 27–56.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Rutledge.

Dastjerdi, H. V., & Farshid, M. (2011). The role of input enhancement in teaching compliments. *Journal of Language Teaching & Research*, *2*(2), 460–466.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morpho syntax. *Studies in Second Language Acquisition, 19*, 195–222.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). Cambridge, UK: Cambridge University Press.

DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). Cambridge, UK: Cambridge University Press.

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 313–348). Oxford: Blackwell.

DeKeyser, R. (2007a). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 97–112). Mahwah, NJ: Erlbaum.

DeKeyser, R. M. (Ed.). (2007b). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology.* Cambridge: Cambridge University Press.

Dodge, M., & Stinson, C. (2010). Microsoft® Excel® 2010 inside out. O'Reilly Media, Inc.

Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research, 4*(3), 193–220.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R. (2005). Principles of instructed language learning. *System, 33*, 209–224.

Ellis, R., Loewen, S., Elder, C., Erlam, R., Philip, J., & Reinders, H. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching* (p. 19). Bristol, England: Multilingual Matters.

Eslami, Z. R., & Eslami-Rasekh, A. (2008). Enhancing the pragmatic competence of non-native english-speaking teacher candidates (NNESTCs) in an EFL context. In E. Alcón-Soler & A. Martinez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 178–200). UK: Cromwell Press Ltd.

*Farahian, M., Rezaee, M., & Gholami, A. (2012). Does direct instruction develop pragmatic competence? Teaching refusals to EFL learners of english. *Journal of Language Teaching &Research*, *3*(4), 814–821.

*Fordyce, K. (2013). The Differential effects of explicit and implicit instruction on EFL learners' use of epistemic stance. *Applied Linguistics*.

*Fukuya, Y. J., & Martínez-Flor, A. (2005). The effects of instruction on learner's production of appropriate and accurate suggestions. *System*, *33*(3), 463–480.

Fukuya, Y. J., & Martínez-Flor, A. (2008). The interactive effects of pragmatic-eliciting tasks and pragmatic instruction. *Foreign Language Annals, 41*(3), 478–500.

*Ghobadi, A., & Fahim, M. (2009). The effect of explicit teaching of english "than ing formulas" on Iranian EFL intermediate level students at english language institutes. *System*, *37*(3), 526–537.

*Gu, X. L. (2011). The effect of explicit and implicit instructions of request strategies. *Intercultural Communication Studies*, *20*(1), 104–123.

Hassall, T. (2008). Pragmatic performance: What are learners thinking? In E. Alcon-Soler & A. Martinez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 74–93). UK: Cromwell Press Ltd.

*Hernández, T. A. (2011). Re-examining the role of explicit instruction and input flood on the acquisition of Spanish discourse markers. *Language Teaching Research*, *15*(2), 159–182.

Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition, 25*(1), 1–36.

Jeon, E. H., & Kaya, T. (2006). Effects of L2 instruction on interlanguage pragmatic development. In M. Norris., & L. Ortega (Eds.), *Synthesizing research on language learning and teaching,* (pp. 165–211). Philadelphia/Amsterdam: John Benjamins.

*Jernigan, J. (2012). Output and english as a second language pragmatic development: The effectiveness of output-focused video-based instruction. *English Language Teaching*, *5*(4), 2–14.

Kasper, G. (2001a). Four perspectives on L2 pragmatic development. *Applied Linguistics, 22*(4), 502–530.

Kasper, G. (2001b). Classroom research on interlanguage pragmatics. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 33–60). Cambridge: Cambridge University Press.

Kasper, G. (2009). L2 pragmatic development. In New *handbook of second language acquisition* (pp. 259–298). Leeds, UK: Emerald,.

Kasper, G., & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in second language acquisition, 13*(2), 215–247.

Kasper, G., & Roever, C. (2005). Pragmatics in second language learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 317–334). Mahwah, NJ: Lawrence Erlbaum Associates.

Kasper, G., & Rose, K. R. (2001). Pragmatics in language teaching. In G. Kasper & K. R. Rose (Eds.), *Pragmatics in language teaching* (pp. 1–10). Cambridge: Cambridge University Press.

Kasper, G., & Rose, K. R. (2002). Pragmatic development in a second language. *Language Learning, 52*, 1.

Kasper, G., & Schmidt, R. (1996). Developmental issues in interlanguage pragmatics. *Studies inSecond Language Acquisition, 18*(2), 149–169.

*Koike, D. A., & Pearson, L. (2005). The effect of instruction and feedback in the development of pragmatic competence. *System*, *33*(3), 481–501.

Kondo, S. (2008). Effects on pragmatic development through awareness-raising instruction: Refusals by Japanese EFL learners. In E. Alcon-Soler & A. Martinez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 153–177). UK: Cromwell Press Ltd.

Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies inSecond Language Acquisition, 25*, 37–63.

Leeman, J. (2007). Feedback in L2 learning: Responding to errors during practice. In R. M. Dekeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 111–183). UK: Cambridge University Press.

Leow, R. P. (2007). Input in the L2 classroom: An attentional perspective on receptive practice. In R. M. Dekeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 21–50). UK: Cambridge University Press.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning, 60*, 309–365.

*Li, S. (2012). The effects of input-based practice on pragmatic development of requests in L2 Chinese. *Language Learning*, *62*(2), 403–438.

Lipsey, M. W., & Wilson, D. (2000). Practical meta-analysis (applied social research methods).

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition, 26*, 399–432.

Lyster, R., & Sato, M. (2013). Skill acquisition theory and the role of practice in L2 development. In M. D. P. G. Mayo, M. J. G. Mangado & M. M. Adrián (Eds.), *Contemporary approaches to second language acquisition* (pp. 71–92). Amsterdam: John Benjamins Publishing Company.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics, 27*, 1–27.

Mackey, A., & Goo, J. M. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Input, interaction and corrective feedback in L2 learning* (pp. 379–452). Oxford, UK: Oxford University Press.

Martínez-Flor, A. (2013). Learners' production of refusals: Interactive written DCT versus oral role play. *Utrecht Studies in Language & Communication*, *25*.

Martínez-Flor, A., & Fukuya, Y. J. (2005). The effects of instruction on learners' production of appropriate and accurate suggestions. *System, 33*(3), 463–480.

*Martínez-Flor, A., & Alcón-Soler, E. (2007). Developing pragmatic awareness of suggestions in the EFL classroom: A focus on instructional effects. *Canadian Journal of Applied Linguistics*, *10*(1), 47–76.

*Narita, R. (2012). The effects of pragmatic consciousness-raising activity on the development of pragmatic awareness and use of hearsay evidential markers for learners of Japanese as a foreign language. *Journal of Pragmatics*, *44*(1), 1–29.

*Nguyena, T. T. M., Phamb, T. H., & Phamb, M. T. (2012). The relative effects of explicit and implicit form-focused instruction on the development of L2 pragmatic competence. *Journal of Pragmatics*, *44*, 416–434.

Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning, 51*, 719–758.

*Nipaspong, P., & Chinokul, S. (2010). The role of prompts and explicit feedback in raising EFL learners' pragmatic awareness. *University of Sydney Papers in TESOL*, *5*(5), 101–146.

Norris, J. M. (2012). Meta-analysis. In C. Chapelle (Ed.), *Encyclopedia of applied linguistics*. Malden, MA: Wiley.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417–528.

Norris, M., & Ortega, L. (2006). *Synthesizing research on language learning and teaching*. Philadelphia/ Amsterdam: John Benjamins.

Norris, J. M., & Ortega, L. (2007). The future of research synthesis in applied linguistics: Beyond art or science. *TESOL Quarterly, 41*, 805–815.

Norris, J. M., & Ortega, L. (2010). Timeline: Research synthesis. *Language Teaching, 43*, 461–479.

Nunan, D. (2004). *Task-based language teaching*. UK: Cambridge University Press.

Ortega, L. (2010). Research synthesis. In B. Paltridge & A. Phakiti (Eds.), *Companion to research methods in applied linguistics* (pp. 111–126). London: Continuum.

Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. Annual challenges. *Annual Review of Applied Linguistics, 30*, 85–110.

Philp, J. (2003). Constraints on "noticing the gap": Non-native speakers' noticing of recasts in NSNNS interaction. *Studies in Second Language Acquisition, 25*, 99–126.

Ranta, L. & Lyster, R. (2007). A cognitive approach to improving immersion students' oral language abilities: The Awareness-Practice-Feedback sequence. In R. M. Dekeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 141–160). Cambridge: Cambridge University Press.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*(2), 283–331.

Rose, K. R. (1994). Pragmatic consciousness-raising in an EFL context. In L. F. Bouton & Y. Kachru (Eds.), *Pragmatics and language learning monograph series* (Vol. 5, pp. 52–63). Urbana, IL: University of Illinois at Urbana-Champaign.

Rose, K. R. (1997). Pragmatics in the classroom: Theoretical concerns and practical possibilities. In L. F. Bouton (Ed.), *Pragmatics and language learning* (Vol. 8). Urbana, IL: University of Illinois at Urbana-Champaign.

Rose, K. R. (2005). On the effects of instruction in second language pragmatics. *System, 33*(3), 385–399.

Rose, K., & Kasper, G. (Eds.). (2001). *Pragmatics in language teaching*. Cambridge: Cambridge University Press.

Rosenthal, M. C. (1994). The fugitive literature. In *The handbook of research synthesis* (pp. 85–94).

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). (2006). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. NY: Wiley.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). Philadelphia/Amsterdam: John Benjamins.

Safont-Jordà, M. P., & Portolés-Falomir, L. (2013). Research method effects on third language learners' refusals. *Utrecht Studies in Language & Communication*, *25*.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11–26.

Shehadeh, A. (2005). Task-based language learning and teaching: Theories and applications. In C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in english language teaching* (pp. 13–30). UK: Palgrave Macmillan.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*, 38–62.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Smith, M. L. (1980). Publication bias and meta-analysis. *Evaluation in Education, 4*, 22–24.

Swain, M. & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Shehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning teaching and testing*. UK: Pearson Education Limited.

Takahashi, S. (2005). Noticing in task performance and learning outcomes: A qualitative analysis of instructional effects in interlanguage pragmatics. *System, 33*(3), 437–461.

*Takimoto, M. (2006a). The effects of explicit feedback and form–meaning processing on the development of pragmatic proficiency in consciousness-raising tasks. *System*, *34*(4), 601–614.

Takimoto, M. (2006b). The effects of explicit feedback on the development of pragmatic proficiency. *Language Teaching Research, 10(4),* 393-417.

Takimoto, M. (2008a). The effects of deductive and inductive instruction on the development of language learners' pragmatic competence. *The Modern Language Journal*, *92*(3), 369–386.

Takimoto, M. (2008b). The effects of various kinds of form-focused instruction on learners' ability to comprehend and produce polite request in english. *TESL Canada Journal*, *26*(1), 31–51.

*Takimoto, M. (2009). The effects of input-based tasks on the development of learners' pragmatic proficiency. *Applied Linguistics*, *30*(1), 1–25.

*Takimoto, M. (2011). Measuring the effects of structured input task repetition on learners' interlanguage pragmatic proficiency. *New Zealand Studies in Applied Linguistics*, *17*(2), 21–36.

*Takimoto, M. (2012). Assessing the effects of identical task repetition and task-type repetition on learners' recognition and production of second language request down-graders. *Intercultural Pragmatics*, *9*(1), 71–96.

Usó-Juan, E. (2013). Effects of metapragmatic instruction on EFL learners' production ofrefusals1. In *Refusals in instructional contexts and beyond* (Vol. 25, pp. 65–99).

VanPatten, B. (2002a). Processing instruction: An update. *Language Learning, 52*(4), 755–803.

VanPatten, B. (2002b). Processing the content of input-processing and processing instruction research: A response to DeKeyser, Salaberry, Robinson, and Harrington. *Language Learning, 52*(4), 825–831.

Vellenga, H. E. (2011). Teaching L2 pragmatics: Opportunities for continuing professional development. *Tesl-Ej*, *15*(2).

Large Class Size and Student–Lecturer Learning Experiences at the Tertiary Level

Ramachandran Ponnan, Marzura Abdul Malek and Balaguru Ambalavanan

Abstract Maintaining harmonious classroom environment is an important aspect of student–lecturer relationship. Harmonious learning environment also greatly depends on the size of the classroom and student–lecturer interactions. Large students' enrollment into tertiary institutions can be attributed to one of the several causes. Of late budgetary concerns have quickly led to the establishment of large-sized classes, affecting students' performance and hence their classroom relationship with lecturers. Literature reviewed suggests that there were several arguments against large-sized classes that form the basis for this research, namely (1) over-reliance on lecture; (2) passive student engagement; (3) reduced lecturer interaction/feedback with students; (4) reduced depth of students' thinking; (5) reduced depth of course objectives and learning strategies used by students outside the classroom; (6) lower-level academic achievement and performance; (7) course dissatisfaction with the learning experience; and (8) lower student rating of course instructions. This is an empirical study by three lecturers involving over three hundred students in a university setting. The study was done by conducting interviews, obtaining students' perceptions and classroom observation over two semesters in the same university setting. The results and findings are presented in the form of discussions and implications for action. The findings in this paper suggest that lecturers' affective behavior such as personal bias had an impact on the classroom relationship. Also discussed are suggestions on how teachers should avoid displaying specific behavior as personal emotion. Student–lecturer relationship along a professional attitude could be an important factor in determining the successful teaching and learning engagement in large–small classrooms. The study would benefit lecturers, education administrators in developing new approaches, assistances, and facilities to teaching and learning.

Keywords Student–lecturer relationship · Large classroom instructions

R. Ponnan (✉) · M.A. Malek · B. Ambalavanan
School of Communication, Taylor's University, Subang Jaya, Malaysia
e-mail: ramachandran.ponnan@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_22

1 Background of the Study

Of late there seems to be great motivation among students to be enrolled in private tertiary institutions. Enrollment numbers in these institutions from 2009 to 2013 have increased by an average growth rate of 5.6 and 13.6 % in public and private universities, respectively (MOHE). The urge to enroll in large numbers can be attributed to one of the several causes. An open policy toward a more liberal education system seems be foremost in the minds of parents as well. At the same time, budgetary concerns also push educationists to contain cost, within which large class size seems like viable cost-cutting strategy, only at the expense of students' performance (Cusoe 2004).

At the university campus itself, do large-sized classes affect students' performance and therefore their classroom relationship with lecturers? Typically, at the outset of university entry, students are met with large foundational classes. The rationale behind such an arrangement is largely the common modules students need to attempt and successfully complete within respective course structures before these students move on to specialized courses (Fig. 1).

The responsibility of teaching large-sized classes and unpacking learning objectives and responding creatively to students' needs are matters that correspond to student numbers. Lecturers who attempt to integrate critical thinking and active students' participation and learning practices do so to achieve better classroom performance. On the other hand, to spend eighteen hours and more per week on large classes, and small learning communities have a bearing on student–lecturer relationship much less the course delivery effectiveness and final students' evaluation on their lecturers. This situation is currently observed at the site of the research (Table 1).

MacGregor et al. (2000) articulate that large class size seems to go against the very elements of students' intellectual development and success. And that,

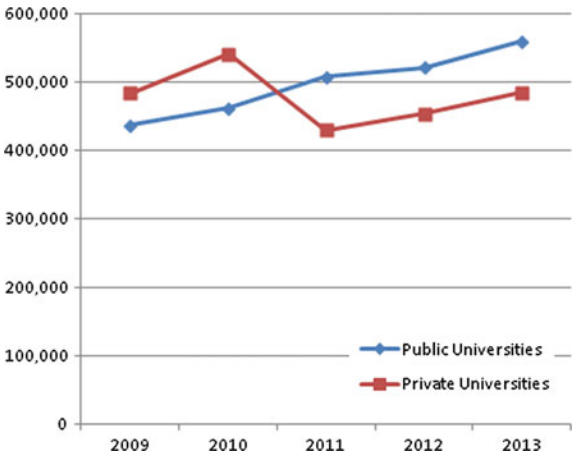


Fig. 1 Public and private university enrollments from 2009 to 2013

Table 1 Student population growth in Malaysia

Type of universities	2009	2010	2011	2012	2013
Public universities	437,420	462,780	508,256	521,793	560,359
Private universities	484,377	541,629	428,973	454,616	484,963

Sources MOHE (Ministry of Higher Education, Malaysia)

"inattention or absence from class and mediocre student performance seem to be tolerated simply as unfortunate realities" (p. 1). Such realities or unfortunate circumstances may be circumventing by carefully strategizing around the problem by early intervention (MacGregor et al. 2000).

As such, the study is motivated by current tertiary environment, scenarios, and facilities available on campus that promoted students' engagement and learning through large-sized classes and small tutorial groups and interjected with blended learning, as well as uniquely purposed classrooms called X-space that provide structured learning experiences.

Based on the experiences of the writers, the article explores the following research questions:

1. What student–lecturer classroom conditions are conducive to the learning process?
2. How large classes affect student–lecturer relationship?
3. How do small groups support student–lecturer relationship?

Literature reviewed suggests that there were several arguments against large-sized classes that form the basis for this study: Spear (1984) laid down a framework of eight statements to be answered, namely (1) over-reliance on lecture; (2) passive student engagement; (3) reduced lecturer interaction/feedback with students; (4) reduced depth of students' thinking; (5) reduced depth of course objectives and learning strategies used by students outside the classroom; (6) lower-level academic achievement and performance; (7) course dissatisfaction with the learning experience; and (8) lower student rating of course instructions.

The study attempts to answer the above questions based on the real-time experiences of the writers. The article explores the following research questions:

1. What student–lecturer classroom conditions are conducive to the learning process?
2. How large classes affect student–lecturer relationship?
3. How small groups can be organized to support effective learning?

2 Method

This is an empirical study by three lecturers involving a student population of 589 students in the third and fourth semesters of undergraduate classes, taught for two modules over two semesters in the same university setting. The study was done by

Table 2 Student populations observed over two semesters and two modules

Lecturer	Semester	Module 1	Module 2	Students interviewed
L1	2013 August 2014 April	131 145	29 34	4 4
L2	2013 August 2014 April	110 82	3 46	3 3
L3	2013 August 2014 April	5 4	- -	1 1
Subtotal		477	112	16
Total		589		

Fig. 2 Thought process of students and lecturers

conducting sixteen interviews with students who were vocal and who were the center of influence in the class to obtain students' insights. Next, classroom observations were done generally upon the whole population of 589 students collectively when they were in class. They were also observed randomly especially when interacting with lecturers outside during learning-related activities such as organization of event management, over two semesters (Table 2). The observation part was assisted by fellow lecturers who visited classes as peer observation every semester. Such observations were noted for the lecturer in attendance as well as part of the university evaluation process.

Initially, the interviews with lecturers and students were conducted as casual chats during their rest time after lecturers' classroom delivery and students' classroom attendance. When casual chats came up with the same comments repeatedly by more than one lecturer and several students, the repetitive comments were verified by all the three researchers and accepted them as perceptions to be further explored for detailed insights (Fig. 2).

The basis for using such a thought process or index is the evaluation survey report on lecturers by students at the end of every semester. Lecturers made frequent reference to their students' evaluation report. Students made direct references to their experiences in class and their relationship with lecturers. The evaluation report contains questions related to the teaching and learning quality experienced by students, and they do have a bearing on how best lecturers can deliver their classroom lectures and small-group tutorials.

2.1 Interviews and Observations

In-depth interviews were conducted with six randomly selected lecturers who were teaching the same cohort of students. Lecturers were asked structured and scripted questions supplemented by additional questions according to responses (Appendix 2). But informants were not asked more than one or two questions at the most at a time to keep it casual and authentic. At the same time, the study tried to avoid obtaining responses that sounded like a Q&A session. The same was done when interviewing students for their perceptions and insights on large class size and their learning outcome. The study ensured that at least two of the three writers were present during the interviews, while observations were done independently following structured guidelines (Appendix 2). Only repeated comments and perceptions during interviews were noted. The frequent occurrence of typical comments and perceptions was further clarified with the other writers for their validity. These comments were noted down and developed as themes (Table 3). The writers made their respective interviews, transcribed, and drew key themes as well as classroom observations and reported back to this study. All interviews and observations have been categorized according to

Table 3 Themes and subthemes emerging from students

Variables	Themes	Subthemes	Experiences
Large classroom conditions	Learning strategies	Assignments; exams; use of technology	Quality of lecturers; over-reliance on lecturer to provide content; low-level assignments and examination questions
	Academic performance	Self-learning	Level achievement and performance below expectation; less involvement
	Lecturer session	Attendance	Shortcut learning. Rote learning; passive learning during lecture; shy; absenteeism
S–L relationship	In-class behavior	Disruptive; participative	Mismatched interest; always ready to respond; enthusiastic in completing work
	Learning experiences	Satisfaction; dissatisfaction	Varying experiences; industry-related knowledge
	Attendance	Positive attitude	Sit close to the front; positive relationship with small class
Effective support for learning	Quality of instructions; out-of-the-class support	Monitoring students; evaluation report	Monitoring disinterested and interested students; lack of interest among students to provide sincere feedback and assessment
	Social media-based activities	Lecturer-led practice or activities	Learning-based activity appreciated; on the bench and out-of-the-class activity better suiting students
	Learning space	Designated space	Self-motivating, relaxed learning space; R&R open space; e-aquarium patronage

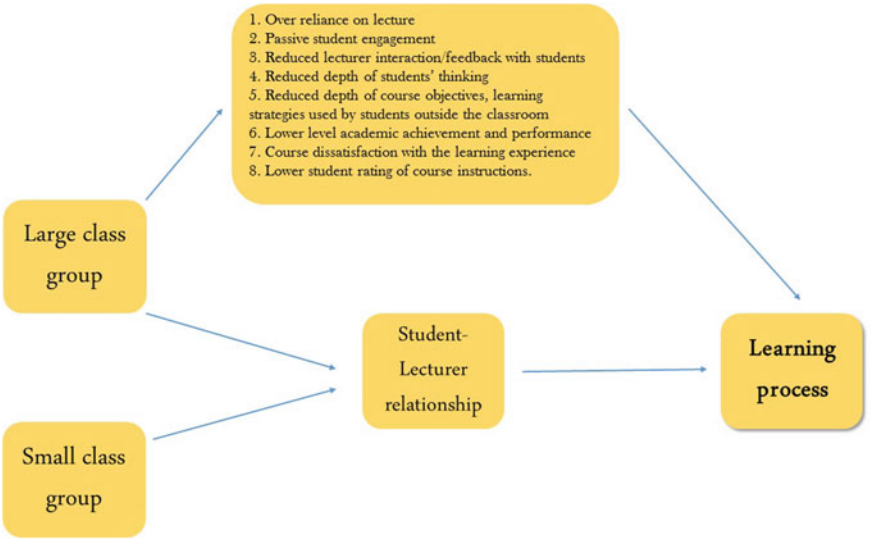


Fig. 3 Conceptual framework of the study

themes and subthemes following the problem and the questions to be answered by the study.

The framework Fig. 3 illustrates how large and small class groups affect student–lecturer relationship. The student–lecturer relationship is further conductive to the learning process. In addition, the writers propose the learning process as being further affected by questions arising from conducting lecturers in large class groups, i.e., over-reliance on lectures, passive student engagement, and lower level of academic performance (Spear 1984).

3 Findings and Discussion

The current discussion on the effectiveness of large classes and/or the redesigning of such classes or modification of classes into smaller tutorials to provide a learning environment that is helpful for student success shall be ongoing. The writers have focused on three emerging matters from the empirical study conducted over two semesters in one tertiary education setting. The data obtained through interviews from students and lecturers and observation of students have lent support for the tabulation of pertinent themes and subthemes given as in Table 3.

The writers make the argument that the success of the students, for which the lecturers and their institution work hard, is very much dependent on the collective responsibility of the students, lecturers, and the institution itself.

3.1 Classroom Conditions and the Learning Process

There is reduced depth of course objectives and learning strategies used by students outside the classroom. From the study made on the curriculum of the modules taught by all three lecturers, it has been learned that the design of the modules is usually likened to a large net to achieve broad objectives, and thus, it becomes excessively ambitious to capture a broad knowledge area. But this is nearly impossible to achieve at the delivery stage.

> I would say it is not so much about in-depth or reduced depth of course objectives as objectives set as a guideline and direction for instructors to conduct their course so that learning outcome will be achieved. However, it is my understanding that classes taught involving student activities will be more effective in achieving course objectives, said a lecturer.

> One student on the other hand, provided a different insight. For reduced depth of learning strategies used by students outside the classroom, I have something to say. A learning strategy is students' approach to learning and using information. Students who do not know or use good learning strategies often learn passively. Students nowadays don't really depend much on getting information by attending lectures regularly, but they will still attend tutorial classes, perhaps due to compulsory attendance.

When students are in class whether large or small, they are surrounded by lecturers, peers, and a support system that are conducive to learning. When they are outside the classroom, the simplistic and the most convenient way for them to obtain information, mostly for their coursework and assignment purposes, would be from the Internet. The study gathers that students most often tend to take the face value of what is printed on paper. It is uncertain that they may be able to make credible and valid judgment on the information obtained from the Web. Students overlook the reliability factor when sourcing for information other than recommended references.

3.1.1 Level of Academic Achievement and Performance

The researchers contend that the experiences, teaching in academic settings, point to students' drop in academic achievement and performance. This seems to be a common factor that cropped up in discussions among fellow lecturers and colleagues. The same thoughts prevail among lecturers from other tertiary institutions. Emanating from such discussions are excuses such as the poor quality of student output that arrive at the doorsteps of universities. It is also the unavailability of suitable and subject matter matching lecturers. Generally, students are well informed and they seem to know much on the surface, but indeed, they lack in-depth knowledge on the same matter. This scenario now points to the lack of interest in reading beyond the needs of comprehension. One factor that points to "shortcut" learning can be realized from, "sir would you be giving examination tips at the end of the module?" ask many students as the semester comes closer to the end.

In terms of creativity, students tend to have similar likes and dislikes, which means most of them might just like K-pop (Korean popular cultural products). But few of them would have differences when it comes to music flavor. This may occur due to differences in cultural background of our students. Most of our students come from culturally Chinese-educated background. This may explain the tendency toward similar taste and thinking, the idea they contribute for assignments or in their case of study (teaching practical module), the idea of producing a short film, is very much the same. It is not to assume that they are not creative, but their very "commercial" and "safe" attitude, not really "out-of-the-box" thinking, results in being so stereo type. This was seldom the case a decade back.

When courses are taught in the English language and students have difficulty in comprehension, students loose interest. Students have openly declared they do not understand their lecturers. For example, when students asked to relate to their favorite radio station in Malaysia, quite a number of students choose to discuss either MyFM or One FM because the main language used in these radio stations is Chinese. They will not choose Hitz.FM or other English radio stations. This exemplifies the scenario where students tend to misread instructions provided during lecture or tutorial sessions.

To a question, "how do you learn in class?" The student response was, "we dislike taking notes or jot down their new experiences and what they hear in class. We prefer taking photos of the notes written on the board. We are pretty much lecturer-dependent," which means leaving no room for deep thinking on cognitive matters deliberated in class. During large class lecturers, students dwell much in the passive learning zone. They are just waiting to mark attendance at the end of the lecture to get on with their own business. In their responses to assignments, it becomes evident that there was not much reference made to key text provided during lecture or searched on their own, which only implies that majority students do not read the text provided at the end of the lecture.

3.1.2 Dissatisfaction with the Learning Experience

There is growing dissatisfaction among students with their learning outcome from the courses taught. Students who arrive into a class at the beginning of every module come with a variety of experiences and at varying standards of cognitive preparedness and learning skills. When these students find themselves put in large classes where the lecturer stands far away and speaks with the aid of a microphone, the students feel distanced. At times, the attitude of students and lecturers is indifferent until students are broken down into smaller tutorial groups.

There were some mature students who found the standard of instructions below their expectation. It was understood that such students were working in the industry at the same time of study and that they knew the subject matter well. As such, their thoughts and experiences would be racing into broader and deeper knowledge when compared to the rest of the class. Such students raised their dissatisfaction in their evaluation report of the lecturer putting the lecturer on the line instead.

When students returned from their internship stint, they always "complained" about the differences in instructions during their course and that they had experienced differences in the industry practices. It is for the same reason that students are sent out to acquire different and up-to-date practices from the industry. It is an additional learning experience introduced at the formative stages of the course to accommodate all the needs of students as the course is being designed. However, course evaluations have always been done to learn whether the curriculum matches up with industry trends with the expert insights of industry panelist. Students have to be made to understand the scenario in the classes and those they experience in the industry are much different and lecturers are trying to simulate such experiences in a classroom environment at best. One lecturer commented:

> "Some lecturers with industry experience have an advantage as it would be good when they can talk the industry language and conduct tasks just as it may happen in the industrial scenario". He continued, "tertiary institutions are very much places about learning a combination of theory, practical skills, and some research in a teamwork environment, involving problem-solving, time management, critical thinking, etc. which are all designed to be learned in a subtle way. Students may not realize this until they go out to work".

3.1.3 Rating of Course Instructions

According to students reporting back on the quality of instructions, they believed that students' evaluation of course instruction does not affect them. The study finds that they lack interest to rate course instructions as they think it is not important to them and it will not affect their learning. As can be understood from annual appraisal of lecturers, only about 5 out of 30 students in a small class perform the ratings. As for the large classes about 25–30, actually perform the evaluation. Students are very much passive in rating their course instructions or the instructors. Unless they had an issue that students had to bring up, they would not provide serious comments on course instructions when there are sufficient formal reporting pathways. At times, they resort to reporting through social media or resort to bringing it up straight to the dean's attention. Lecturers have the burden of explaining the importance of such evaluation and that it leads to better quality of instruction in the future.

Next, students' ratings are somehow influenced by their peers and collective motivation. Students generally tend to give lower ratings for course instruction delivered in large classes citing instances of lack of clarity, student behavior, disruptions, and monotony of lectures.

3.1.4 Lecture Sessions

Based on the experiences gathered, lecture sessions with more than 100 students are thought to be ineffective for instructions and too big for effective class management. One senior lecturer asserted, "… most students in such classes were not really paying attention as everyone minded their own business even if you were for

instance talking to them about some interesting stuff. They would be at their FB or devices." Although at the beginning of the semester, students seemed to behave well and attended lectures regularly (*attending lecture is not compulsory in our institution but tutorials are*), gradually week after week, it becomes obvious that the number of students attending class would drop. This is quite common as large-lecture attendance decreases throughout the semester and it often drops down to about 30 % by the end of the semester. The common excuse is that students were rushing to complete assignments.

One factor for poor class attendance is modules with 100 % coursework and no examination. Another factor is "attention-grabbing" in a large class is harder as there would be too much distraction in the hall. Distractions include noise and students' conversation. Poor attendance and inattention do contribute to poor performance at the end of the semester. With a large class, motivating students is an uphill task given the impersonal nature of the class and the lecturer and the lack of individual accountability. Students' attention-grabbing and concentration tend to drop dramatically after 20 min into the lecture. At times, lecturers do not even know the names of their students given the size of the campus and students as the groups are too big.

A large number of students in class do deter students from asking questions, even if the lecturer encouraged them to do so. Large class size reduces the quality of lecturer's interaction. Lecturers tend to provide brief feedback and want to get on with their planned lesson. It is noticed that teaching effectiveness decreased with the increase in class size.

3.2 Student–Lecturer Relationship

In teaching core modules where the size of students is normally large, it is always difficult to measure the effectiveness of the lectures. Students' understanding is tested and measured mostly during tutorials when the class size is about 20–30 students. This allows for the creation of conductive student–lecturer relationship that will enable the lecturer to adopt specific techniques in content delivery or structure hands-on exercise. In the experience of one of the writers, to reinforce learning in class, students with some knowledge in the subject matter or who know better would be asked to explain and share their opinion. Such exercises do boost the confidence of students and at the same time get the attention of their peers.

3.2.1 Students' in-Class Behavior

In building student–lecturer relationship, it is important to know the characteristics of young students in their class. Students are by nature active, well informed, tech-savvy and brave enough to explore. However, systematic traditional education is something that they claim to be boring as such they will mind their own business.

Students have made it known that they prefer lecturers who adopt a motivational role; they prefer to engage in an interactive learning process. Also, students nowadays do not adapt well to lecturers who typically put students at a higher self-directed learning position; students preferred lecturers who provided direct and simple instructions and who interact more on a personal basis with them, be it in the classroom or outside of the classroom. When it is a large class, students expect lecturers to afford some personal attention when the class is broken up into smaller classes.

For close interaction with students, lecturers required to spend more time, a precious resource to be shared with students in large-sized classes. Yet, they prefer personalized information that will simplify their tasks; they want you to hear their own voice in action as part of their learning process.

3.2.2 Small Class Learning Experience

Data collected from students who were from small classes manifested a positive student–lecturer relationship. "The good number for student–lecturer relationship to work would be 15–20 students to a tutorial class," the students claimed. They also believed that small-sizes class provides ample space and opportunity for the creation and enhancement of learning experiences. What students felt was important, "here students get instant feedback and they can realize the learning that takes place immediately." With a small number of students, attention-grabbing is possible to be achieved. Time can be devoted to discuss subject matter; students will be encouraged and confident to ask questions in their group instead of asking questions along the hallway. Students shy asking questions in front of their peers. Observations of such shy students reveal that students are afraid of asking wrong questions or questions that sound silly, especially among those who come from non-English-speaking or vernacular schools. Some students feel there are much "cultural" restrictions that hold them back to be standing or talking before their classmates.

In contrast, small-group discussion helps the lecturer to deliver materials with ease and gets to know each individual's strength and weakness by observation throughout the semester. Small groups definitely encourage students' active involvement. Lecturer has control over the arrangement or design of the classroom to effect better learning experience, such as sitting in a circle for discussion or meeting arrangement for a media strategy sales pitch.

3.3 Organization of Small Groups' Support

One student who has been in the university for two years said,

> "… but I would say it's more for questions and answers being discussed during tutorial classes, and these are what students think they really need for exam preparation". But

> students tend to forget that in the lecture (which usually it's a big classroom), information is given by the instructor and it's supported by audio-visual materials. Students would think a lecture is boring but instructors nowadays have been thinking of how to enhance the lecture experience supported with technology different from a typical lecture session.

A few students have also cited that social networking sites are a great resource popular among student. For instance, Facebook has become a platform for lecturers and students to discuss, to inform, and to share course-related matters. Faculty members are encouraged to set up social media groups to interact with students about coursework-related matters. This is especially becoming useful as students may not be accessing student portal or email frequently. But they are surely logging on to Facebook every day.

It was learned that to simulate in-class or real-time experience to support understanding, especially in small-group classes, lecturers tend to ask the entire class a practical question that relates to the topic under study, where it was sort of compulsory for everyone to answer the question in an effort to enlist their active participation in the assignment. They would role-play in a rota of different positions (as in a production) and at the same time work with different partners. When they are active, that is having the opportunity to move a little, shout opinions and suggestions by which time students get active. Whether they learn the salient content while they are active is not quite clear unlike rote learning.

The learning environment be it large-sized class or small tutorials has been steadily moving away from one-way instructions to self-motivated social learning facilitation, and the e-aquarium is one of them. That urges students to learn on their own in groups doing their assignments. There are sufficient deliberately designed learning spaces strategically located all over the university for this purpose. These spaces are often near food and drinks outlets, powered and wired for connectivity. This is a great challenge for the tertiary institution administration to implement and is specifically for most lecturers to monitor an enhanced learning experience.

4 Recommendations and Conclusions

The study finds that teaching in large classes generally has adverse effects on students' learning and performance that are contextual and reliant on three factors. The writers conclude that the typical organization of large classes and the subsequent breaking down into smaller tutorial groups have a bearing on both student and lecturer's quality of performance. Next, student–lecturer relationship nurtured that determines the quality of engagement over time in and out of the class contributes to the satisfaction and dissatisfaction of students' experiences. And thirdly, the effectiveness of support is provided by lecturers and the institution in realizing students' objectives and organizational goals.

The study also makes a few suggestions on how lecturers may want to avoid displaying specific behavior.

- Avoid personal emotion, especially for female lecturers, emotional rollercoaster tends to happen occasionally, should be avoided.
- Avoid being bias. It is important for a lecturer to be fair and reasonable to each of the students.
- Always be well prepared to the class. Students would know that if a lecturer is prepared for class or if a lecturer is just dragging time.
- Maintain a professional relationship with students. It is always advisable to not engage in any form of conflict.
- Students constantly need much motivation and inspiration and help they can get from lecturers.

Lastly, young students are tech-savvy that they shun from the typical teaching methods. They resent to lecturer who speaks from a higher position, talking down at students in a lower position. Students prefer active interaction among students and with lecturers. They want to treat lecturers as their friends and guiding pillars. Lecturers are asked to make a connection with students, which also means the relationship is built not only around the classroom setting, but out of the classroom. It is quite possible to encourage, motivate, and persuade students to the right track with some outdoor activities including a simple lunch, a coffee break time, or a simple gathering after class or during weekends. There are even some colleagues holding consultation hours out of official working hours. And the consultation is not necessarily to be conducted in the university, but out-of-the-campus activities. Some form of engagement is needed when it comes to teaching the younger generation. Besides, lecturers also have to be bold enough to make changes to their instructional methods in tandem with the needs of students.

Appendix 1

Observation Guidelines

Classroom Behavior—students:

- Students' arrival to class
- Leaving class with/without permission
- Disruption, talking, playing the fool
- Inattention
- Preoccupied with devices

Interested in class:

- Asking for clarification
- Asking questions
- Responding with suggestions
- Showing signs of prior reading, preparedness
- Taking notes
- Having completed prior work assigned
- Able to recall previous lessons covered

Appendix 2

Interview Questions for Students

1. How do you like you're learning to be delivered in the classrooms?
2. Do you depend on activities organized by lecturers in trying to learn in and outside class?
3. To what extent can students be expected to do their own research to learn?
4. How do you learn face to face while in class attendance?
5. What approach do you adopt to learning using the facilities in the university?
6. What industry-related experience do you get at the university
7. Do students understand the consequences of students' course evaluation?
8. Generally, what is the satisfaction level of the students in terms of their end-of-semester performance?

Appendix 3

Interview Questions for Lecturers

1. Do you always follow course objectives closely when teaching your module?
2. Is there a drop in student academic achievement and performance?
3. What is the student classroom behavior like?
4. How would you organize classroom participation especially when student numbers are large?
5. Is the standard of instruction suitable for all students?
6. How do you gauge the effectiveness of your teaching especially in large classes?
7. Mr. ….., as you are teaching in both large and small classes, would you be able to identify students' level of response to oral questions?
8. How do you think students are rating instructions delivered in large classes?

References

Cusoe, J. (2004). The Empirical Case Against *Large Class Size:* Adverse effects on the teaching, learning and retention of first-year students. Retrieved from http://www.classsizematters.org/wp-content/uploads/2012/11/Week-13-Cuseo-1.pdf on the 12 June 2015.

MacGregor, J., Cooper, J. L., Smith, K. A., & Robinson, P. (2000). Editor's notes. In J. MacGregor., J.L. Cooper., K.A. Smith & P. Robinson (Eds.), *Strategies for energizing large classes: From small groups to learning communities.* New Directions for teaching and learning, no. 81. San Francisco: Jossey.

MOHE, http://www.iie.org/en/Services/Project-Atlas/Malaysia/Higher-Education-Sector. Retrieved 1 Nov 2015. http://www.ncte.org/positions/statements/why-class-size-matters. Retrieved 30 Oct 2015.

Spear, K. I. (1984). Editor's notes. In K. I. Spear (Ed.), *Rejuvenating introductory courses* (pp. 1–9). San Francisco: Jossey-Bass.

Map Interpretation: Tool for Rapid Learning and Assessment Lens for Cognitive Engagement

Arniza Ghazali

Abstract The final-year students electing materials analysis as their course were coached in understanding the concepts of analysis for one semester. Analysis of students' learning experience shows that systematic development of understanding of the topical contents enabled them to construct the topical maps. Despite students' inept mapping skill and inability to see the whole course contents as a network of linking topical maps, students who progressively encoded the knowledge by full attendance of lecture sessions were able to rapidly grasp the course's masterpiece map constructed by the instructor (i-MAC). Given the same time frame as other students, these regular participants also performed extremely well in the final analytical challenge assessing their ability to solve analytical problems, commensurate with the set course learning outcomes. The leap from B to A and even C to A of this group was the outcome of the systematic arrangement of important concept gathered through progressive intellectual development. Upon introducing the topical linkages in i-MAC, grasping of the concept among them was automatic, unlike the extra struggling required of the counterpart. This was assessed from the correctness of their verbal map interpretation (VMI), which is reflective of the final assessment scores. Deliberation on students' learning identifies the importance of (1) systematic encoding of knowledge, (2) intensified neural activation especially achieved via digestion of i-MAC and the subsequent VMI and (3) instructor's reflections, needs analysis and carefully designed intervention in the form of i-MAC and feedback to VMI, to engender high-order cognitive engagement. Proving correlation between extent of VMI correctness and performance in the final analytical challenge, verbal i-MAC interpretation, therefore, also serves as a reliable tool for assessing students' degree of cognitive engagement and their preparedness to function in problem-solving situations.

Keywords Learning · Cognitive · Engagement · Mapping

A. Ghazali (✉)
Universiti Sains Malaysia (USM), 11800 Gelugor, Penang, Malaysia
e-mail: arniza@usm.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_23

1 Introduction and Literature Review

By the embedded meanings in the proposition, "No true educator simply wants to 'teach'; educators want students to 'learn'" (Banikowski 1999), the teaching paradigm has shifted from emphasising "what students are and how to teach them" to "what teachers do and how to improve them" to the recent, "What students do…" as phrased by Biggs in 1999 (Davies 2010) to make meaning of their learning. Within the existing paradigm, teaching activities are to be designed to ensure that students are engaged in meaningful learning, as a measure for bridging the gaps in performance by providing support for the less engaged to move their learning to a higher continuum: to develop the ability to link ideas, remembering them and using the acquired skills and knowledge later on.

Good learning is indeed facilitated by thinking ability (Cole 1971; Cromley 1998), which is in the true sense, cognitive engagement. The process is fuelled by motivation, a subcomponent of affective engagement, which ultimately manifests in the learners' actions (behavioural engagement). Not limited to the fields of medicine, science, technology, engineering and mathematics, learning is fruitful when accompanied by an active brain work. Synonymous to cognitive engagement, true learning process involves the working mind, i.e. the ability to process information online (*that in turn enables cognitive flexibility, abstract thinking, strategic planning, access to long-term memory and sentience*) (Dietrich 2004) as recorded through the lenses of neurosciences and brain-based learning (Davies 2010).

To engender the aforesaid cognitive engagement, knowledge mapping, known as a tool for transforming abstract knowledge and understanding into concrete (linking) visual representation (Hay et al. 2008), is one of tools that instructs thinking activities and promotes meaningful learning. The effects are made more profound by executing the activities by hand-drawing (Bounds 2010) the principal ideas, connecting them with curved lines, organising the points in hierarchy or radial arrangement and associating the main points to the breeding points. Inclusive of hand-drawing mode, mapping amplifies learning activities by "*processing information verbally and pictorially (that) helps learning happen* via *more than one modality*" (Davies 2010). This is achieved through separate encoding of information in the memory by visual and propositional forms. The processes involved, which activate multiple senses associated also to intense activation of various parts of the brains, help enhance students' readiness to learning and thus access more meaning in the learning material (Shams and Seitz 2008). The hand-drawing mode also allows adoption of the best usable features of the mentioned digital maps. Moreover, map construction, which demands for high-intensity mind work analogous to struggling (Davies 2010), is posited by Stigler (2012) as the slow and sticky processes essential of learning.

Learners are, however, diverse in their abilities, and even the powerful approach such as mapping needs careful design to best engage learners to learning. The workability of an instructional approach cannot be judged by evidence of physical engagement of learners, which is oftentimes deceiving. Ineffective instructional

approach may have born attribution to the current phenomenon of graduates having lower-than-expected thinking proficiency (Kuldas et al. 2015). Students' inability to construct a map of learning materials is a sign of setback in cognitive proficiency, and this was exactly the situation of a group of final-year students of focus in this study. The incident needs immediate fixing in order to prepare the cohort to achieve the final course learning outcome (CLO 3), which is to be "able to solve issues in industrial analytical challenge", as simulated in the final examination questions. The inquiries at hand were (1) how would verbal interpretation of an instructor's course map (VMI) as a last minute intervention support learning and (2) does VMI serve as a good gauge of students' learning?

To answer the aforesaid questions, two of the unique principles of the Scholarship of Teaching and Learning (SoTL) were embraced to understand students' learning experience, i.e. inquiry into students' learning (de Braga et al. 2015), and conducted in partnership with them (Poole and Chick 2015). The whole idea is to understand the way VMI facilitates last minute learning of a group of students who were unable to construct the map of the course. This is a serious sign of poor cognitive level analogous to the current issue in higher education reported by Nagappan in 2010 as posited by Kuldas et al. (2015). By charting the cognitive milestones in students' learning processes, a model is to be developed to identify the suitable approach that supports attainment of higher order thinking skills, HOTS. In the teaching contingent, the study also aims to demonstrate the possibility of VMI as an assessment tool to gauge learning.

2 Methodology

In the 2014/2015 academic year, ten out of nineteen candidates were enrolled for a third-year elective course, IWK307 Advanced Paper Technology—Instrumental Analysis for Pulp and Paper (hereby denoted as "materials analysis"). In each class, a 40-min lecture was delivered as a guide to the scope of each topic. While various activities were designed throughout the semester, students were assigned the task of constructing the course concept map as the major coursework with the aim to indirectly instruct a powerful cognitive engagement.

Students' engagement experience was studied using four instruments: teacher ratings of students (TRS), instructor's cyclical reflections (ICR) and student self-report (SSR) as suggested by Fredericks and McColskey (2012) (Table 1).

The last ICR was performed to assess the needs for corrective measures for the discovered pseudo-learning of the course materials. This prompted the design of a final intervention to support the attainment of CLO 3. As a result, a course map was constructed by the instructor (i-MAC), and students' verbal interpretation of i-MAC

Table 1 Instruments and their uses

Instruments	Frequency and domain of engagement in focus	Monitored variables
• Teacher ratings of students (TRS)	TRS1: Weekly (A&B) ….	Attendance, participation
	TRS2: Bi-weekly (C) …..	Assignments
	TRS3: Week 14 (C) …….	Map construction ability (MAC)
	TRS4: End of semester (C)	Final assessment
• Instructor's cyclical reflection (ICR)	Weekly (A, B & C) …….	Learning needs
• Student self-report (SSR)	Once: Start of semester break (A, B and C)	Course learning experience
• Interview	Multiple phases during semester break (C)	Analysis on requirements for intervention to work
Key—Engagement domain: A ≡ Affective; B ≡ Behavioural/Physical; C ≡ Cognitive		

(VMI) was heard in order to offer instantaneous feedback as corrective measures for misconception. Students were later given the freedom to have a snapshot of the map as a revision material. The i-MAC was for students to digest before the next hearing of their VMI.

3 Results and Discussion

From TRS1, the students were found to appear in the grouping presented in Table 2, with the count of "+" marking the instructor's impression of learners' affective, behavioural and cognitive engagement.

Table 2 Grouping of students by Instructor's impression and TRS1

Candidate		Attendance & state (2014)	State (2011)
1.	G1-1	100 %; +++	+++
2.	G1-2	100 %; ++	+++
3.	G1-3	100 %; +++	+++
4.	G2-1	93 %; ++	+++
5.	G2-2	93 %; +	+++
6.	G2-3	80–90 %; ++	+++
7.	G3-1	80–90 %; +	+++
8.	G3-2	<80 %; +++	++
9.	G3-3	<80 %; ++	++
10.	G3-4	<60 %	Unnoticed

3.1 Assessment of Students' Initial Learning

Performance in the major coursework, however, was substantially challenged by their normal pseudo-learning practices such as rote learning, and this denies correlation with the findings of TRS1 in Table 2 due to the unexpected students' ineptitude in constructing the course map. This was evident from the submission of individual topical maps presented on loose sheets of A4 papers. The important diagnosis was not just inability to link the topical maps but also students' inability to visualise the big picture of the course, indicative of the lack of mastery of the taught concepts. Of primary concern is the dysfunctionality in problem-solving situation. The glaringly low marks obtained by G1-2, for instance, were due to the incompleteness of the map, on top of the inability to connect the topical maps as one complex course map. Consequently, the presented maps were disconnected

Table 3 Leap in the final analytical challenge

Students	Rank and grade		
	Pseudo-course map*	Analytical challenge**	Students' feedback
G1-1	1 (B+)	2 (A)	"(i-MAC) provides eye bird view (bird's-eye view), connection, fast scanning, power memory, creative, simple, attractive and comprehensive (presentation of the learning materials)" "Verbal description of the constructed map helped a lot" "I am happy with my final score"
G1-2	**10 (C)**	**1 (A)**	"Description and discussion of the constructed map (i-MAC) helps a lot" "I am contented"
G1-3	2 (B−)	3 (A)	"Map is easy to understand. Easy to digest. Discussion on map helps a lot. I get results on target" "Make more mind map"
G2-1	4 (B)	5 (A−)	"Map is easy to understand. Easy to digest. Discussion on map helps a lot. I get results on target" "Make more mind map"
G2-2	5 (B)	6 (B+)	"Discussion on map helped a lot"
G2-3	**6 (B)**	**4 (A)**	"I tried my best to score. The subject is very useful and the final mindmap that Dr made for us helps A LOT! Thank you ☺"
G3-1	8 (B−)	7 (B+)	"I find the course very tough." "Map discussion and description helped a lot. I prefer one-on-one consultation" (Translated)
G3-2	7 (B−)	8 (B+)	"Description of the constructed map and discussion helped a lot" "I am happy with my final score because the score reflect(s) on my effort."
G3-3	9 (B−)	9 (B+)	"Map description and discussion helped a lot" "I am ok with my final score"
G3-4	3 (B)	10 (B+)	"Map description and discussion helped a lot but it should have been given much earlier"

course maps, hereby described as "pseudo-course map", suggesting the lower-than-expected cognitive engagement (60–75 %, C to B+). To rectify the situation, ICR was performed and an instructor's map (i-MAC) was introduced to the students who were to verbally interpret the map before the instructor. The outcome of the intervention is depicted in the grade for the analytical challenge (Table 3).

VMI as an intervention was witnessed as engendering improvement. The transformative outcome of i-MAC with VMI intervention is portrayed in the improvement in the scores in Table 3. The favourability is justified by the positive learning experience recorded in student self-reports, SSR, hereby noted as "feedback".

3.2 Postmortem Analysis of i-MAC with VMI Intervention

To understand the way i-MAC works, a flashback on the VMI session is required. Candidate 1 (G1-1) came in a group with G2-1, G2-3 and G3-4. At the end of the discussion, G1-1 seemed to master the concept and G2-1 required a little bit more digestion, while G2-2 remained contaminated with misconception of the topical relations. Apparently, G2-2's misleading understanding of the contents was not of the school-made (Barke et al. 2009) type, i.e. by acquisition of misleading contents from the subject matter expert (SME). Rather, it was the candidate's own misinterpretation of the learning materials. Despite resolving with other candidates (G1-1, G2-1, G1-2, G2-3), G-2-2 seemed to have the encrusted misconception and it was very difficult to disrupt and subsequently rectify the old understanding. Regarded as "misconception" or the self-developed concepts that do not match the taught concept, Thompson and Logue (2006) amplified the seriousness of misconception as posited by Eggen and Kauchak in 2004 as follow:

> The biggest issues are that once a misconception has been formed, it is extremely difficult to change and that processing misconceptions can have serious impacts on learning.

Indeed, this explains the fate of G2-2 who only managed a B+.

The i-MAC could be relatively harder for G3-4 who recorded poor attendance (Table 2). This denied systematic encoding of the knowledge, which is essential in learning. The importance of this is portrayed by the top score of G1-1 who engaged with full attendance and G1-2 who asserted:

> Attending lecture is important to slowly build an understanding of the course contents thoroughly (Communication via Whatsapp translated from Bahasa Melayu to English).

In the neurosciences, the orderly knowledge development is analogous to the embryological development of the brain involving rough drafts (in the absence of experience) advancing to a more complex circuit structure critical to the nervous system development (Marcus 2015). Apparently, G3-4 exhibits none of the expected cognitive engagement throughout the coaching period. The condition was

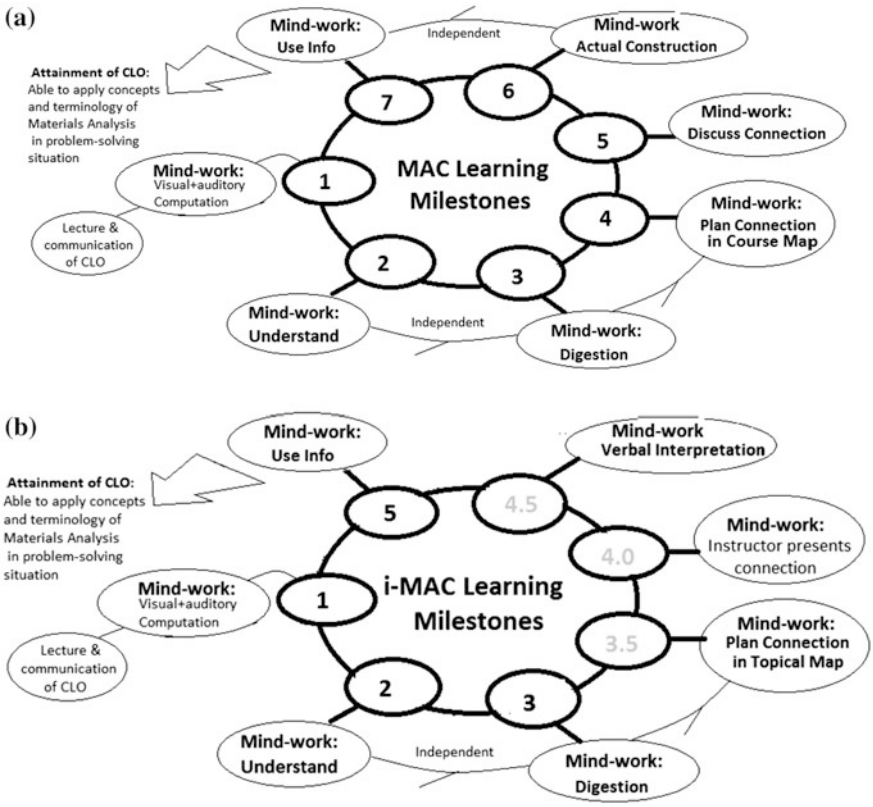


Fig. 1 Learning milestones in mapping activities. **a** The expected seven milestones from course map construction (MAC). **b** The five milestones from interpretation of i-MAC

hard to support considering the time constraint and the history of the missing basic milestones (Fig. 1a) and candidate's unreadiness to discuss the course materials (milestone 5, Fig. 1a) at the problem-solving stage that assumed prior preparation (milestone 2 and 3, Fig. 1a) by the students.

Despite the absence of MAC (course map construction by students) learning milestones, the i-MAC still offers five of the mind work avenues (Fig. 1b). The first layer of knowledge came through attendance of lecture session, which is accompanied by activities assessing students' learning and further needs. It is an added learning mile if students had spent time looking through the materials if not learning to understand and reinforcing their understanding by digesting the materials. While construction of the course map completes the learning loop with major milestones 5 and 6, construction of unlinked individual topical maps removed milestones 5 and 6 and offered a partial milestone 4, as portrayed in the coursework scores in Table 2. With i-MAC, students transferred milestones 5 and 6 (Fig. 1a) to the efforts

underpinning verbal interpretation of i-MAC (or VMI). Thus, the extent of correctness of VMI seemed reflective of the level of learning–assessment tool.

VMI and i-MAC, however, could be maximised for equivalent effects of the seven thinking milestones in MAC process (Fig. 1a). The emergence of the right-to-left brain function that could have been an intensive cognitive engagement activities achieved via milestones 4, 5 and 6 in MAC (course map construction) was lightly experienced in milestones 4–4.5 through the i-MAC (Fig. 1b) due to the ruled out planning and organising and intensive linking tasks. It is paramount to note, however, that repetitive VMI as strategised by G1-2 supported summative learning effects analogous to undergoing the total of seven thinking milestones in MAC (Fig. 1a). This is demonstrated by the leap from C to A cognitive performance proven by G1-2 (Table 3). Furthermore, the gradual enhancement of correctness of this VMI further supports the validity of VMI as gauge of cognitive engagement.

3.3 Rationale to Rapid and Visible Cognitive Engagement

The commonality shared among the A scorers was appreciation of i-MAC, which partly overlaps with being highly teachable. This is portrayed by the rapid adoption of a new corrected idea substituting the misconceived ideas detected through verbal description of i-MAC. Such level of appreciation could be attributable to the strong logical flow established by perfect continuity and order behind the knowledge encoding process. TRS4 (Table 1, analytical challenge in Table 3) shows that a majority of these students turned up for lectures fully and this marks the important role of attendance as support for learning fruitfulness previously reported by Rolka and Ramshagen (2015). McLeod (2007a), in this regard, posited that the way we store info affects, the way we retrieve it. Thus, a gradual stepwise input of knowledge provides a systematic development of the body of knowledge into the memory bank. This involves stages of encoding, storage and retrieval (McLeod 2007a), making the information more accessible and highly usable. The systematicity ultimately leads to students' rapid switch to the corrected conceptions due to the mind's preference for the more logical relations of the systematically learnt concepts. In problem-solving situations, the knowledge becomes more easily retrieved by sequence (from short-term memory bank) as well as by the logical association (from long-term memory bank) (McLeod 2007a), resulting in better performance among the regular participants (Tables 2 and 3).

Moving on to learning depth, McLeod (2007b) pointed out that deep learning involves semantic processing of knowledge achievable through elaboration and rehearsal. Interventions giving more meaningful analysis of images and thinking associations analogous to i-MAC, therefore, had provided a deep learning approach, despite MAC inability. The exercise of giving link to the previously acquired knowledge mimicked the bottom brain and top brain relations proposed by Kosslyn and Miller (2013). Being newer, the i-MAC intervention, thus, was

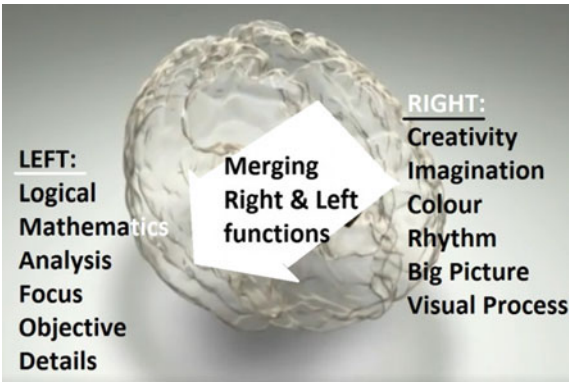


Fig. 2 Learning milestones supported by verbal interpretation of i-MAC merged the *left* and the *right brain* functions, making rapid grasp of learning materials

received by the top brain that drew information on the bottom brain—where the prior knowledge acquired from lectures, activities and taken notes were stored.

Marcus (2015), in this regard, offered an insight into deep learning as an unsupervised system that can form categories for themselves, also with emphasis of step-by-step computation. On a similar account, McLeod (2007a) emphasised that assigning difficult task is a gateway to deeper learning. By the same principle, apparently, verbal interpretation of the course map, which students regarded as hard-to-construct work piece, instructs learning by making sense of the texts, colour and associations. The so-called complex but rather superb process of engendering more learning worked through the bridging between the predominantly left brain function to the right brain functions (Fig 2). VMI of i-MAC, therefore, is a sticky learning process worth experiencing due to the accumulative intense neural activation, or the high-order cognitive engagement.

4 Conclusion

Verbal interpretation of the instructor map, i-MAC (VMI) was powerful at enhancing learning via several milestones of cognitive engagement. Owing to the intense brain activity involved in VMI, it promoted rapid focused learning analogous to the way learning is enhanced in the course map construction activity. The rapidity was supported by several factors including systematic encoding of knowledge in the memory by fully attending to the input process designed by the SME (instructor). VMI being one of the strategies for high-order cognitive engagement simultaneously acts as a valid gauge of students' learning and preparedness to function in problem-solving situation. A longitudinal research is recommended to demonstrate correct adaptation of the approach to the various learning contexts to scaffold better functioning knowledge workers to man the various sectors in an industrialising country.

References

Banikowski, A. K. (1999). Strategies to Enhance Memory Based on Brain Research Focus on Exceptional Children 0015511X32(2):1–22. Retrieved 17 July 2015 from, http://sc-boces.org/English/IMC/Focus/Memory-strategies2.pdf.

Barke, H. D., Hazari, A., & Yitbarek, S. (2009). Chapter 2 students' misconceptions and how to overcome them. In H. D. Barke, A. Hazari, & S. Yitbarek (Eds.), *Misconceptions in chemistry.* doi:10.1007/978-3-540-70989-3-2. (pp. 21–24), Berlin, Heidelberge: Springer-Verlag.

Bounds, G. (2010). How Handwriting Trains the Brain Forming Letters is Key to Learning, Memory, Ideas, *The Wall Street Journal.* Retrieved 4 March 2015 from http://www.wsj.com/articles/SB10001424052748704631504575531932754922518.

Cole, M. (1971). *Cultural Context of Learning and Thinking: An Exploration in Experimental Anthropology* Eric Number ED062465 (p. 298). Retrieved 31 Aug 2015 from, http://eric.gov/?ti=Racial+Differences.

Cromley, J. (1998). *Learning to think, learning to learn: What the science of thinking and learning has to offer adult education*. National Institute for Literacy.

Davies, M. (2010). Concept mapping, mind mapping and argument mapping: What are the differences and do they matter? *High Education.* doi:10.1007/s10734-010-9387-6.

de Braga, M., Boyd, C., & Abdulnor, S. (2015). Using the principles of SoTL to redesign an advanced evolutionary biology course. *Teaching and Learning Inquiry—The ISSOTL Journal, 3*(1), 15–29.

Dietrich, A. (2004). The cognitive neuroscience of creativity. *Psychonomic Bulletin & Review, 11*(6), 1011–1026.

Fredericks, J. A., & McColskey, W. (2012). The measurement of student engagement: A Comparative analysis of various methods and student self-report. In S. L. Christenson (Ed.), *Handbook of research on students engagement* (pp. 763–782). New York, NY: Springer.

Hay, D., Kinchin, I., & Lygo-Baker, S. (2008). Making learning visible: The role of concept mapping in higher education. *Studies in Higher Education, 33*(3), 295–311.

Kosslyn, S. M., & Miller, W. (2013). *Top brain, bottom brain: Surprising insights into how you think.* (ebook pp. Location 90 of 2995). New York: Simon and Schuster Inc.

Kuldas, S., Hashim, S., & Ismail, H. N. (2015). Malaysian adolescent students' needs for enhancing thinking skills counteracting risk factors and demonstrating academic resillience. *International Journal of Adolescence Youth, 20*(1), 32–47.

Marcus, G. (2015). The Computational Brain. In M. Marcus & J. Freeman (Eds.) *The future of the brain: Essays by the world's leading neuroscientist* (pp. 205–214). USA: Princeton University Press.

McLeod, S. (2007a). Stages of Memory–Encoding, Storage and Retrieval Simply Psychology. Retrieved July 17 2015 from, http://www.simplypsychology.org/memory.html

McLeod, S. (2007b). Levels of Processing. Retrieved 17 July 2015 from, http://www.simplypsychology.org/levelsofprocessing.html

Poole, G., & Chick, N. (2015). Weaving SoTL into our everyday lives. *Teaching and Learning Inquiry—The ISSOTL Journal, 3*(1), 1.

Rolka, C. & Ramshagen, A. (2015) Showing up is half the battle: Assessing different contextualized learning tools. *International Journal Scholarship of Teaching and Learning 9* (1), Article 10. Retrieved 8 Sept 2015 from, http://digitalcommons.georgiasouthern.edu/ij-sotl/vol9/iss1/10

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences, 12*(11), 411–417.

Stigler, J.W. (2012). *Struggle for smarts? How Eastern and Western cultures tackle learning*. Interview Transcript hosted by Renee Montagne, 12 Nov 2012.

Thompson, F., & Logue, S. (2006). An exploration of common student misconceptions in science. *International Education Journal, 7*(4), 553.

Assessment in Mechanical Engineering Capstone Project: The Case of Taylor's University

Douglas Tong Kum Tien and Mohd Hardie Hidayat Mohyi

Abstract The Mechanical Engineering Group Project is a capstone project offered in the third year of the B.Eng (Hon) Mechanical Engineering programme at Taylor's University. It consists of two modules, namely Mechanical Engineering Group Project 1 and Mechanical Engineering Group Project 2, offered in semesters 5 and 6, respectively. These modules are crafted based on the CDIO educational framework. The CDIO framework is founded on the premise that engineering graduates should be able to Conceive, Design, Implement and Operate (CDIO) complex value-added engineering systems in a modern team-based engineering environment to create systems and products. In this capstone project, students worked in groups where they were given a task to conceive and design a human-powered vehicle (HPV) in MEGP1. Following this, in MEGP2 the students were tasked to implement and operate their design, so that a safely constructed and fully functioning artefact is produced. MEGP2 culminates in a HPV race among the project groups with the race also forming part of an assessment. The students worked in project teams consisting of 4–7 students per team. Each team created a unique HPV design dissimilar from the other groups. Throughout the two semesters, the students had to CDIO their HPV within the time frame and the budget allocated. This paper specifically describes the assessment strategy adopted for MEGP2 and provides key observations on the learning attainments made possible through it in achieving the module's and the CDIO's learning outcomes.

Keywords Capstone project · Mechanical engineering · Assessment · CDIO

D.T.K. Tien (✉) · M.H.H. Mohyi
Taylor's University, Subang Jaya, Malaysia
e-mail: DouglasKumTien.Tong@taylors.edu.my

M.H.H. Mohyi
e-mail: Hardie.Hidayat@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_24

1 Introduction

The Mechanical Engineering Group Project is the capstone design and build module offered to third-year mechanical engineering students in Taylor's University's School of Engineering's B.Eng (Hon) Mechanical Engineering programme. This module is an opportunity for the students to apply their engineering knowledge and to further develop their engineering skills in a realistic environment. It builds upon and further reinforces the CDIO concepts that the students have been introduced to when they undertake group projects in year one and year two.

CDIO is the acronym for Conceive–Design–Implement–Operate. It refers to the CDIO initiative co-founded by Massachusetts Institute of Technology and three Swedish universities with the aim of educating engineering students "who are able to Conceive–Design–Implement–Operate complex, value-added engineering products, processes and systems in a modern, team-based environment" (Crawley et al. 2007).

The CDIO approach recommends having a capstone course that seeks to provide an authentic engineering environment for learning and its practitioners observed that such an approach surpassed other experiences in teaching important lessons that can only be learned in this kind of setting (Miller and Bodeur 2002). Taylor's University's School of Engineering is a member school of the worldwide CDIO initiative and as such is fully committed to implement the CDIO approach in its curriculum in order to develop its students to become proficient engineers who are required to meet the complex challenges of the twenty-first century.

This capstone module is divided into two modules that span over two semesters. In this module, the students are allowed to form their own project teams, select their own team leader and work together in the same teams over both semesters 5 and 6. The first module, which is the Mechanical Engineering Group Project 1 (MEGP1), is offered in semester 5. This module emphasizes the Conceive and Design elements of CDIO in which the student teams are required to create, design, analyse and evaluate a human-powered vehicle (HPV). This challenge requires them to use relevant engineering design and analysis tools and techniques.

For the Mechanical Engineering Group Project 2 (MEGP2) module which is offered in the subsequent semester, the emphasis is on the Implement and Operate elements. At this stage, the student teams proceed to manufacture and test the HPV that they have designed in the previous semester. This is where hands-on application is emphasized. The module finally culminates in a HPV race among the teams.

Engineering skills such as planning, project management, procurement, fabrication, testing, problem solving, team work and communication are very much the theme for the Implement and Operate stage. The Implement and Operate stages can be considered as equally important as the Conceive and Design stages. Many good plans and designs fail due to poor implementation of the four components in CDIO practices (Al-Atabi 2014).

In order to facilitate learning, having the right assessment strategy is crucial (Lawson et al. 2014). This particular case study focusses on the assessment strategy for MEGP2 which is aimed at developing Implement and Operate capabilities in the students while still reinforcing the Conceive and Design elements already introduced to them through MEGP1.

2 Mechanical Engineering Group Project 2

MEGP2 is conducted through weekly discussion meetings with the supervisor who is also the module coordinator. Each project team is assigned a weekly time slot for meeting. During these meetings, the teams update the supervisor, who is the module leader, of their progress which they track using a Gantt chart prepared by themselves. The supervisor notes any concerns and provides feedback or affirmations as appropriate. The teams are also required to have weekly meetings of their own where the discussions of these meetings are formally minuted. These minutes would then be shown to the supervisor during the scheduled weekly discussions. With this system in place, the supervisor is able to closely monitor the ongoing progress of each team through their Gantt chart, meeting minutes and verbal discussion during these weekly discussion meetings.

In addition, each student updates a team analysis form on a weekly basis. Included in the form is a personal assessment by the students concerning their team's stage of group development based on the Tuckman model (Tuckman 1965). According to Tuckman, there are 4 stages of group development, namely forming, storming, norming and performing stages. By having each team member rate, their own team's group development stage enables the supervisor to discern the progress and needs of the team.

The weekly discussion meetings serve to keep the students' progress on track. It makes the teams accountable to the supervisor on a regular basis. Building a well-constructed and fully functioning HPV within the given time frame and budget requires considerable initiative, self-discipline, strong team work and high level of commitment from the students. This challenging activity has to be carried out in the midst of tests and assignments due from other modules, which could potentially derail the project's progress and direct the focus away from it. Hence, the weekly meetings where attendance is compulsory help to maintain the students' ongoing focus on the project.

Finally, the supervisor conducts several lectures during the first half of the semester to ensure that the students fully understand the requirements of the module, the assessments, the project deliverables, timelines and specific topics that would be helpful for the students during this stage. A module handbook is prepared for the students, and it is given at the start of the semester. The handbook covers all relevant information pertaining to the module with an entire set of assessment rubrics given in its appendix.

By semester 6, most of the students should already be familiar with some of the workshop practices through their participations in earlier projects, workshops and laboratory experiences. Otherwise, they are expected to learn on their own whatever is required to fabricate their HPV. There are technical staffs overseeing these workshops to provide guidance to them.

The module is conducted in such a manner as to ensure that the expected learning outcomes for MEGP2 as stated below are achieved.

1. Create a functioning prototype based on design.
2. Evaluate the functionality of prototype against design.
3. Evaluate the design based on performance, cost and sustainability and to optimize the design if necessary.
4. Execute project closure.

3 Assessments in Mechanical Engineering Group Project 2

The assessment components for the MEGP2 are listed in Table 1. These will be discussed in the sections to follow.

3.1 Logbook

The importance of the logbook in the students' learning cannot be underestimated. In MEGP2, each student records not only their tasks but also their learning over the previous fortnight. They then meet with the supervisor to show their logbook updates and to discuss on their progress. These fortnightly updates are meant to be compiled as one of the assessment submissions at end of the semester. While the general weekly discussion involves the entire team, the logbook discussion involves each student individually.

For the fortnightly logbook update, the students are required to record more than just a list of activities that they have carried out over the past two weeks. They are

Table 1 Assessment components of MEGP2

Type of assessment	Learning domain	Mark (%)
Logbook	Cognitive	10
Interim report	Cognitive	30
Final report	Cognitive	30
Presentation and peer assessment	Cognitive and affective	10
Artefact assessment	Cognitive, affective and psychomotor	20

also expected to reflect on these activities and garner any learning experiences from it. These learning reflections are to be documented in the logbook. Attainment of new or the enhancement of existing skills would also be duly recorded. In addition, the students also need to make an effort to link these learnings to the relevant learning outcomes and/or the relevant programme outcomes. At the end of the logbook meeting, the supervisor signs the student's current logbook update. The student compiles all the signed copies which would be submitted for a final assessment at the end of semester.

Besides the logbook update, the students are required to prepare a return-on-failure (ROF) analysis form twice during the semester on the sixth week and the twelfth week. The students identify mistakes or failures that they have made during the course of this project. It must be something that has already happened. They must identify the root cause, prescribe corrective actions and recognize learning opportunities. The purpose of this activity is to alleviate the fear of failure which is an inhibition to innovation and future success.

The message to the students is that any unforeseen failure can lead to positive returns if learning takes place. Initially, there were students who mistook return-on-failure with predicting potential design failures in their HPV. While being able to foresee and predict possible failures so as to mitigate and prevent it is strongly encouraged in an engineer, this is not considered as return-on-failure (ROF) analysis. ROF concerns the failures that have already occurred. The failures could be in the area of design, manufacture, project management, communication, documentation and leadership, basically anything that pertains to their project work for MEGP2.

As with the logbook update, the contents of the ROF form are likewise to be discussed individually with the supervisor and the supervisor signs the ROF form as acknowledgement. Together with the logbook updates, the ROF forms are compiled and submitted at the end of semester. A written summary and analysis accompany these submissions which are then assessed by the supervisor according to the rubric given in the handbook.

3.2 Interim Report

The interim report is basically a midpoint evaluation to ensure that the project is progressing in the right direction and according to schedule. The interim report is due for submission on week 5 of the semester. Midpoint evaluations are meant to improve the awareness of team performance and lead to enhanced self-correction and motivation (Zhou and Pazos 2014). It is also an opportunity for the supervisor to provide written feedback by assessing the interim report.

An analysis of the design and the prototype which was completed during MEGP1 is required to be written in the interim report. At this stage, the students would have progressed beyond their previous level of understanding and therefore would be able to analyse their design better. They are also expected to be able to

identify weaknesses and/or possible improvements compared to the analysis made at MEGP1. This report also serves as a progress update and it is assessed according to its rubric which is available in the student handbook.

3.3 Final Report

The final report is usually due in week 12 of the semester, and it is meant to capture all activities related to the project that a student has undertaken in MEGP2. It also requires the students to analyse these activities in depth. The analysis would include the implication of the design on manufacturability, testability, usability, ease of maintenance and sustainability. It should also analyse design decisions, manufacturing process options, procurement issues, planning and scheduling judgments, etc.

The final report is an individual submission since each student plays a different role in the team. In the final report, each student is required to focus primarily on his/her contribution to the team effort. Although it is an individual submission, an overall description and analysis of the entire product are also expected. Students are also expected to discuss possible areas of improvement. They need to develop at least 5 feasible and appropriate solutions for improvement and possible ways to implement them.

Together with the final report submission, the entire team also submits a Technical Operations and Maintenance Manual (TOMM) for their HPV which they are expected to produce. The TOMM is meant to assist users in obtaining a general understanding of the engineering system of the HPV and to provide solutions that may arise from the operation of the HPV. Safety of the user is a paramount concern, and its emphasis must be clearly highlighted in the TOMM.

3.4 Presentation and Peer Assessment

This assessment consists of two parts, a group presentation and a peer assessment, with each part contributing 5 %. These assessments help in the development of the affective learning domain.

3.4.1 Group Presentation

This presentation is conducted and assessed at the end of the semester Engineering Fair. The assessor is the supervisor. All team members are required to take turns presenting their artefact in front of them with the help of PowerPoint slides. During the Engineering Fair, each team needs to display their artefact at an assigned location for an entire day. During the Engineering Fair, judges, academic staff, visitor and other students may visit their HPV booth and pose any questions.

Besides the fluency and clarity of presentation, students are expected to highlight all project deliverables as well as recommendations for improvement.

3.4.2 Peer Assessment

The students are required to assess their peers, i.e. their team members by giving a rating of 0–10. They are assessed for their contributions, problem solving, their attitude, focus on task and working with other team mates. This assessment is done through a survey conducted anonymously using Google Drive. The peer assessment contributes 5 % to the total marks of the module.

3.4.3 Artefact Assessment

The artefact assessment also takes place during the Engineering Fair. Two judges other than the supervisor will conduct this assessment. One of the judges would be from another university or from any engineering industry invited for the Fair, while the other judge would be from Taylor's School of Engineering. The artefact is assessed for its innovation, safety, design and aesthetics. Having an external judge is beneficial to serve as a benchmark against other universities, to provide students with the opportunity to showcase their work to other experts and to receive their feedback and serve as a learning opportunity.

3.4.4 HPV Race

The HPV race is a part of artefact assessment but is conducted on a different day other than the Engineering Fair. It is the concluding highlight of the module. Through this race, students' products are pushed to the limit. They will quickly discover what works and what does not work. It is impossible to hide any deficiencies which might have escaped judges' notice during the artefact assessment in the Engineering Fair. This is as real world as it gets. A product must work as intended under the actual working conditions for which it was designed. Not only it is meant to function when tested under ideal conditions, but it must survive the rough and tumble of end users' handling.

Although this activity is classified as an "operate" activity in CDIO, there is no less of design and implementation learning that can take place. If a product fails to operate as intended, there are usually unforeseen design deficiencies, implementation mistakes or judgmental errors in decision-making usually pertaining to its design and implementation. Even if the product survives the race and does well, the students still have the opportunity to discover how further improvements can be made whether to go faster, to provide better handling, to improve safety and the like. Herein can be found tremendous opportunity for learning as the students will not likely forget the race experience.

The race was conducted in the vacated staff car park area which was not a proper race track. For safety reasons, it was decided that the race had to be a timed race. Personal protective equipment is provided for safety precautions. Each team would perform three laps around the circuit and the places were awarded according to the time taken to complete the three laps. Every team was given two attempts with the better attempt's time used for comparison against the rival teams. Figures 1 and 2 show the teams preparing and one of the teams racing, respectively.

Fig. 1 Students preparing for the HPV race

Fig. 2 One of the student HPVs in action

To make this a race proper, marks were awarded to the students within the artefact assessment rubric, according to their teams' placement. Although the race only constituted a small fraction of the artefact assessment marks, it introduced a healthy spirit of competition which facilitated learning.

There were instances where students' HPVs failed to complete the race. However, as a general observation in capstone projects, it has been noted that failure of the end product does not mean that learning has not occurred (Mosher and Ramaswamy 2014). The main objective of this module was to help students learn, both through their successes and their "failures", and it is believed that this objective was accomplished.

There is an ongoing debate about process versus product in capstone projects. Should students deliver a successful product or should the focus be more on the learning process itself (Paretti et al. 2011)? For this capstone module, both of these objectives were aimed for and both were observed to be largely attained. The effectiveness of learning process has been explained in this paper while the artefacts from the majority of the students' teams managed to return a decent performance in the HPV race. Therefore, they successfully completed the CDIO cycle with regard to both the learning process and the product delivered.

4 Conclusion

The assessment process of a mechanical engineering capstone module conducted at Taylor's University and how it impacts on students' learning experience has been presented. The intention of this capstone module is primarily to help the students become conversant with the CDIO approach. Consequently, the assessment process reflects this objective. It has been demonstrated through this case that it is possible to achieve the goal with the current assessment strategy. Nonetheless, in the spirit of continual improvement, the assessment process and rubrics will continue to undergo review from semester to the semester.

References

Al-Atabi, M. (2014). *Think like an engineer*. Subang Jaya, Malaysia: Mushtak Al-Atabi.

Crawley, E. F., Malmqvist, J., Ostlund, S., & Brodeur, D. R. (2007). *Rethinking engineering education: The CDIO approach*. New York: Springer.

Lawson, J., Rasul, M., Howard, P., & Martin, F. (2014). *Getting it right: assessment tasks and marking for capstone project courses*. Capstone Design Conference: 2014 Conference Proceedings. Retrieved 31 July 2015, from http://www.capstoneconf.org/resources/2014%20Proceedings/Papers/0017.pdf

Miller, D. W., & Bodeur, D. R. (2002). *The CDIO capstone course: An innovation in undergraduate systems education. Proceedings of the 2002 American Society for Engineering Education Annual Conference.* Retrieved 1 Aug 2015 from http://ssl.mit.edu/spheres/library/CDIO_Capstone_030402.pdf

Mosher, G. A., & Ramaswamy, S. K. (2014). *Role of senior capstone courses in 21st century technology undergraduate programs*. White Paper by the Association of Technology, Management, and Applied Engineering. Retrieved 30 July 2015, from http://c.ymcdn.com/sites/www.atmae.org/resource/resmgr/White_Paper/JUNE_2014_White_Paper%5B2%5D.pdf

Paretti, M., Layton, R., Laguette, S., & Speegle, G. (2011). Managing and mentoring capstone design teams: Consideration and practices for faculty. *International Journal of Engineering Education, 27*(6), 1–14.

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin, 63*(6), 384–399.

Zhou, Z., & Pazos, P. (2014). *Managing engineering capstone design teams: A review of critical issues and success factors. Proceedings of the 2014 Industrial and Systems Engineering Research Conference*. Retrieved 30 July 2015, from http://xcdsystem.com/iie2014/abstract/finalpapers/I1107.pdf

Development and Assessment of Learning Outcomes of Graduate Course on Structural Rehabilitation

Tamer El-Maaddawy

Abstract Learning outcomes of a civil engineering graduate-level course on rehabilitation of structures are presented and assessed in this paper. The course is offered to Master's students by the Civil and Environmental Engineering (CEE) Department at the United Arab Emirates University (UAEU). It comprises two modules. Module I is about condition assessment and conventional repairs whereas Module II is about innovative strengthening with composites. Students are exposed to real-life problems and hands-on training. They conduct research, analysis, and comparisons. They present and critique recent research articles, report findings and submit technical reports. Course learning outcomes (CLOs) have been developed and mapped to the program learning outcomes. Direct assessment instruments have been developed to assess the level of achievement of the CLOs including homework assignments, exams, laboratory project, research paper, and presentations. Student performance has been compared to faculty and student self-perceptions. All CLOs have been adequately achieved. The level of achievement of all CLOs, based on student performance, was on average 85 % with a standard deviation of 2 % and coefficient of variation of 2.5 %. The CLO pertaining to student communication skills featured the lowest level of achievement of 82 %. The faculty and student self-perceptions were in agreement. Nevertheless, the faculty self-perception was closer to the student performance than the student self-perception. In future terms, students should deliver multiple presentations throughout the semester in order to improve their communication skills, and hence, advance the program competencies and student learning.

Keywords Assessment · Learning outcomes · Civil engineering · Graduate level

T. El-Maaddawy (✉)
United Arab Emirates University, Al Ain, United Arab Emirates
e-mail: tamer.maaddawy@uaeu.ac.ae

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_25

1 Introduction

In 2007, the department of Civil and Environmental Engineering (CEE) at the United Arab Emirates University (UAEU) has established a Master of Science (MSc) graduate program in Civil Engineering (CE), leading to M.Sc. degree in CE. The program is designed to provide the community and industry with distinguished national manpower and highly qualified civil engineers for the sustainable development of the country. The program also trains students intending to purse their doctor of philosophy (PhD) studies in CE. Graduates of the program are anticipated to provide the link between the advancements in civil engineering sciences and corresponding applications. The program covers various CE disciplines including structural engineering, geotechnical engineering, construction management, highway and transportation, water resources, environmental engineering, and surveying. The program degree requirements are given in Table 1. The program has two modes of enrollment, namely "thesis" and "non-thesis." The "thesis" option requires successful completion of at least 30 credit hours, among which 6 credit hours of Master's research thesis that should be completed over a minimum of two semesters. The "non-thesis" option requires successful completion of at least 33 credit hours.

The National Qualifications Authority (NQA) of the United Arab Emirates has established a National Qualifications Framework (QF*Emirates* 2012) that provides a frame of references concerning the quality of qualifications awarded for learning. The QF*Emirates* define qualifications based on "learning outcomes" rather than content or time spent on a program. It focuses on the outcome of the learning experience and the achievement of a set of learning outcomes that reflect what is expected to be achieved at a specific level of qualification, from the most simple to the most advanced levels of learnings. The QF*Emirates* describes the learning outcomes in terms of knowledge, skills, and aspects of competence. The aspects of competence are expressed in terms of autonomy and responsibility, role in context, and self-development.

The graduate studies committee of the CEE department at the UAEU, chaired by the author, has developed a set of program learning outcomes (PLOs) at the MSc level that are in compliance with the QF*Emirates*. Learning outcomes of the MSc program in CE adopted by the CEE department at UAEU are given in Table 2. The outcome-based education concept has been implemented in several engineering programs worldwide (Felder and Brent 2003; Wahab et al. 2011; Arshada et al.

Table 1 Degree requirements of the MSc in CE

Courses	Thesis option (credit hours)	Non-thesis option (credit hours)
Core	12	12
Electives	12	21
Research thesis	6	0
Total	30	33

Table 2 Program learning outcomes (PLOs) of the MSc in CE

Outcome	Description
PLO1	Describe highly specialized civil engineering principles, concepts, and methodologies
PLO2	Evaluate the performance of advanced civil engineering systems and components through the use of applicable research principles, analytical methods and modeling techniques
POL3	Conduct advanced applied research to develop innovative solutions for highly complex civil engineering problems through the use of appropriately selected research methodologies and modern engineering tools
PLO4	Apply advanced multidisciplinary problem-solving approaches to critically analyze contemporary, sophisticated, and highly complex civil engineering problems.
PLO5	Present and critique highly complex civil engineering issues and communicate effectively at a high level of proficiency
PLO6	Lead professional activities and manage ethical issues in highly complex civil engineering projects
PLO7	Implement the social, environmental, ethical, economic, and commercial aspects to develop valid decisions affecting highly complex civil engineering projects

2012). Assessment of PLOs is essential for obtaining accreditation by various accrediting organizations (QF*Emirates* 2012; ABET 2013; Osman et al. 2012; El-Maaddawy 2015a). Results of assessment of PLOs will help the graduate studies committee in making decisions that can greatly improve the program competencies and student learning (Afida et al. 2011; Lindholm 2009).

This paper presents results of assessment of learning outcomes of the civil engineering graduate-level course Rehabilitation of Structures—CIVL 616, developed recently by the author (El-Maaddawy 2015b). The study was conducted on a batch of 12 Master's students enrolled in the course offered in the second semester of the academic year 2014–2015. Results of the present study will be used as input data to evaluate the overall achievement of the PLOs and to make decisions pertaining to curricular and program development.

2 Course Description

The course Rehabilitation of Structures—CIVL 616 is an elective graduate-level course offered for Master's students by the CEE Department at the UAEU. The course is three credit hours. It offers practical tips, design details, and case studies collected by the author upon comprehensive research and outreach activities. The course consists of two modules. Module I is about condition assessment and conventional repairs whereas Module II is about innovative strengthening composites. Each module comprises four topics as given hereafter.

- Module 1—Condition assessment and conventional repairs
  - Defects and deterioration,
  - Structural condition assessment,
  - Repair procedure and materials, and
  - Structural strengthening and stabilization.

- Module II—Innovative strengthening with composites
  - Principles of strengthening with composites,
  - Flexural strengthening with composites,
  - Shear strengthening with composites, and
  - Column strengthening with composites.

3 Development of Course Learning Outcomes

The CLOs describe what a learner will know, understand, and be able to do upon successful completion of the course. Each CLO should start with an action verb. The CLOs should be observable, measurable, and capable of being understood by the learners, faculty, external agencies, and stakeholders. They should also contribute to the achievement of one or more of PLOs. The learning outcomes of the graduate course CIVL 616 are given in Table 3 and mapped to the PLOs in Table 4. The assessment units used to measure the level of achievement of each CLO is also given in Table 4.

Table 3 Learning outcomes of the course CIVL 616

Outcome	Description
CLO1	Describe damage mechanisms and principles of structural strengthening
CLO2	Identify causes of defects, cracks, damage and deterioration of concrete structures
COL3	Develop an appropriate repair strategy for a deficient structure taking into consideration the social, economic and commercial aspects
CLO4	Perform analysis and design of reinforced concrete elements strengthened with advanced composites using appropriate problem-solving approaches and international code provisions
CLO5	Conduct experiments for condition assessment, corrosion monitoring, and strengthening of columns using advanced techniques
CLO6	Report findings and critique recent research on assessment and rehabilitation of structures
CLO7	Communicate effectively with peers and clients at a high level of proficiency

Table 4 Mapping of CLOs to PLOs

CLOs	Program learning outcomes (PLOs)							Assessment units[a]
	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	
CLO1	x							HW1, HW2, EM1, EM2
CLO2	x							HW1, EM1
COL3	x	x		x			x	HW1, EM1
CLO4	x	x		x				HW2, EM2, LR
CLO5		x	x					LR, LP
CLO6			x		x	x		LR, LP, RR, RP
CLO7					x	x		LP, RP

[a]*HW* homework, *EM* Exam, *LR* Laboratory report, *LP* Laboratory presentation, *RR* Research paper report, *RP* Research paper presentation

4 Assessment Instruments

Direct and indirect assessment instruments have been developed to evaluate the level of achievement of the CLOs. The direct assessment instruments used to assess student learning throughout the course included homework assignments (groups assessment), exams (individual assessment), laboratory project (group assessment), research paper (group assessment), and presentations (individual assessment). The indirect assessment instruments included student assessment of course survey where a questionnaire survey was distributed on all students to fill out at the end of the semester. In the questionnaire, students were asked to rate their level of achievement of CLOs from *very low* to *very high* on a 5-point scale. The course instructor (faculty) rated also the level of achievement of CLOs from *very low* to *very high* on a 5-point scale based on direct observation of student performance in class activities.

4.1 Laboratory Project

The laboratory project has been designed to contribute to the achievement of course learning outcomes CLO4, CLO5, CLO6, and CLO7. The laboratory project aims to expose students to the following activities:

- Hands-on training on nondestructive tests and electrochemical techniques for structural condition assessment and corrosion monitoring.
- Hands-on training on strengthening of concrete columns with carbon-fiber-reinforced polymers (CFRP).
- Application of analytical approaches for prediction of load capacity of concrete columns wrapped with CFRP.
- Comparative analysis with predictions of international guidelines and standards.
- Data analysis, reporting, and effective communications with peers and clients.

Corrosion monitoring test | Bulk resistivity test | UPV test

Fig. 1 Tests in progress conducted by students

Students have been divided into groups, three students each. The project has been divided into two phases. In Phase I, students performed corrosion monitoring of steel reinforcing bars in concrete prisms subjected to wet–dry cycles. In Phase II, students conducted nondestructive testing, strengthening, and destructive testing on concrete columns. Tests of Phase I included covermeter survey, half-cell potential, corrosion rate with linear polarization technique, and concrete resistance using 4 probe Wenner resistivity meter. Students conducted also Schmidt hammer test, ultrasonic pulse velocity (UPV) test, and bulk concrete resistivity test using Giatec RCON™ electrical resistivity meter on specimens of Phase II prior to wrapping with CFRP. Sample tests in progress conducted by the students are shown in Fig. 1.

Strengthening of concrete columns with CFRP included surface preparation, cutting of carbon fiber fabric sheets, mixing of epoxy adhesive, application of epoxy and wrapping of impregnated carbon fiber fabrics. The concrete columns were tested to failure by the students under uniaxial compression. The measurements recorded during testing by students included loads, longitudinal strains, and lateral strains in the CFRP-wrapped columns. Students performed a comparative analysis between the experimental results and analytical predictions of international guidelines and standards (ACI 440.2R-08 2008; Intelligent Sensing for Innovative Structures Canada Educational Module 2004).

4.2 Research Paper

The research paper assignment has been designed to contribute to the achievement of course learning outcomes CLO6 and CLO7. In the research paper assignment, students reviewed and discussed original and recent journal articles, describing a major scientific advancement in a research area related to condition assessment and rehabilitation of existing structures. Students made presentations, submitted reports and participated in discussions. Students were divided into four groups, three students each. The research topics selected by student groups are given below.

- Service life predictions of deteriorated concrete structures,
- Monitoring of corrosion activity and measurements of corrosion rates,

- Structural health monitoring (SHM) and damage detection, and
- Seismic retrofitting and shake table tests of concrete structures.

5 Assessment Methodology

The direct assessment units used to measure the level of achievement of each CLO are given in Table 4. Approximate weight for each assessment unit is given below:

- Exam (Module 1) = 25 %,
- Exam (Module 2) = 25 %,
- Homework 1 = 10 %,
- Homework 2 = 10 %,
- Laboratory project report = 10 %,
- Research paper report = 10 %, and
- Presentations = 10 %.

The student performance indicated by the average earned grades in the direct assessment measures pertaining to each CLO provides direct evidence of student learning. Accordingly, student grading linked to each course learning outcome has been used as a direct measure to assess the level of achievement of CLOs. The student average performance has been presented on a 5-point scale to be compared with results of the student and faculty self-perceptions. The assessment procedure has been implemented into an automated spreadsheet capable of developing assessment reports and graphs.

6 Assessment Results

Results of assessment of CLOs are summarized in Table 5 and presented in Fig. 2. A specific CLO was considered "achieved" when the corresponding student performance weight value was in the range of 3.5–4 (i.e., 70–80 %). For student performance weight values greater than 4 (i.e., >80 %), the CLO was considered "adequately achieved." From Table 5, it can be seen that all CLOs have been adequately achieved. The level of achievement of all CLOs, based on student performance, was on average 85 % with a standard deviation of 2 % and coefficient of variation of 2.5 %. The lowest level of achievement of 81.7 % was featured by CLO7 pertaining to student communication skills. On the other hand, CLO2 and CLO3 related to student knowledge on causes of defects and student skills on development of repair strategies, respectively, featured the highest level of achievement of 87.4 %.

The consistency between student performance, faculty, and student self-perceptions was estimated by calculating the corresponding alignment index I_i as shown in Table 6. The alignment index values were calculated by dividing

Table 5 Results of assessment of CLOs

CLOs	Student self-perception	Faculty self-perception	Student performance		Level of achievement of CLOs based on student performance
			Percent	Weight	
CLO1	4.27	4.25	86.6	4.33	adequately achieved
CLO2	4.64	4.25	87.4	4.37	adequately achieved
COL3	4.36	4.25	87.4	4.37	adequately achieved
CLO4	4.18	4.25	85.9	4.30	adequately achieved
CLO5	4.36	4.25	83.5	4.18	adequately achieved
CLO6	4.36	4.25	85.2	4.26	adequately achieved
CLO7	4.36	4.25	81.7	4.08	adequately achieved

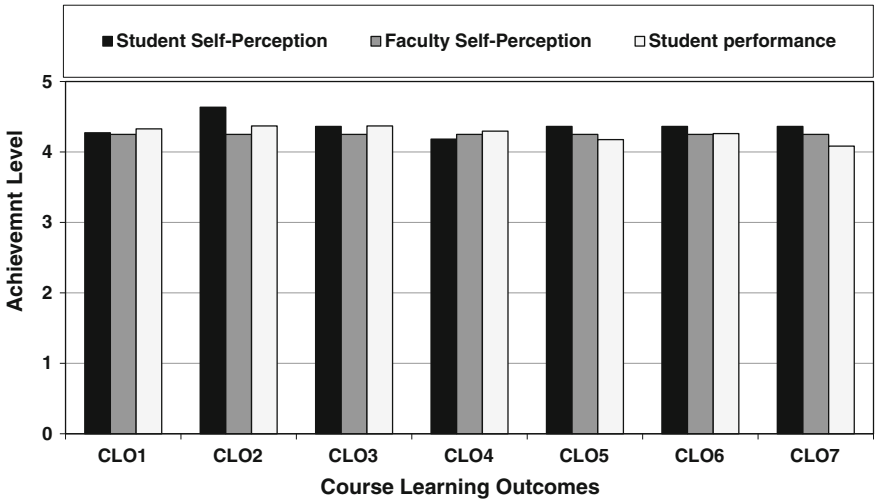


Fig. 2 Results of assessment of CLOs

results of faculty or student self-perceptions by those based on student performance. Index values in the range of $0.85 \leq I_i \leq 1.15$ were considered indicative of consistency/alignment between results. From Table 6, it can be seen that the faculty and student self-perceptions were generally consistent with student performance. Nevertheless, students tended to overestimate the level of achievement of CLOs. This was more evident for CLO7 with a student alignment index value of $I_s = 1.07$. Results of faculty and student self-perceptions were considered in agreement when the difference between them was within a 15 % error band (i.e., $\Delta I \leq 0.15$). Although the faculty and student self-perceptions were in agreement, results of the faculty self-perception were closer to student performance. In future terms, students should deliver multiple presentations throughout the semester to improve their communication skills and increase the level of achievement of CLO7.

Table 6 Outcome-based alignment indices

CLOs	Alignment index based on student self-perception		Alignment index based on faculty self-perception		Faculty versus student self-perception	
	Index (I_s)	Observation	Index (I_f)	Observation	$\Delta I = \lvert I_f - I_s \rvert$	Observation
CLO1	0.99	Aligned	0.98	Aligned	0.01	Agreement
CLO2	1.06	Aligned	0.97	Aligned	0.09	Agreement
COL3	1.00	Aligned	0.97	Aligned	0.03	Agreement
CLO4	0.97	Aligned	0.99	Aligned	0.02	Agreement
CLO5	1.04	Aligned	1.02	Aligned	0.02	Agreement
CLO6	1.02	Aligned	1.00	Aligned	0.02	Agreement
CLO7	1.07	Aligned	1.04	Aligned	0.03	Agreement

7 Conclusions

Course learning outcomes (CLOs) of the civil engineering graduate-level course Rehabilitation of Structures—CIVL 616 have been assessed in this paper through direct and indirect assessment instruments. All CLOs have been adequately achieved. The level of achievement of all CLOs, based on student performance, was on average 85 % with a standard deviation of 2 % and coefficient of variation of 2.5 %. The lowest level of achievement of 81.7 % was featured by CLO7 pertaining to student communication skills. Course learning outcomes CLO2 and CLO3 related to student knowledge and skills, respectively, featured the highest level of achievement of 87.4 %. The faculty and student self-perceptions were generally consistent with student performance. Students tended, however, to overestimate the level of achievement of CLOs. The faculty and student self-perceptions were, generally, in good agreement. Nevertheless, results of the faculty self-perception were closer to student performance than those pertaining to student self-perception. In future terms, students should deliver multiple presentations throughout the semester to improve their communication skills and increase the level of achievement of CLO7. Results of the present study will be used as input data to evaluate the overall achievement of PLOs which will help the program executive committee in making valid decisions pertaining to program development.

References

ABET. (2013). *Criteria for Accrediting Engineering Program*. Engineering Accreditation Commission of the Accreditation Board for Engineering and Technology (ABET), p. 26.

ACI 440.2R-08. (2008). *Guide for the design and construction of externally bonded FRP systems for strengthening concrete structures*. MI: Farmington Hills.

Afida, A., Hamimi Fadziati, A. W., Norhana, A., Ahmad Ashrif, A. B., & Hafizah, H. (2011). Assessment of student program outcomes through a comprehensive exit strategy. *Procedia-Social and Behavioral Sciences, 18*, 33–38.

Arshada, I., Razalia, S., & Mohamed, Z. (2012). Programme outcomes assessment for civil & structural engineering courses at University Kebangsaan Malaysia. *Procedia-Social and Behavioral Sciences, 60*, 98–102.

El-Maaddawy. (2015a). Assessment of learning outcomes of reinforced concrete design courses: UAEU experience. *7th International Forum on Engineering Education (IFEE2015). Quality Assurance in Engineering Education*. 17–19 March, Sharjah, UAE.

El-Maaddawy. (2015b). Development of a multi-disciplinary graduate course on rehabilitation of structures. *4th International Conference on Concrete Repair, Rehabilitation and Retrofitting (ICCRRR 2015). Education, Research, and Specifications Theme*. 5–7 October, Leipzig, Germany.

Felder, R. M., & Brent, R. (2003). Designing and teaching courses to satisfy the ABET engineering criteria. *Journal of Engineering Education, 92*(1), 7–25.

Intelligent Sensing for Innovative Structures Canada Educational Module No. 4. (2004). *An introduction to FRP-strengthening of concrete structures. ISIS Canada*. Winnipeg, Manitoba, Canada: Canadian Network of Centres of Excellence on Intelligent Sensing for Innovative Structures.

Lindholm, J. A. (2009). *Guidelines for developing and assessing student learning outcomes for undergraduate majors* (1st ed.).

Osman, S., Jaafar, O., Badaruzzaman, W., & Rahmat, R. (2012). The course outcomes (COs) evaluation for civil engineering design II course. *Procedia-Social and Behavioral Sciences, 60*, 103–111.

QF*Emirates*. (2012). *Qualifications framework emirates handbook*. Abu Dhabi, UAE: National Qualifications Authority (NQA).

Wahab, H., Ayob, A., Zaki, W., Hussain, H., Hussain, A., & Mokri, S. (2011). Program outcomes measurement and assessment processes. *Procedia-Social and Behavioral Sciences, 18*, 49–55.

Standard-Based Performance Assessment's Scoring Practice in Primary Schools

Rubiah Dalail, Yuen Fook Chan and Gurnam Kaur Sidhu

Abstract There is a global shift from teaching of learning to teaching for learning in school. The Ministry of Education (MOE), Malaysia, is also emulating the global shift by implementing standard-based educational system starting from 2011. The system upholds formative assessment as a tool to integrate assessment into teaching and learning. However, the scoring in formative environment that is supposed to be conducted as authentic assessment is subjected to reliability issues. Hence, this study was conducted specially to identify the reliability of teachers' scoring in formative assessment environment and to identify whether there was a significant difference between formative and summative assessment scoring in the study. A total of 11 teachers and 530 students from grade A national primary schools in Selangor have been selected for the study. The data collected from test instruments and performance reports were analysed using quantitative method. The Cohen's kappa statistical test indicated that teachers' scoring in formative environment was subjected to low reliability. The result of inter-rater scoring agreement indicated that most of the teachers overestimated their students' performance level (>80 %, $n = 423$). In terms of inter-rater agreement in teachers' scoring, the chi-square analysis showed that there was no significant difference between scoring practices with three variables, namely (1) students' grade level, (2) students' gender, and (3) teachers' option in teaching. The analysis found that there was a significant difference between scoring practices based on school location. The findings further indicated that there was a higher agreement (14 %) identified on scores given to rural students as compared to urban students (7 %). Hence, better professional development and training in quality standard-based performance assessment scoring are seen as ways to improve teachers' scoring practice in standard-based performance assessment.

R. Dalail (✉)
Examinations Syndicate, Putrajaya, Malaysia
e-mail: drubiah@ymail.com; rubiahdalail@moe.gov.my

Y.F. Chan · G.K. Sidhu
Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia
e-mail: yuenfook@salam.uitm.edu.my

G.K. Sidhu
e-mail: gurnamsidhu@salam.uitm.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_26

Keywords Standard-based performance assessment · Scoring practice · Primary school

1 Introduction

Standard-based performance assessment is a process of deciding a score by forming sound judgement based on a prescribed level of performances, on what students are expected to know, understand, and able to do. However, not much research has been done on teachers' scoring of students' performance in standard-based referencing schools, especially at the primary school level. In fact, the scoring approaches need competent teachers who are able to authentically integrate this kind of assessment in the teaching and learning processes. Indeed, authenticity in assessment is important as students in primary schools still lack the awareness, skills, and courage to use their knowledge and make advancement in thinking and learning (Gullo 2005).

2 Standard-Based Performance Assessment

The concept of standard-based performance assessment system has been adopted in Malaysian schools since 2011. The aim of the adoption is to shift from 'teaching of learning' to 'teaching for learning' in schools. This shift is emulating the engineering of assessment in other developed nations. Segers et al. (2003) envisioned the direction of the new assessment characteristics in six characteristics. The first characteristic indicated that the assessment specialists have been shifting from decontextualised to more contextualised assessment and from paper–pencil tests to more task-driven performance assessment. The second characteristic shows a tendency for using a norm-referenced interpretation to a standard-referenced interpretation. Whereas traditional assessment reported on students' achievement, the current assessment reports on four aspects, namely student profile, achievement, development, and involvement (Examinations Syndicate 2009). The third characteristic depicts the movement of assessing the low levels of cognitive ability towards high levels of cognitive ability. Memorising and being able to repeat information are no longer favourite keys to examination questions. Instead, the relation to real-life contexts and the ability to critically apply it in a way that creates value are imposed. The fourth characteristic refers to the multidimensionality of intelligence. Whereas traditional assessment measures more on cognition aspect, the current assessment covers the wholeness of a student (Examinations Syndicate 2009). It is based on progressive learning approach in which each student is looked upon as a special subject. The fifth characteristic concerns the move towards integrating assessment into the learning process. It upholds formative assessment, which blends assessment with the function to improve learning for both teachers

and students (Brookhart 2007). The sixth characteristic refers to the change in responsibilities in the assessment process. Students are involved in the attainment of assessment objectives by equipping themselves with self-assessment tools such as targeted performance standards (PSs) and exemplars. Therefore, the adoption of self-assessment and peer assessment is an example of the shift in responsibility from teacher-centred to student-centred.

The Examinations Syndicate, in ensuring the success of the standard-based performance assessment, has tailored the PS document guide as shown in Fig. 1. The conceptual framework of PS shown in Fig. 1 is in view of the revised Bloom's taxonomy on human growth ranging from fundamental to complex thinking skills. PS consists of statements or rubrics that function as indicators of how well students should know, understand, and are able to do what is required by the curriculum context. The PS is partitioned into six levels of performance known as Bands. A Band benchmarks students' performance. Therefore, students' performance in a class may vary within these six Band levels. Figure 1 shows the two-tier division in the Band. The first tier is students' performance at the developmental stage, i.e. Band 1 to Band 3, and the second tier is students' performance at the accomplishment stage, i.e. Band 4 to Band 6. This reflects the wholeness of students intensifying with their mastery of each Band of the PS in upward migration. The higher the students move up in the Bands, the bigger the elements of 'know', 'understand', and 'able to do' that the students possess. The quality of performance expected of a student rises steadily as the student progresses through the Bands. Ultimately, at Band 6, students are able to communicate their knowledge in a clear and intelligent manner as they bring in real-life examples.

The second part of Fig. 1 shows an extract of PS document tailored for Year 2 mathematics. The PS document is drafted in a four-column table. The heading of the first column is labelled Band, indicating the level of students' standing. The second column, performance standards, is a generic description of their level of growth in learning. Descriptor, the third column, relates to the content standard of learning. This describes what students know and are able to do on a topic-by-topic basis. In addition, the fourth column, evidence, is a guide to assess students. It

BAND	PERFORMANCE STANDARDS	PERFORMANCE STANDARDS FOR MATHEMATICS
6	Know, understand and able to do in an exemplar manner	Master and apply mathematical knowledge and skills in solving complex mathematical problems creatively and innovatively
5	Know, understand and able to do in an admirable manner	Master and apply mathematical knowledge and skills in solving complex mathematical problems using different applicable strategies
4	Know, understand and able to do with manner	Know and understand mathematical knowledge to do mathematical computational procedure with guidance
3	Know, understand and able to do	Know and understand basic mathematics knowledge to do basic mathematics operation
2	Know and understand	Know and understand basic mathematics knowledge
1	Know	Know basic mathematics knowledge

Accomplishment stage

Developmental stage

Band	Performance Standard	Descriptor	Evidence
1 Know	B1 Know basic mathematics knowledge	B1D1 Whole Number	B1D1E1 State the whole numbers till 1000 by way of counting, in one, in two, in five, in ten, in hundred

Fig. 1 Extracted document of Year 2 performance standard for mathematics. *Source* Adapted from Examinations Syndicate (2012)

refers to information relevant to the conclusion about the mastery of a student's performance. Teachers monitor students' performance against these predetermined standards, the descriptors, and the evidence that students need to demonstrate. A Band is awarded to students if they are able to give evidence that fulfils all of the requirements of PS in that band. Teachers use their sound professional judgement to determine either one of the two possibilities of scoring to students: mastered or not yet mastered. The score is given by comparing the concrete evidence gathered from students with the predetermined statement in the PS document. The practice of giving score in this manner is called standard-referenced assessment (SRA) (Kubiszyn and Borich 2003, p. 57). SRA describes the standing of what students know and are able to do in comparison with the academic performance standards for a particular subject area. Therefore, teachers could scaffold by using various sequential performance evidences to help all students climb the mannerism ladder in the standards.

3 Statement of Problems

Teachers' sound judgement on students' performance is pivotal in bringing about transformation in assessment practice. However, there are many combinations of reasons on why teachers are not performing as well as the standard-based assessment system hopes. One of the reasons is mainly on the issue on validity and reliability of scoring. Taras (2005) identified that the evolving of formative assessment process was not well adopted by teachers. Teachers were seen as not able to distinguish between formative assessment (FA) and summative assessment (SA) accordingly. Most of them used SA for formative purposes. They had been unconsciously neglecting FA in classroom assessment as teaching to the test dominates the practice. Taras (2005, p. 476) believed that '*Until the centrality and indeed neutrality of SA is acknowledged, the real blossoming of FA will not and cannot occur*'. For inexperienced and not well-trained teachers, the scoring of performance assessment in classroom was too subjective as it could be done in any form, authentic or real-life contexts. Previous research findings showed that teachers expressed frustration with the amount of time required to complete students' assessment portfolios and concern over scoring reliability and validity (Lawrence 2011; Ebel and Frisbie 1991).

4 Objective of the Study

The objective of the study is to identify the reliability of the teachers' scoring in formative assessment environment. To drive the objective, the following questions were posed:

(a) How do teachers do scoring using standard-based performance assessment?
(b) How reliable are teachers' scoring in a formative assessment environment?

5 Methodology

5.1 Sampling

Overall, the analyses were conducted with data from two groups comprising of students from two randomly selected grade A schools in Selangor: an urban school and a rural school. Multistage sampling technique was used to select Year 1, Year 2, and Year 3 students as respondents. Each grade level was stratified into three different subgroups of students: high performers, average performers, and low performers. Finally, students from three different classrooms were selected from each grade level to be representative sample accordingly. Table 1 illustrates the aggregation of these respondents.

A number of 11 teachers who taught mathematics to the corresponding students in Table 1 were selected as the first raters, scoring students' performance in formative environment. In contrast, a practitioner from the Examinations Syndicate was appointed as the second rater, scoring students' performance in summative environment.

5.2 Instruments

Two instruments, administered at two different points in time, were used to analyse teachers' scoring practice of students' performance.

(a) Students' Formative Assessment Report (SFAR)

Students' formative assessment report (SFAR) is a secondary data accumulated throughout the school days by 11 mathematics' teachers under the study. The data are stored in a custom-made database known as Schools Based Assessment Management System, *Sistem Pengurusan Pentaksiran Berasaskan Sekolah* (SPPBS). Each time a student mastered a learning standard, his teacher would

Table 1 Selection of respondents according to grade level

Area	Grade levels	No. of classrooms	No. of students
Urban	Year 1, Year 2, and Year 3	3	308
Rural	Year 1, Year 2, and Year 3	3	222
Total			530

recognise this by putting a tick on the assumed PS in the SPPBS. Consequently, the summary of all performance standards mastered by a student formed the student's formative assessment report (SFAR). Students' score on whole number topic from this report was then cross-checked with students' score in summative assessment as discussed in the following subpoint (b).

(b) Students' summative assessment (SSA)

The SSA was administered to get evidence of scoring on students' performance in summative test. It was a 45-min test in paper-and-pencil format. The scope of the assessment was on whole numbers. The items' specification was based on all performance standards on the topic. To be fair and to motivate students to do the assessment, they were assessed according to the level of performance attained in SFAR. For example, students, whose SFAR scoring was Band 1, were tested on items from Band 1 only. The same were applied to students rated Band 2, where they were answering items from Band 2 and so on. This meant that in the SSA setting, students were answering either one of the six booklets (six bands) that matched their SFAR scoring. It was assumed that students were able to complete the customised tests with ease as they were assessed according to the band level they attained in the SFAR scoring. The field experience helped the researcher in developing the students' test booklets. The institutional procedures and guides of items' construction were strictly followed by the researcher. For the purpose of SSA, the researcher had developed PS items according to grade level—Year 1, Year 2, and Year 3. Each grade level had six booklets, namely Band 1, Band 2, Band 3, Band 4, Band 5, and Band 6. The number of items varied between bands as it represented the number of evidences students needed to perform to acquire mastery of performance standard in that particular band. For example, there were five performance standards in Band 1 of Year 1, and therefore, students needed to answer 15 items as evidence of their mastery in those standards. The mastery of each PS was measured using five different items, as according to the rule of thumb, five items were needed to measure a single standard or skill (McMillan 2007). As we know, the greater the number of items, the greater the reliability (McMillan 2007, p. 75). Besides that, the difficulty of the items also affected the reliability. Therefore, the items presented in the tests were designed to have more moderate-level items and small sum of difficult and easy items to be valid and reliable.

To validate the instruments, an expert from the Examinations Syndicate, two academicians, and two teachers were sought to provide advice and comments on the content.

6 Findings and Discussions

The discussion in this section focuses on the comparison between formative assessment scoring by teachers as the first rater via SFAR and summative assessment scoring by a practitioner as the second rater via SSA.

RQ1: How do teachers do scoring using standard-based performance assessment?

The result of comparing agreement rates of scoring between raters, as in Table 2, shows that scoring agreement was only 20 % (n = 107) of the total cases. On the contrary to the SFAR, the practitioners' scoring signifies that 80 % (n = 423) of the students had not yet mastered some standards. Therefore, the standing of these students should, by right, be placed at a band level below the identified band level. In other words, the study's finding suggests that teachers overestimated their students' performance level most of the time.

The study's findings show that none of the teachers observed or scored any of their students' standing at Band 6, referring to the independent efforts of students to expend their own knowledge. Can this finding be related to the norm of the traditional practice? The traditional practice of teaching views students as highly dependent on their teachers' teaching and feedback. They are not encouraged to take responsibility and exercise freedom and power regarding their abilities and own creative process of adapting it to the real world. Are these skills being sufficiently emphasised in the classroom? In fact, the degree of consistency required of scoring depends on how well teachers know their students and not on their idiosyncrasies. The way teachers respond to students' ongoing learning responses and their authenticity in assessing is a key dimension of quality scoring. Furthermore, teachers must also collect comprehensive information about each student so that all areas of development and learning are addressed. There are at least three indicative findings in the study that are consistent with many scholars' views on a better scoring practice.

Table 2 Cross-tabulation of students' performance level by inter-rater scoring agreements

Students' performance level indicators	Inter-rater scoring agreement (%)		
	Agreement	Disagreement	Total
Band 1	38.46 (15)	61.54 (24)	100.00 (39)
Band 2	46.67 (49)	53.33 (56)	100.00 (105)
Band 3	7.43 (20)	92.57 (249)	100.00 (269)
Band 4	22.55 (23)	77.45 (79)	100.00 (102)
Band 5	Nil	100.00 (15)	100.00 (15)
Band 6	Nil	Nil	Nil
Total	20.19 (107)	79.81 (423)	100.00 (530)

Note Number in parentheses is the frequency of students being assessed in each category

Table 3 Teachers' scoring agreement by school location

School location	Inter-rater agreement (%)		
	Agreement	Disagreement	Total
Urban	6.60 (35)	51.51 (273)	58.11 (308)
Rural	13.58 (72)	28.30 (150)	41.89 (222)
Total	20.19 (107)	79.81 (423)	100.0 (530)
Chi-square	35.541		
P	0.000		
phi	0.26		

Note Number in parentheses is the frequency of students being assessed in each category

6.1 Smaller Number of Students in a Class

Table 3 shows that there is a significant difference between urban and rural teachers' scoring on students' performance level ($X = 35.541$, $p < 0.05$). Comparatively, this association might be due to the number of students per class in respective locations. The small number of students per class in rural areas makes a more personalised learning environment; therefore, a higher number of scoring agreement have been achieved. This assumption is supported by a longitudinal study of the effect of small class size in Tennessee, known as the Student/Teacher Achievement Ratio (STAR) project. The study concluded that students who were enrolled in small classes in the early elementary grades were more likely to be successful in school.

6.2 The Age of the Students

Even though there is no significant difference between scoring practice and students' age ($X = 4.187$, $p > 0.05$) according to chi-square analysis, the comparative difference shown in Table 4 shows that the percentage scoring agreement is highest among students at Year 3, followed by Year 2 and Year 1. This finding is consistent with Gullo's (2005, p. 20) observation that 'assessment policies should be designed recognising that reliability and validity of assessment increase with children's age. The younger the child, the more difficult it is to obtain reliable and valid results from assessment procedures. Hence, many types of assessment should not be administered until the child is older'.

Table 4 Teachers' scoring agreement by students' grade level

Students' grade level	Inter-rater agreement (%)		
	Agreement	Disagreement	Total
Year 1	5.85 (31)	27.55 (146)	33.40 (177)
Year 2	6.04 (32)	27.74 (147)	33.77 (179)
Year 3	8.30 (44)	24.53 (130)	32.83 (174)
Total	20.19 (107)	79.81 (423)	100.0 (530)
Chi-square	4.187		
P	0.123		
phi	0.09		

Note Number in parentheses is the frequency of students being assessed in each category

6.3 Grouping Students in Similar Ability Classroom

The result in the study indicates that homogeneous performance level grouping is superior for promoting students' learning outcome. Teachers in rural schools that adopt grouping students with similar ability were better able to identify their students' performance level compared to their companions at the urban schools who have heterogeneous group of students in a class. This finding suggests that grouping students with similar ability works well with standards' performance assessment. The practice of manipulating students' performance becomes easier as noted by Hargreaves (1972, p. 139) who stated that 'teacher is a position where if he wishes he can exploit the pupils at will'.

RQ2: How reliable is teachers' scoring in a formative assessment environment?

The Cohen's kappa statistical test ($p < 0.05$) proved that the teachers' scoring in the study was subjected to low reliability. The reliability can be discussed in the following two points.

6.4 The Substance of Learning

It is not surprising that there was low agreement between the teachers' and practitioner's scoring, as according to findings, teachers were not actually practising meaningful learning as proposed by the new concept of classroom assessment. The highest disagreement on teachers' scoring in this study was positioning students at Band 3 and Band 5 as shown in Table 2. Band 3 is on students' performance of doing basic mathematics operation, and Band 5 is on injecting creativity to think out of the box. Teachers either were not good in subjective cognitive level or did not expose the students to open-ended problem-solving. Students' scripts show that they failed even simple basic operation problems when these were expressed in wording. Novak (2010, p. 24), in analysing other studies, found that '*…changing*

teaching patterns and student learning patterns towards more meaningful learning was not easy, in fact the percentage of students passing State exams that year declined. Nevertheless, teachers and students felt that good things were happening with the new emphasis on meaningful learning…'.

6.5 Scoring Interpretation

Currently, in Malaysian primary schools, two coexisting assessment systems are in place: formative assessment and summative assessment. This dualistic nature of scoring provides useful data as it serves as a multidimensional assessment that assesses both academic knowledge and performance of a student. The issue that popped up is that teachers' assigned scoring of FA has low reliability and has not been systematically used for accountability reporting as it has weak association with SA. In fact, the more comprehensive and contextualised the assessment looks at students' performance by means of combining FA and SA, the more scoring increases the likelihood of valid and reliable assessment results because teachers have assessed more information during the teaching learning activities. Indeed, Cizek and Bunch (2007, p. 11) believed that it is impossible to express a judgement about an 'absolute' level of performance without incorporating normative referencing and standards' referencing.

7 Conclusion

Prompted by the global trends towards teaching for learning, the introduction of standard-based performance assessment in Malaysian schools is timely and necessary. However, for the betterment of the quality of teachers in scoring practice, teachers cannot rely solely on the ministry alone to spoon-feed what and how to do assessment. Indeed, the implementation of standard-based performance assessment is seen as pressure to teachers to wake up from their comfort zone to search for better 'scoring' by experimenting new techniques of integrating performance assessment into the teaching and learning processes. In order to get a better result of the implementation of standard-based assessment, teachers have to be willing to give the new method of assessment a fair tryout and to really want to change over to the new assessment methods.

References

Brookhart, S. M. (2007). Expanding views about formative classroom assessment: A review of the literature. In J. H. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 43–62). New York: Teachers College Press.

Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards for tests*. London: Sage Publication Ltd.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs: Prentice Hall.

Examinations Syndicate. (2009). *The briefing of National Educational Assessment System (NEAS) to 500 pilot schools*. Presentation material, Putrajaya: Lembaga Peperiksaan.

Examinations Syndicate. (2012). *Standard prestasi.* http://www.moe.gov.my/lp. Accessed January 28, 2012.

Gullo, D. F. (2005). *Understanding assessment and evaluation in early childhood education* (2nd ed.). New York: Teachers College Press.

Hargreaves, D. H. (1972). *Interpersonal relations and education*. London: Routledge and Kegan Paul Ltd.

Kubiszyn, T., & Borich, G. (2003). *Educational testing and measurement: Classroom application and practice* (7th ed.). USA: Wiley.

Lawrence, J. A. (2011). *An examinations of high school assessment and grading practices within a professional learning community*. Northern Illinois University, ProQuest LLC.

McMillan, J. H. (2007). *Classroom assessment: Principles and practice for effective standard-based instruction* (4th ed.). USA: Pearson Education Inc.

Novak, J. D. (2010). Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations. *Journal of e-Learning and Knowledge Society, 6*(3), 21–30.

Segers, M., Dochy, F., & Cascallar, E. (2003). The Era of assessment engineering: Changing perspectives on teaching and learning and the role of new modes of assessment. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (pp. 5–12). Dordrecht: Kluwer Academic Publishers.

Taras, M. (2005). Assessment—Summative and formative—Some theoretical reflections. *British Journal of Educational Studies, 53*(4), 466–478. ISSN 0007-1005.

Assessing Writing Readiness for Academic Purposes Among Students in Higher Education

Lai Fong Lee, Gurnam Kaur Sidhu, Yuen Fook Chan, Narayanan Geethanjali and Sian Hoon Teoh

Abstract The ability to write well is important for academic success in higher education. The questions that arise are whether students are ready for academic writing in higher education and their awareness of their readiness. These questions were a catalyst for the researchers to develop a self-assessment checklist to investigate the readiness of students in higher education for academic writing. This paper describes the development of the writing readiness checklist based on theories and pertinent literature. The resulting checklist comprises two categories: writing skills and writing strategies. This checklist aims to create awareness among students on their readiness in writing and to guide them to obtain feedback to improve writing. A preliminary study using this checklist was conducted with a group of 40 ESL students in a local public university. This paper reports the findings for one writing skill, drafting. The students depicted higher capacity for organisation in writing their drafts but avoided writing multiple drafts. They indicated a degree of readiness for writing with regard to drafting. These findings were corroborated by interview findings and the checklist score. These findings have implications for using the checklist to guide students in higher education to develop writing readiness and proficiency and to take an active role in learning and for faculty to use the checklist to provide pertinent feedback.

Keywords Academic writing · Writing checklist · Self-assessment · Higher education

1 Introduction

Effective writing, an essential aspect of student learning outcome in higher education, is viewed as an important intellectual skill and is essential for academic and career success (Hasegawa 2013; Lee 2013; Lee et al. 2014a, b; Sparks et al. 2014).

L.F. Lee (✉) · G.K. Sidhu · Y.F. Chan · N. Geethanjali · S.H. Teoh
Universiti Teknologi MARA, Shah Alam, Malaysia
e-mail: leela679@salam.uitm.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_27

This need for writing competency in higher education demands assessments which guide curriculum and instructional revision with regard to enhancing student's writing skills, informing hiring decisions as well as in giving feedback to students on their readiness for academic and career pursuits (Sparks et al. 2014). In addition, lifelong learning in the knowledge society and pedagogies which are learner-centred has tasked language assessment to enhance the holistic growth of students and their autonomy in learning (Kostopoulou 2009). Lynch (2001) highlights that the practice of assessment should "actively enable the construction of the self as subject, rather than the object of assessment" (cited in Oscarson 2009, p. 58).

These views of assessment are in line with formative assessment which is frequently termed as assessment for learning. Formative assessment is mainly utilised to enhance student learning by providing them feedback on their progress in the process of learning (Black 1998; Yorke 2003; Oscarson 2009). Formative assessment supports learning through the following aspects. It focuses on learning instead of performance, develops skills for self-assessment skills, works on the premise that support is needed by students in learning and creates awareness among students that their beliefs about their abilities as learners impact the outcomes. It should also be in tandem with practices in teaching and learning, make a distinction between feedback from learning and grading, adopt a standard-based framework, as well as develop self-improvement, and enhance reflective learning among students (Black 1998; Black and William 1998; Singh and Terry 2008). In line with this, Singh and Terry (2008) stress that there is an increasing shift to formative assessment in higher education in an attempt to get students to focus on the processes of assessment and view them as a part of learning.

An integral aspect of formative assessment is self-assessment practices. Oscarson (1989) points out that self-assessment practices serve to enhance learning, increase awareness among students, enhance orientation of goals, promote learner autonomy and contribute to democratic learning processes and need analysis. Self-assessment also enables students to diagnose their abilities and problems and to find out what they lack and need to learn and the level they have attained for an assignment or a course. With regard to writing, Schendel and O'Neill (1999) assert that self-assessment heightens students' self-awareness of their writing and provides control to students and a measure of rhetorical agency (cited in Oscarson 2009). Research has indicated that students can self-assess their writing. They are able to develop their ability to perceive their strengths and weaknesses in writing and consequently become more autonomous and more responsible for their own writing. Their writing skills and performance also improve (Zheng et al. 2012; de Assis 2012; Honsa 2013; Mohammad Reza 2010).

One way to self-assess writing is through the use of checklist. Existing writing checklists for self-assessment purpose tend to be focused on processes of writing or are based on writing rubrics. An example of the former is Honsa's (2013) self-assessment writing checklist on writing processes in general, but the constructs

are derived from a writing textbook. He found improvements in students' linguistic quality in writing and an increase in their autonomy after they used the checklist. Another example is Nimehchisalem et al.'s (2014) self-assessment checklist which concentrates on processes before writing, while writing and after writing specifically for argumentative writing genre. It is based on theories: Pyramid of Argumentation and Process Approach in writing. There are also checklists developed to guide students in other academic writing genres such as research proposals (Emery et al. 2006) and qualitative research processes and writing (Frels et al. 2011). Meanwhile, Covill (2012) utilised writing rubrics to get students to check on their ability and confidence in writing tasks, and her findings were positive. In sum, different writing checklists exist, but Nimehchisalem et al.'s (2014) have questioned the lack of theoretical background for some of them.

Taking into considerations the positive implications of formative assessment, self-assessment, and checklists on writing, a self-assessment checklist was developed to guide students on their readiness for academic writing in higher education. The following section provides the framework for the checklist.

2 Framework for Self-assessment Checklist for Writing Readiness for Academic Purposes

The self-assessment checklist for writing readiness for academic purposes in this study was developed based on theoretical underpinnings and literature on learning and writing which are discussed briefly as follows: they are sociocultural theory, Vygotsky's zone of proximal development (ZPD), theories of writing, i.e. product approach and process approach, and writing strategies.

Sociocultural theory frames cognition in a social context, i.e. social relationship and culturally constructed artefacts where learning takes place, is integral in the learning process (Lantolf 2001; Lattuca 2002; Lantolf and Thorne 2006). Basically, sociocultural theory views learning as related to historical, cultural and institutional concepts. This theory also views learning as a socially mediated process whereby mediation is utilised to help people participate in an activity (Mitchell and Myles 2004; Johnson 2006). In this research, writing is contextualised in the academic writing of higher education. The self-assessment checklist is a mediatory tool to help students understand their readiness to write, and the feedback which follows from the completion of the checklist is also a mediation to develop student writing. Research on sociocultural theory and writing has furthered understanding of contexts of students' writing and mediation in writing (Thorne 2004; Lantolf and Thorne 2006; Lei 2008; Lee 2012).

In line with the sociocultural theory, ZPD was also used as a theoretical basis to develop the checklist in this study. ZPD is defined as a tool to comprehend the

internal course of development (Daniels 2001; Thompson 2013). Vygotsky argues that "the only good kind of instruction is that which marches ahead of development and leads it; it must be aimed not so much at the ripe as at the ripening functions" (Vygotsky, 1986 in Thompson 2013, p. 248). He also asserts that a person's development within a ZPD involves social interaction, dialogue and mediated activity between learners and their teachers.

Moving on to theories of writing, Leki (1992) states that writing as a text product and the formal aspects of writing should be accounted for in writing instruction. Accordingly, Grabe and Kaplan (1996) highlight that there are seven components which interact in a text construction. They are "syntactic structure, semantic senses and mappings, cohesion signalling, genre and organisational structuring to support coherence interpretations, lexical forms and relations, stylistic and register dimensions of text-structure, and non-linguistic knowledge bases, including 'world knowledge'" (Grabe and Kaplan 1996, p. 62). Next, focusing on writing as a process approach, it is divided into four stages: expressive stage, cognitive stage, social stage and discourse community stage. Models of writing such as Flower and Hayes model of writing and Bereiter and Scardamalia model of the writing process are widely used in the process approach. The process approach in writing promotes self-discovery, planning and pre-writing, and multiple drafting, while writing, revising and editing and all these processes occur recursively (Grabe and Kaplan 1996). The checklist in this study incorporates both aspects of the product and process approaches, with the intention of raising L2 students' awareness of rhetorics, language and vocabulary in the writing process. Such integration of product–process approach has been found to be positive to L2 students in writing (Lee 2001; Davies 2004; Tangkiengsirisin 2006; Pasand 2013).

The framework adopted for writing strategies in developing the self-assessed checklist for academic writing readiness in this study was Mu's (2005) Taxonomy of ESL Writing Strategies. Mu's (2005) taxonomy was selected due to its theoretical nature in contrast to other writing strategies. Mu devised the Taxonomy of ESL Writing Strategies by synthesising research on writing strategies and four vital theories in ESL writing. The four theories are Contrastive Rhetoric Theory, Cognitive Development Theory, Communication Theory and Social Constructionist Theory. There are five categories of strategies: rhetorical strategies, metacognitive strategies, cognitive strategies, communicative strategies and social/affective strategies in this taxonomy. Research on writing strategies among L2 students has linked writing strategies to improved proficiency in L2 writing (Sasaki 2000, 2007; Mu 2005; Mu and Carrington 2007; Saeid Raoofi et al. 2014). Research has also given support to writing strategies as essential to facilitate L2 writers' abilities in writing (Lei 2008; Muhammad Ridhuan et al. 2011; Nooreiny and Mazlin 2013; Saeid Raoofi et al. 2014) and they should be taught (Mimi Estonella and Nooreiny 2014).

3 Self-assessment Checklist for Writing Readiness for Academic Purposes

This section describes the checklist for students to self-assess their writing readiness for academic purposes in higher education. The idea to develop the checklist was spurred on by the problems students faced in writing in higher education and the quality of their writing texts (Lee 2012, 2013; Lee et al. 2014b). This checklist was developed based on learning and writing theories and research literature (refer sections above). The checklist consists of two categories: writing skills and writing strategies to help students to assess their capacity and readiness for academic writing. These two categories were incorporated as existing writing checklists tended to focus more on writing skills (Covill 2012; Honsa 2013; Nimehchisalem et al. 2014). The checklist also addressed theoretical basis missing in the development of some writing checklists (Nimehchisalem et al. 2014). The rationale for this checklist is that students should be cognizant of the skills and strategies involved in writing. Here, the checklist acts as a form of reference, a criterion for skills and strategies to achieve effective writing. Students need to know whether their repertoire of writing skills and strategies is sufficient to accomplish effective writing in the academic context. The checklist as a tool for self-assessment is hoped to create awareness among students on their readiness for academic writing with regard to their writing skills and strategies. It can indicate to students the writing skills and strategies which they lack and guide them to take the necessary steps to address this problem. Reflecting formative assessment, this checklist also serves as an interactive tool, i.e. promoting feedback between students and the faculty, and provides a basis for faculty to input feedback. The self-assessment checklist is also envisioned to promote better writing proficiency as well as autonomy and self-regulation among students.

Specifically, the self-assessment checklist for writing readiness for academic purposes consists of writing skills and writing strategies. Writing skills comprise of five categories: analysis, obtaining information, planning, drafting and revising. There are four subskills for analysis, five for obtaining information, six for planning, eighteen for drafting and nine for revising. Aspects of the product (rhetorics, language and vocabulary) and process (pre-, while and post-writing) approaches are incorporated in these writing skills. These are derived from the product and process approaches in writing. Writing strategies comprise of five categories derived from Mu's (2005) Taxonomy of ESL Writing Strategies. They are rhetorical strategies, metacognitive strategies, cognitive strategies, communicative strategies and social/affective strategies with various substrategies. There are eight substrategies for rhetorical strategies, eleven for metacognitive strategies, twelve for cognitive strategies, four for communicative strategies and nine for social/affective strategies.

Students self-assess their writing skills and strategies by rating their level of agreement based on a four-point scale ranging from 1 for strongly disagree, 2 for disagree, 3 for agree and 4 for strongly agree. This self-assessment checklist was programmed in software. Marks are assigned corresponding to the Likert scales: 1

mark for strongly disagree, 2 marks for disagree, 3 marks for agree and 4 marks for strongly agree. Once students have completed the checklist, the average score for every skill and strategy and the scores for overall skills and overall strategies are displayed. A rubric of the scores and their interpretations of students' level of writing readiness was also developed for students' reference. It was determined that 1.0–1.9 marks indicated poor writing readiness, 2.0–2.9 indicated average readiness and 3.0–4.0 indicated good writing readiness. Students can then check their level of writing readiness and consult their writing instructors for further guidance.

The items in the self-assessment checklist were checked by two experienced teaching-English-as-a-second-language lecturers for content validity. These two lecturers whose area of expertise was in writing had more than 20 years of experience in teaching in university. They gave feedback for the checklist, and some items were refined and others added on. The modified checklist was checked again by the two same experts. Next, a pilot test was conducted with 30 students. The majority of them indicated that they understood the items in the checklist. Next, the Cronbach's coefficient alpha was calculated to check the internal reliability of the instrument. George and Mallery (2003) indicated that "_>.9—Excellent, _>.8—Good, _>.7—Acceptable …" (cited in Gliem and Gliem 2003, p. 87). The Cronbach's alpha value for this self-assessment checklist was 0.855, and was considered good. The same two experts also validated the checklist with feedback for improvement.

This paper goes on to report the findings obtained for one aspect in the self-assessment checklist: drafting. The findings were triangulated with the checklist rubrics scores for drafting and interview findings to ascertain whether students could self-evaluate their writing readiness.

4 Methodology

The population of this study consisted of students from the social sciences discipline in a local public university in Malaysia. The sample was 40 ESL undergraduates who were randomly selected. They were in their first and second year at the selected local public university. This study employed a mixed methods design. The students were given the self-assessment checklist for writing readiness for academic purposes developed for this study. They were asked to rate the writing skills and strategies they used based on a four-point scale ranging from 1 for strongly disagree, 2 for disagree, 3 for agree and 4 for strongly agree. Semi-structured personal interviews were conducted with five students. The results from the checklist were analysed and presented descriptively and inferentially. The interview data were transcribed, coded into themes and triangulated with the checklist data. It should be noted that this paper only focused on one aspect in the checklist, which is drafting in the findings and discussion section. The Cronbach's alpha for drafting skill in the self-assessment checklist was 0.912, and was considered as excellent.

5 Findings and Discussion

Table 1 shows that in the self-assessment checklist for writing readiness for academic purposes, the mean for drafting was 2.88 and the standard deviation was 0.385. The findings for drafting skill in Table 1 are further explained and elaborated after the explanation for Table 2.

Since the mean of 2.88 (drafting skill) is descriptively higher than the tested value (the boundary between 'agree' and 'disagree'), a one sample *t* test was conducted to find the significance of the level. Table 2 shows that the level of usage is significantly different from the test value 2.5 with sig <0.05; thus, there is sufficient evidence to conclude that the students significantly agree on the usage of the skills in drafting.

Focusing back on Table 1, it shows that in drafting, the skill, "I make sure my writing has introduction, developmental paragraphs (body) and conclusion", was rated descriptively highest (mean = 3.27 with standard deviation = 0.549) among the skills. All the five students who were interviewed were very clear about these aspects in writing their drafts. Nevertheless, they still had the perception that they should write only one paragraph each for introduction and conclusion. This can be because they had yet to expand their repertoire of genres for academic writing (Lee 2013).

The drafting skill rated descriptively second highest was "I ensure that my introduction provides a clear context for my topic" (mean = 3.07 with standard deviation = 0.469). This was followed by "I am able to ensure that the ideas in every paragraph are logically related" (mean = 3.05 with standard deviation = 0.444). The findings for these two items indicated that students were aware of cohesiveness in writing. To illustrate, Student 1 noted that the introduction "must fit the topic" and "the ideas in the paragraphs are together". Student 3 pointed out that "linking topic, introduction and content is important" in writing. Student 5 also agreed but admitted that the ease of relating ideas in paragraphs depended on the writing task.

Next, the skill "I am able to develop supporting details for the developmental paragraphs" (mean = 3.0 with standard deviation = 0.447) was rated descriptively fourth highest for drafting. When interviewed, the students generally stated that they were capable of developing details for their writing as they could obtain information from reading online sources. This concurs with Lee et al.'s (2014b) findings. Student 2 added, "I do a lot of citations as I want my writing to have content and proof". Meanwhile, Student 4 stated, "I focus on the content… it's ok". The findings can be linked to students' awareness that academic writing demands robust support and expansion of information utilised in the written text. Lea and Street (1998) pointed out that students' writing in higher education needs to reflect knowledge of their discipline subjects as this is often demanded for in the writing tasks given to students.

Table 1 Students' drafting skills in self-assessment checklist for writing readiness

Item	Mean (*M*)	Standard deviation (SD)
Drafting skill	2.88	0.385
I write multiple drafts before handing in my work	2.46	0.778
I make sure my writing has introduction, developmental paragraphs (body) and conclusion	3.27	0.549
I am able to write a clear thesis statement for my introduction	2.88	0.640
I ensure that my introduction provides a clear context for my topic	3.07	0.469
I am able to form a stated or an implied topic sentence for each developmental paragraph	2.95	0.590
I am able to develop supporting details for the developmental paragraphs	3.00	0.447
I am able to ensure that the ideas in every paragraph are logically related	3.05	0.444
I am able to paraphrase information accurately for my writing	2.88	0.600
I am able to summarise information accurately for my writing	2.85	0.573
I am able to quote correctly for my writing	2.90	0.735
I am able to synthesise information accurately for my writing	2.68	0.687
I am able to accurately cite my sources	2.90	0.700
I am able to summarise the main points in my conclusion	2.93	0.608
I am able to relate my conclusion back to my topic and focus	2.93	0.565
I am able to use a variety of vocabulary when I write	2.56	0.634
I am able to write grammatically correct sentences	2.66	0.693
I am able to use direct and clear language for my writing	2.90	0.625
I am able to use appropriate formal academic writing style in my writing	2.88	0.510

Table 2 One sample *t* test for drafting skill

Test value = 2.5						
	t	df	Sig. (2-tailed)	Mean difference	95% confidence interval of the difference	
					Lower	Upper
Total	6.235	40	0.000	0.375	0.25	0.50

On the other hand, the skill in drafting rated descriptively lowest was "I write multiple drafts before handing in my work" (mean = 2.46 with standard deviation = 0.778). When interviewed, Students 1 and 4 stated that they wrote two drafts before submitting their work, whereas Students 2, 3 and 5 wrote only one draft. The findings indicate that students may not be well attuned to the fact that good writing

is not the product of a single draft but multiple drafts of revising (Morrison-Saunders 2015). Moreover, the findings reflect that during writing, students should be more aware that they need to rethink and reflect consistently on their written work (Chien 2008). Other research also found students tend to view their writing as a single draft process (Nooreiny and Mazlin 2013).

The drafting skill rated second lowest was "I am able to use a variety of vocabulary when I write" (mean = 2.56 with standard deviation = 0.634). Preceding this was "I am able to write grammatically correct sentences" (mean = 2.66 with standard deviation = 0.693). These findings on vocabulary and grammar depict that the students were aware of the need to use appropriate language in academic writing. Their view of their abilities on their vocabulary and grammar was comparatively lower which indicated that they may need help in these areas. This was supported by the interview findings. For example, Student 1 said, "Hard to elaborate points… need grammar"; Student 2 stated, "The problem… you need good grammar"; Student 3 indicated, "… I can't find words I want"; and Student 4 found that in trying to write paragraphs, she was "confused with grammar and sentence". The situation can also be due to students' level of proficiency as the majority of them are at the intermediate level. These findings with regard to language in writing are supported by similar findings in Lee's (2013), Lee et al.'s (2014b) and Giridharan's (2012) research.

The skill in drafting rated descriptively fourth lowest was "I am able to synthesise information accurately for my writing" (mean = 2.68 with standard deviation = 0.687). Students 2 and 5 were inclined to put information "one by one" as it was less stressful. Students 1 and 4 tried to synthesise but found it "confusing" and "take up time", respectively. These findings reflect that students still need guidance in synthesising information. This finding concurs with the findings from Lee's (2014b) research that students in higher education indicated difficulties in synthesising in writing. This problem also raises the question of the level of students' ability at critical analysis.

Overall, the findings indicated that the skills related to drafting in the self-assessment checklist for writing readiness for academic purposes which were rated descriptively higher by the students were linked to organisation and development of content. In contrast, those which were rated descriptively lower were to do with revising through multiple drafts, language and synthesising of information. It should be noted that the items rated descriptively lower were still higher or at the acceptable mean level (2.5). The checklist rubrics score of 3.0 indicated good level of readiness among the students [note that this is at the border between average (2.0–2.9) and good (3.0–4.0)]. The interview findings offered further insights, and taking into account the students' views for the items rated descriptively lower, it can be surmised that to an extent, the students showed readiness and awareness of drafting but this needs to be further enhanced. These findings can serve as a reference to indicate to students their level of readiness for writing for academic purposes with regard to drafting.

6 Conclusion

This preliminary study on the checklist for writing readiness for academic purposes indicates that it can be a useful tool to promote effective academic writing in higher education. As in the findings for the writing skill, drafting reported in this paper, it informs on the students' agreement of their use of aspects related to drafting, which in line with the aim of the checklist offers a view of students' preparedness for writing. Further studies hope to look into the link between the level of students' writing readiness indicated by the self-assessment checklist and their writing texts. It should be noted that the checklist as Nelson et al. (2012) reminded is not "an end in itself" but is "an effective departure point for the ongoing dialog" to guide students' writing (p. 380). Similarly, students can use the checklist developed in this study as a mediatory tool to gauge their level of readiness for academic writing. The checklist can shed light on writing skills and strategies which students need to demonstrate or utilise in writing. It can help to raise students' consciousness on what is lacking and to take autonomy for their learning by seeking guidance to enhance their writing ability. The faculty too can pinpoint students' needs in writing and work interactively with students in providing them the necessary feedback. To conclude, self-assessment practice such as the checklist for writing readiness for academic purposes has positive implications to help students write effectively, a much required skill in the twenty-first century.

Acknowledgements The authors gratefully acknowledge funding by Fundamental Research Grant Scheme (FRGS) from the Ministry of Education (MOE) and Universiti Teknologi MARA (UiTM).

References

Black, P. (1998). Formative assessment: Raising standards in the classroom. *School Science Review, 80*(291), 39–46.

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–74.

Chien, S.-C. (2008). A cognitive analysis of the relationships between Chinese EFL writers' strategy use and writing achievement performance. *Cambridge Occasional Papers in Linguistics (COPiL), 3*, 44–61.

Covill, A. E. (2012). College students' use of a writing rubric: Effect on quality of writing, self-efficacy, and writing practices. *The Journal of Writing Assessment, 5*(1), 1–19.

Daniels, H. (2001). *Vygotsky and pedagogy*. London: Routledge Falmer.

Davies, R. J. (2004). Writing discourse across cultures I: Towards an integrated approach to EL2 composition pedagogy. *Ehime University Faculty of Education Bulletin, 36*(2), 77–100.

de Assis, B. M. (2012). *Self-assessment of writing skills: A reliable and valid tool in an EFL classroom*. Master's project. Universidade Nova De Lisboa.

Emery, L. J., Harvey, C., & Andersen, C. M. (2006). Formative evaluation using checklists to improve research proposals. *Perspect Health Information Management*, *3*(2). Retrieved July 10, 2015 from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2047304/

Frels, R. K., Sharmab, B., Onwuegbuziec, A. J., Leechd, N. L., & Stark, M. D. (2011). The use of a checklist and qualitative notebooks for an interactive process of teaching and learning qualitative research. *The Journal of Effective Teaching, 11*(1), 62–79.

Giridharan, B. (2012). Identifying gaps in academic writing of ESL students. *US-China Education Review, A6*, 578–587.

Gliem, J. A., & Gliem, R. A. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. 2003 *Midwest Research to Practice Conference in Adult, Continuing, and Community Education*. Retrieved July 25, 2015 from www.ssnpstudents.com/wp/wp-content/uploads/2015/…/Gliem-Gliem.pdf

Grabe, W., & Kaplan, R. B. (1996). *Theory and practice in writing: An applied linguistics perspective*. London: Longman.

Hasegawa, H. (2013). Students' perceptions and performances in academic essay writing in higher education. *International Journal of Innovative Interdisciplinary Research, 4*, 1–14.

Honsa, S, Jr. (2013). Self-assessment in EFL writing: A study of intermediate EFL students at a Thai University. *Voices in Asia Journal, 1*(1), 34–57.

Johnson, K. E. (2006). The sociocultural turn and its challenges for second language teacher education. *TESOL Quarterly, 40*(1), 235–257.

Kostopoulou, S. (2009). Democratising and enhancing the quality of institutionalised language assessment through the European language portfolio. In L. Taylor & C. J. Weir (Eds.), *Language testing matters* (pp. 263–277). Cambridge: Cambridge University Press.

Lantolf, J. P. (2001). (S)econd (L)anguage (A)ctivity theory: Understanding second language learners as people. In M. P. Breen (Ed.), *Learner contribution to language learning: New directions in research* (pp. 141–158). England: Pearson.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.

Lattuca, L. R. (2002). Learning interdisciplinarity: Sociocultural perspectives on academic work. *The Journal of Higher Education, 73*(6), 711–739.

Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education, 23*(2), 157–173.

Lee, B. P. H. (2001). Integrating product-process considerations in teaching business letter writing. *Asian Journal of English Language Teaching, 11*, 113–128.

Lee, L. F. (2012). Changing participation and identity in writing transition: A case study. *Asian Journal of University Education, 8*(1), 67–84.

Lee, L. F. (2013). Writing transition in postsecondary education: A case study. *6th International Conference on University Learning and Teaching (InCULT 2012)* (*Procedia—Social and Behavioral Sciences*) (No. 90, pp. 267–277).

Lee, Lai Fong, Gurnam Kaur Sidhu, & Chan Yuen Fook (2014a). Exploring 21st century skills among postgraduates in Malaysia. *Procedia-Social and Behavioral Sciences, 123*, 130–138.

Lee, Lai Fong, Gurnam Kaur Sidhu, Chan Yuen Fook , Teoh Sian Hoon, Geethanjali Narayanan , & Mohd Ismail Azizi bin Md Ishak (2014b). *Examining contributory factors and barriers to success in academic writing among undergraduates*. Paper presented at regional conference on science, technology and social sciences, 23–25 November, Pahang, Malaysia.

Lei, X. (2008). Exploring a sociocultural approach to writing strategy research: Mediated actions in writing activities. *Journal of Second Language Writing, 17*(4), 217–236.

Leki, I. (1992). *Understanding ESL writers: A guide for teachers*. London: Heinemann.

Mimi Estonella Mastan, & Nooreiny Maarof. (2014). ESL learners' self-efficacy beliefs and strategy use in expository writing. *5th World Conference on Educational Sciences* (Vol. 116, pp. 2360–2363) (*Procedia—Social and Behavioral Sciences*).

Mitchell, R., & Myles, F. (2004). *Second language learning theories*. London: Arnold.

Mohammad Reza Javaherbakhsh. (2010). The impact of self-assessment on Iranian EFL learners' writing skill. *English Language Teaching, 3*(2), 213–218

Morrison-Saunders, A. (2015). Writing about writing in higher education: Modelling good practices. In *Teaching and Learning Forum 2015—Teaching and learning Uncapped, 29–30 January*. Perth: University of Western Australia.

Mu, C. (2005). A Taxonomy of ESL writing strategies. *Proceedings Redesigning Pedagogy: Research, Policy, Practice, May 30-June, Singapore* (pp. 1–10). Retrieved March 2, 2015 from http://eprints.qut.edu.au/view/person/Mu,_Congjun.html

Mu, C., & Carrington, S. (2007). An investigation of three Chinese students' english writing strategies. *TESL-EJ*, *11*(1). Retrieved June 23, 2015 from www.tesl-ej.org/ej41/a1.pdf

Muhammad Ridhuan Tony Lim Abdullah, Zulqarnain Abu Bakar, Razol Mahari Ali, Raja Ahmad Iskandar Raja Yaacob, Abdur-Rahman M. A., Abd Mutalib Embong, & Am Zamri Amar. et al. (2011). Writing strategies of Malaysian ESL undergraduate engineering learners. *International Journal of Engineering and Technology, 11*(2), 1–9.

Nelson, J. S., Range, L. M., & Ross, M. B. (2012). A checklist to guide graduate students' writing. *International Journal of Teaching and Learning in Higher Education, 24*(3), 376–382.

Nimehchisalem, V., Yoong Soon Chye, D. Sheena Kaur A/P Jaswant Singh, Siti Zaidah Zainuddin, Sara Norouzi, & Sheren Khalid. et al. (2014). Self-assessment checklist for undergraduate students' argumentative writing. *Advances in Language and Literary Studies, 1*, 65–80.

Nooreiny Maarof & Mazlin Murat (2013). Writing strategies used by ESL upper secondary school students. *International Education Studies, 6*(4), 47–55.

Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and application. *Language Testing, 6*(1), 1–13.

Oscarson, A. D. (2009). *Self-assessment of writing in learning English as a foreign language: A study at the upper secondary school level*. Sweeden: University of Gothenburg.

Pasand, P. G. (2013). Process-product approach to writing: The effect of model essays on EFL learners' writing accuracy. *International Journal of Applied Linguistics and English Literature, 2*(1), 75–79.

Saeid Raoofi, Chan Swee Heng, Jayakaran Mukundan, Sabariah Md Rashid. (2014). A qualitative study into L2 writing strategies of university students. *English Language Teaching, 7*(11), 39–45.

Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing, 9*(3), 259–291.

Sasaki, M. (2007). Effects of study-abroad experiences on EFL writers: A multiple-data analysis. *The Modern Language Journal, 91*(4), 602–620.

Singh, K., & Terry, J. (2008). Fostering students' self assessment skills for sustainable learning. *EDU-COM International Conference*. Edith Cowan University.

Sparks, J. R., Song, Y., Brantley, W., & Liu, O. L. (2014). Assessing written communication in higher education: Review and recommendations for next-generation assessment. *ETS Research Report Series, 2014*(2), 1–52. doi:10.1002/ets2.12035.

Tangkiengsirisin, S. (2006). Approaches to teaching second language writing. *Language Institute Journal, 3*, 1–26.

Thompson, I. (2013). The mediation of learning in the zone of proximal development through a co-constructed writing activity. *Research in the Teaching of English, 47*(3), 247–276.

Thorne, S. L. (2004). Cultural historical activity theory and the object of innovation. In O. St. John, K. van Esch, & E. Schalkwijk (Eds.), *New insights into foreign language learning and teaching* (pp. 51–70). Frankfurt: Peter Lang Verlag.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*, 477–501.

Zheng, H., Huang, J., & Chen, Y. (2012). Effects of self-assessment training on Chinese students' performance on college English writing tests. *Polyglossia, 23*, 33–42.

Perceptions on Love of the Young People: An Assessment Based on the Love Attitude Scale

Fides del Castillo

Abstract The construct of love spans different epochs, cuts across cultures, and presents varied meanings that it has received considerable attention and discussion from various philosophers, psychologists, scholars, as well as the Catholic Church. The classical philosopher Socrates describes love as a "hunger of the human soul for divine beauty" while Plato asserts that love is an emotion that exists in hierarchical order (Levy and Davis in Journal of Social and Personal Relationships 5:439–471, 1988). From a clinical psychologist's perspective, love is an act of will that enables a person to commit oneself to another and carry out loving acts toward the other (Peck 1989; Fromm 2000). While, some scholars argue that certain human needs must first be fulfilled before a person can share his or her love with another (Judd in Love and lifestyles. Saint Mary's Press, Minnesota, 1981; Maslow in Psychological Review 50:370–396, 1943); others maintain that love can only be correctly described according to its different levels: *eros* or sexual desire, intimacy, commitment, passion, and agape (Sternberg in Love is a story: a new theory of relationships. Oxford University Press, USA, 1998; Fisher in Why we love: the nature and chemistry of romantic love. Holt, USA, 2004; Hendrick et al. in Journal of Sex Research, 43:76–86, 2006). With these various meaning of love presented by different authors, what does the young people perceive about it? This paper intends to access the understanding of the youth toward love in parallel to the definitions presented by the different discipline. The researcher employed quantitative survey research method as the basic design of study. The quantitative research design includes three stages of preparations in order to create an attitudinal scale, namely (1) development of the love attitude scale (LAS) developed by this researcher, (2) the administration of the LAS to a new set of 1000 student respondents, (3) analysis of the results using descriptive statistics, *t* test, and analysis of variance (ANOVA).

Keywords Love scale · Perceptions of youth · Attitude toward love

F. del Castillo (✉)
De La Salle University Manila, Manila, Philippines
e-mail: fides.delcastillo@dlsu.edu.ph

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_28

1 Introduction

Love is understood in different meanings and cuts across culture, beliefs, and references. Due to the great interest on the topic of love, there is a rich base of literature and studies in the area of religious education, psychology, anthropology as well as researches on love attitude scale (LAS) that are indispensable tools for an in-depth study on love attitudes. This research reviews local and foreign sources and presents them in a thematic way to show the relevance and uniqueness of this study.

1.1 In Religious Education

"God is love, and he who abides in love abides in God, and God abides in Him" (I Jn 4:16). Benedict XVI emphasized in *Deus Caritas Est* that Jesus is the incarnate love of God. By dying on the Cross in order to raise and save man, Jesus expressed love in its most sublime form (#12–15). Also, the love of neighbor must be grounded in the love of God. Hence, as responsible members of the Church, one must reflect the Trinitarian love in service and charity (# 16–18).

The *Dogmatic Constitution on Divine Revelation* (Dei Verbum) of Vatican II succinctly tells us the ministry and work of Jesus (4): Jesus Christ sent as a man among men; speaks the words of God (cf. Jn 3:34) and accomplishes the saving work, which the Father gave him to do (Jn 5:36; 17:4). He did this by the total act of his presence and self-manifestation—by words and works, signs and miracles, but above all by this death and glorious resurrection from the dead, and finally by sending the Spirit of truth.

Jesus told his disciples "as the Father loves me, so I also love you" (Jn 15:9). Jesus constantly manifested his love for people which culminated in his passion, death, and resurrection. The whole life of Jesus is a perfect life of charity in which man can share through the Holy Spirit. Thus, the very foundation of love for others (*neighbor*) is the love revealed by Jesus Christ who shares this love with us through the Holy Spirit.

Christ reveals to man their basic vocation as persons by giving himself and dying on the Cross. The Theology of the Body of John Paul II (2006) presents the Gospel message of love in a new, deep, and profound way because "man cannot live without love (p. 1)." John Paul II focused on the biblical teaching that man are made in the image of God. Hence, every person is sacred and a gift and this includes the human body. To love is a commitment and task that require self-giving and respect. In the book Love and Responsibility (1993), John Paul II also presented how love is seen as a responsibility. It is a commitment founded in the love of God. This confirms that love is based on a self-giving act. It is not reduced to sex neither a commodity.

de Mesa and Cacho (2011) said that "falling and staying in love with God is crucial (p. 71)." But since God has shown us how to truly love (agape), it would be easier for us to stay in love with Him and thus show a self-giving love to other people. Staying in love is a choice and a personal commitment. It does not linger on what is superficial. Every person called to happiness is invited to love since love is in the center of one's life (Rahner 1967). This kind of love must be nurturing which will allow the person to become aware of his or her goodness (Judd 1981). This affirms the love is more than what the external eyes can see. Love goes beyond and is expressed by acts of kindness, goodness, and commitment.

The following literatures present a clear narrative of the meaning of love in religious education. They support the concept of love in this study. The predefined factors of LAS state that love is a self-giving act manifested in loving action that includes commitment and service. Commitment in this respect is oriented toward God and other persons in the community. On the other hand, service finds its realization in actualization of commitment seen in communities. This concept is deeply rooted in one's faith in God in imitation of Christ's deeds.

1.2 In Psychology

Hazan and Shaver (1987) describe love as an "attachment process" wherein both healthy and unhealthy forms of love originate as reasonable adaptations to specific social circumstances. The portrait of love offered by attachment theory includes negative as well as positive emotions: for example, fear of intimacy (discussed by Hatfield 1984), jealousy (e.g., Hindy and Schwarz 1984), and emotional ups and downs as well as caring (Rubin 1973), intimacy (Sternberg 1986), and trust (Dion and Dion 1985). Also, attachment theory describes "separation and loss" and helps explain how loneliness and love are related to each other (Shaver et al. 1986). Finally, attachment theory links adult love with socio-emotional processes evident in children and non-human primates; it places love within an evolutionary context. It proves that human beings go through different love experiences based on their situation. The portraits of love offer positive as well as negative experiences. Hence, love starts from emotion and builds within us once it is nurtured. This can be acquired based on the person's experience or non-experience of love.

From another perspective, Fisher (2004) uses the constructs lust, attraction, attachment, intimacy, commitment, and passion to describe love. However, the humanistic philosopher Erich Fromm (2000) describes love as not merely a feeling but also as an act. The feeling of love is superficial in comparison with one's commitment to love via a series of loving actions over time (Fromm 2000). Stenberg (2011) enumerates intimacy, commitment, and passion as the different components of love. He claims that those who share confidences and various details of their personal lives with other people are considered to be in an intimate relationship. Examples of these are friendships and romantic love affairs. Also, the

love between two committed persons means that they strive to keep the relationship permanent. Furthermore, Sternberg (2011) claims that passionate love is founded on sexual attraction. Because of this, a very passionate love often leads to heartaches, frustrations, and various neuroses. Hence, love should develop from a feeling to commitment.

The aforementioned literature explained that love is acquired and developed from a person's experience and non-experience of love. Love can be described based on how it was encountered and learned as well as from the circumstances and factors that surround a person. This paper gathered the love constructs of students, and the survey results have shown that the students' concept of love is based on their experience and non-experience of love. Also, the students' constructs on love vary from love as a mere feeling to love as a commitment.

The LAS instrument will be used to survey 1000 Lasallian students. The instrument will also cover the antecedent such as gender, family structure, and relationship background. This will show the relevant insights on the result of the survey between male and female, students who live with parents and those who live with their guardians, and also those students who have been in a love relationship as well as those who have not. Knowing the students' background and experiences will allow the researcher to determine significant insights about their concept of love. Hence, this new knowledge will be a valuable contribution to religious education.

1.3 In Other Areas

From the eye lens of anthropology, the experience of love is described and divided into three partly overlapping stages: lust, attraction, and attachment (Fisher 2004). "Lust" is the feeling of sexual desire of human beings to other people. It is a natural occurrence and response to a feeling or an intense sensation. What follows lust is "attraction." It determines what people find attractive in another person and sets the level of desire and value for another person. The last stage is "attachment" which involves a higher level of love like sharing a home, parental duties, mutual love, and security. Ultimately, "attachment" necessitates a desire to care for another person.

The Typology Love Theory describes six love styles divided into two categories. The primary love includes *Eros, Ludus,* and *Storge* while the secondary love consists of *Mania, Pragma,* and *Agape*. Eros refers to the strong physical and emotional attraction followed by commitment to the loved one (Lee 1973). *Ludus* love refers to playful love or game-playing love. In this love style, there is no commitment toward love and the partner (Lee 1973). *Storge* refers to love relationship developed slowly from friendship (Lee 1973). Those who practice *Storge* have a strong commitment toward the relationship. *Pragma* love is a combination between *Ludus* and *Storge* where love is perceived as realistic and practical. In *Pragma* love, there is less focus on physical attraction but emphasizes the conscious

search for a compatible partner. *Manic* love or possessive love is characterized as a combination of erotic and *Ludic* love styles which results in an obsessive, intense, full feeling and possessive kind of love. People who engage in *Manic* love have a strong need to be loved. *Agape* love or altruistic love is a combination of erotic and *Storgic* love. This type of love needs attention from the loved ones without having personal interest. This love is seen as intense and full of friendship and increases the quality of love with the need for mutual helping (altruism).

The aforementioned studies help us understand that love is a natural occurrence. Everyone undergoes the process of being loved as well as expressing it. Also, love has various levels. The intensity starts from desire and may progress to agape or just stop on the other stages. The LAS will show the progression of the students' concept of love. Also, the LAS factors confirm that love starts from a personal affinity toward someone and then gradually develops to an act of service and commitment toward God and others.

Underwood (2009) proposed the model of compassionate love wherein an individual expressing compassionate love begins with a base of individual variations in personality, biology, and developmental patterns. Also, compassionate love is shaped within and by cultural, historical, family, and social environments. As such, people who were well-loved when they were children affect their capacity to love others. The expression of compassionate love can also be affected by the specific situation and the relationship to the person being loved. For example, people typically express different attitudes and behaviors to "ingroup" members (family, friends, relatives, and other similar groups) than to "outgroup" members. How each individual defines our "outgroup" varies, but most people do have a distinct sense of the strangers that affects how they relate to people. Thus, both situational factors and specific relationships can affect how compassionate love is expressed. Different elements may affect the individual's capacity to love depending on their experiences in life.

Although the individual person and final actions are important parts of the model, at the center of the model of compassionate love are motivation and discernment. Motivation and discernment are integral parts for the "moment of choice." At some point, a person internally reflects and makes a choice to move, to act, to express something, that is centered on the good of the other. At this moment, the person balances the various aspects of the situation, for example their own needs, priorities of obligation, fairness assessments, and perceived urgency. They also discern the appropriateness of action, sometimes explicitly and analytically, and sometimes with more of a gut sense, a more implicit process. Thus, both motives and discernment are keys in the moment of choice as well as the behavior flow from that choice (Underwood 2009).

The results of discernment are actions and attitudes. Positive behavior can result from compassionate love or from non-loving choice. Because of this, it is very difficult to judge the compassionately loving quality of an action. Underwood (2009) adds that an example of the complexity of judging love by actions is the case of someone who wants to donate money to a university, but will only do so if it

goes toward a specific building, and that building must have the name of the donor on the front. Hence, what is the central motive of the giver—is it love, "centered on the good of the other," or is it "centered on the good of oneself"? This is not an easy call, but if the money will be given only if the name appears on the building, then that is a hint regarding motivation. Therefore, we can say that genuine caring attitudes can in and of itself the suffering of another even before we take any action. The mere willingness to give of oneself for the good of another can produce a positive result in another person even when the actions and expressions are quite subtle.

1.4 Attitude Scale

Attitude is defined as the degree of positive and negative affect associated with some psychological object. Psychological objects are referred by Thurstone (1946) to be as "symbol, phrase, slogan, person, institution, ideal or idea toward which people can differ with respect to positive or negative affect." While Hogg and Vaughan (2005) stated that "an attitude is a relatively enduring organization of beliefs, feelings and behavioural tendencies toward socially significant objects, groups, events or symbols (p. 150)." An attitude has three components, namely affective, behavioral, and cognitive. The affective component is primarily involved in a person's feelings or emotions. The behavioral component is defined as the way the attitude have influences a person act or behave (Hogg and Vaughan 2005). And lastly, the cognitive component is involved in a person's belief and understanding about an attitude object.

The attempt to measure the traits and attitudes of persons is as nearly as old as the technique in measuring intellectual capacity. There are many researchers who have surveyed and attempted to measure attitudes. Measuring attitudes have been developed and done by Murphy and Murphy (1931), Bain (1930), Vetter (1930), Katz et al. (1931), Allport (1929), Likert (1932) and other researchers. In measuring attitudes, the link between the affect, behavior, and cognate lies in the consistency of the components. This means that the person's attitude is reflected in its consistency.

In this paper, the attitude toward love will be the main focus of the study. This section discusses the different studies on love attitudes made by different researchers.

Hendrick and Hendrick (1986) have developed a scale measuring the degree of endorsement of the love attitudes proposed by Lee (1973). Their study presented that the attitudes that people have on love direct their behavior and experiences toward the individual they love. Originally, Hendrick and Hendrick (1986) constructed a love scale for an extensive study of love and sex attitudes among undergraduate students. The scale was developed based on Lasswell and Lasswell's (1976) previous work, which was the sample profile. They converted the categorical nature of the sample profile to a Likert scale. Later, Hendrick and Hendrick (1986)

improved the sample profile which produced the LAS. Similar to the sample profile, the scale consists of six dimensions—*Storge, Agape, Mania, Pragma, Ludus and Eros*. Hendrick and Hendrick (1986) also built a short version of the LAS to provide researchers with an instrument to measure love styles. They produced two sets of short form of LAS which were the 4-item short form of LAS and the 3-item short form of LAS. To illustrate, each short form of LAS has six dimensions of love, *Storge, Agape, Mania, Pragma, Ludus and Eros*. In the 3-item short form of LAS, there was 18 items where all three items were used for every love dimension. Conversely, the LAS 4-item short form of LAS with 24 items was applied using four items for each dimension of love. It was proven that the 4-item and or 3-item versions have stronger psychometric properties compared to the earlier version of LAS (Hendrick and Hendrick 1986). This version was shown to be psychometrically sound compared to the first version (Tzeng 1993).

Previous researches have proven the LAS to be psychometrically sound (Tzeng 1993). This is because LAS has the ability to measure love at different stages in a relationship including for people who are in the "falling in love stage" (Hendrick and Hendrick 1990). In addition, the acknowledgment of a specific sexual item in the *Eros* dimension of love also allows researchers to study the existence of sexuality within the love context. LAS also allow researchers to understand how love actually functions rather than depending on just "liking" and "loving" factors of love proposed earlier by Rubin (1970).

Also, there are researchers who used the LAS of Hendrick and Hendrick (1986) to measure the similarities of the love attitude between parents and between parent and children (Amos et al. 2001), love attitudes in relationships (Hammock and Richardson 2011), love attitudes of Malaysians (Chong et al. 2012) and love attitudes across family life stages (Montgomery and Sorell 1997).

These researches and scale development on love prove that the study of love attitudes is important and the subject matter is measurable. Although one will find the LAS developed by the researcher in this study to be unique and different from the other love scales for the following reasons: (1) The LAS in this study was primarily developed to address the construct of love among Lasallian college students in the Philippines. The students' experience and non-experience of love were covered in the development of the instrument; (2) the LAS in this study has a different definition of love. Love is defined as a self-giving act manifested in loving action that includes commitment and service. Commitment is oriented toward God and the other persons in the community while service finds its realization in actualization of commitment seen in communities. This concept is deeply rooted in one's faith in God in imitation of Christ's deeds. As such, the current study is valuable and distinct. It is a newly developed research that will help religious educators increase knowledge about the Lasallian youth's attitude toward love in the Philippines.

2 Methodology

2.1 Research Design

The researcher will employ quantitative survey research method as the basic design of study. The quantitative research design will include four stages of preparations in order to create an attitudinal scale, namely (1) development of the LAS developed by this researcher, (2) the administration of the LAS to a new set of 1000 student respondents, (3) analysis of the results using descriptive statistics, *t* test, and analysis of variance (ANOVA).

Attitudinal scale measures the individual's beliefs, attitudes or feelings and behavioral tendencies about a specific topic. As for this research, the specific attitude that it intends to measure is the senior college students' attitude toward love. This attitudinal test used the Likert scale (1932) that allows respondents to indicate the extent of their agreement with a series of statements.

2.2 Respondents

The researcher targeted student respondents covering the age range of 15–24 years old from tertiary Lasallian institutions. For the purpose of clarity and consistency of this research, the researcher based the age range of the Lasallian youth from the United Nation's (UN) definition of youth (see www.unesco.org). Youth, according to UN definition, are people between 15 and 24 years of age.

2.2.1 Research Locale

The student respondents in the study are limited to particular tertiary Lasallian schools, namely De La Salle University, Manila; De La Salle College of St. Benilde; De La Salle Lipa; University of St. La Salle; and La Salle University, Ozamis. Three (3) of which is located in Luzon, one (1) in Visayas, and one (1) in Mindanao. The reason for choosing the university is determined by the number of tertiary schools across the regions.

3 Perceptions of Filipino youth Toward Love: Analysis of Variance (ANOVA)

In order to further determine the perceptions of Filipino youths toward love, the responses of the participants on the LAS were analyzed in *Statistica* (a statistics and analytics software developed by *StatSoft*). *Statistica* provides data analysis, data

management, statistics, data mining, and data visualization procedures. For this section, ANOVA will be utilized to determine the degree of difference of two or more groups of data. ANOVA is based on the comparison of the average value of a common component. It compares the mean scores between the groups and determines whether any of the means are significantly different from the others (Springer 2010). Since the mean scores of respondents in the four factors shall be compared, a one-way ANOVA between groups was used for analysis.

Also, to determine notable differences in mean scores from the four factors between the five schools, a post hoc test is employed after ANOVA. The post hoc test is used to explore the differences among means of the variables so as to know which group means are significantly different from each other (Stevens 1999). The specific analytical tool (Lowry 2008) used for the post hoc test in this study is the Tukey's honest significant difference (HSD).

4 Results of ANOVA and Post Hoc Test

Below are the results and descriptions of the one-way analysis of variance (ANOVA). The ANOVA table will show the sum of square (SS), degrees of freedom (df), mean square (MS), F-ratio (F), and significance (P). In ANOVA, the total SS helps express the total variation that can be attributed to various factors. Every time, the SS is computed the df appear. It is the minimum number of independent coordinates that can specify the position of the system completely (Iversen and Norpoth 1987). On the other hand, MS refers to an estimate of the population variance based on the variability among a given set of measure while F-ratio is the ratio of the variance between groups to the variance within. It means that MS measures the population variance while F-ratio test whether or not two variances are equal. Lastly, if the p-value of the ANOVA is below 0.05, it means that there is a significant difference in the variables. Meanwhile, the post hoc table will show the mean score, standard deviation, and significant difference of the variance of each school.

In order to fully appreciate the post hoc test, charts are provided. These will show and present the result of ANOVA in line chart. The number codes for the five schools are as follow: 1—DLSU, 2—CSB, 3—DLSL, 4—USLS, and 5—LSU (Table 1).

Table 1 shows that there is a significant difference among the students' perception toward love on all four factors. The Filipino youth's perceive the quality of love at the $p < 0.05$ level with $F(1, 4) = 5.8$, $p = 0.0001$. The mean scores of the student respondents for Committed subscale has a significant difference at the $p < 0.05$ level with $F(1, 4) = 7.33$, $p = 0.000008$. Moreover, the significant difference on the student respondents perception toward love as a Rooted in God is at

Table 1 ANOVA results

		SS	Df	MS	*F*	*P*
Engaged	Between groups	2,265,024	1	2,265,024	100,583	0.00*
	Within groups	525	4	131	5.8	0.000123*
	Total	2,265,549	5			
Committed	Between groups	1,594,534	1	1,594,534	52,295.6	0.000000*
	Within groups	894	4	224	7.33	0.000008*
	Total	1,595,428	5			
Rooted	Between groups	921,684	1	921,684	39,488.1	0.000000*
	Within groups	1916.3	4	479.1	20.53	0.000000*
	Total	923,601	5			
Integrated	Between groups	798,161	1	798,161	41,639.9	0.000000*
	Within groups	230.1	4	57.5	3	0.017834*
	Total	798,391	5			

Note $*p < 0.05$

Table 2 Post hoc test results

	Mean	SD	DLSU	CSB	DLSL	USLS	LSU
Engaged							
DLSU	4.53	0.3		0.334841	0.308744	0.219261	0.028719*
CSB	4.46	0.4	0.334841		0.010396*	0.995357	0.539065
DLSL	4.59	0.3	0.308744	0.010396*		0.006972*	0.000876*
USLS	4.45	0.5	0.219261	0.995357	0.006972*		0.768235
LSU	4.38	0.5	0.028719*	0.539065	0.000876*	0.768235	
Committed							
DLSU	4.31	0.3		0.295935	0.000045*	0.001759*	0.112059
CSB	4.4	0.2	0.295935		0.113002	0.388236	0.847831
DLSL	4.53	0.3	0.000045*	0.113002		0.990393	0.939517
USLS	4.5	0.3	0.001759*	0.388236	0.990393		0.995553
LSU	4.48	0.3	0.112059	0.847831	0.939517	0.995553	
Rooted							
DLSU	4.38	0.1		0.064787	0.000017*	0.000017*	0.000081*
CSB	4.53	0.1	0.064787		0.000375*	0.001712*	0.085855
DLSL	4.8	0.1	0.000017*	0.000375*		0.999976	0.982977
USLS	4.79	0.1	0.000017*	0.001712*	0.999976		0.992042
LSU	4.75	0.1	0.000081*	0.085855	0.982977	0.992042	
Integrated							
DLSU	4.22	0.3		0.041646*	0.270493	0.136713	0.316185
CSB	4.36	0.2	0.041646*		0.968609	0.999892	1
DLSL	4.33	0.4	0.270493	0.968609		0.992114	0.989482
USLS	4.35	0.3	0.136713	0.999892	0.992114		0.999957
LSU	4.36	0.3	0.316185	1	0.989482	0.999957	

Note $*p < 0.05$

$p < 0.05$ with $F(1, 4) = 20.53$, $p = 0.0000$. Lastly, the expression of love subscale is significantly different at $p < 0.05$ with $F(1, 4) = 3.0$, $p = 0.018$.

Post hoc comparisons using the Tukey's HSD test in Table 2 indicated that the mean score for the Engaged subscale is at $F(4, 899) = 5.8319$, $p = 0.00012$. It means that there is a significant difference among the Lasallians perception on Love as being Engaged. The attitude of students from De La Salle Lipa (DLSL) toward love is significantly different from De La Salle University, Manila (DLSU); De La Salle University College of St. Benilde (CSB); University of St. La Salle Bacolod (USLS); and La Salle University, Ozamis (LSU).

The chart in Fig. 1 shows De La Salle Lipa having the highest mean score of ($M = 4.59$, SD = 0.33). De La Salle University, Manila, got the second highest mean score ($M = 4.53$, SD = 0.34), followed by De La Salle University College of St. Benilde ($M = 4.46$, SD = 0.42), then University of St. La Salle Bacolod ($M = 4.45$, SD = 0.50), and lastly La Salle University, Ozamis ($M = 4.38$, SD = 0.51).

For the "Engaged" subscale, it is interesting to note that the top three schools are from Luzon. This indicates that demographics affect the students' perception toward love specifically on love as Engaged. Students from urban areas are more likely to express and affirm their love openly compared to students in rural areas. This claim is supported by studies on the differences of urban and rural settings (Freudenburg 2007; Stoodley 1959). Filipino youth from urban areas may have been exposed to increased modernization and changing times compared to those in rural areas who have a more conservative and simple life.

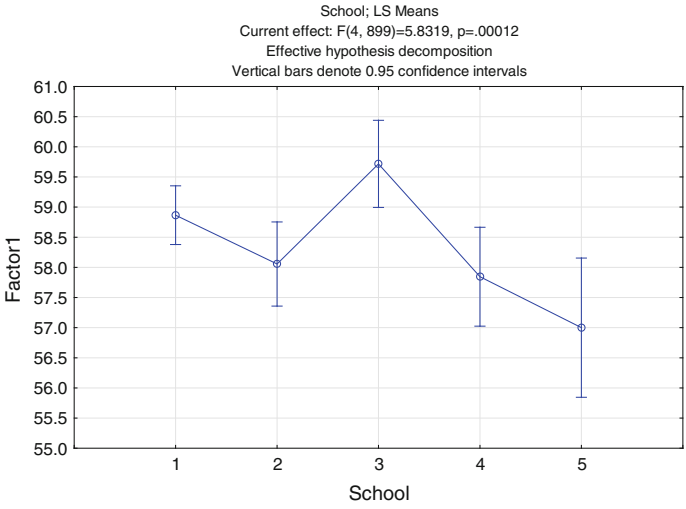


Fig. 1 Chart for Factor 1 Engaged

Furthermore, post hoc results in Table 2 indicate that the perception of Lasallian students on the Committed subscale has $F(4, 899) = 7.3334$, $p = 0.00001$. It is very interesting to note that in this factor, only De La Salle University, Manila, had a low mean score ($M = 4.31$, SD = 0.29). It means that students from De La Salle University, Manila, significantly differ from the other four schools in terms of perceiving love as Committed.

The chart in Fig. 2 shows De La Salle Lipa having the highest mean score for "Committed" subscale ($M = 4.53$, SD = 0.3). University of St. La Salle Bacolod follows, ($M = 4.5$, SD = 0.3), then La Salle University, Ozamis ($M = 4.48$, SD = 0.3), and lastly De La Salle University College of St. Benilde ($M = 4.4$, SD = 0.2).

The National Filipino Youth Survey (2002) described many young Filipino Catholics having a personalist or individualist faith. This means that the youth are religious, but they do not see the need to get involved in Church activities or community service. The Episcopal Commission on Youth (2003) adds that 44.9 % of young Filipino Catholics are nominal Catholics who seldom practice their faith, while about 3.8 % of the youth do not practice their faith at all. This is commonly true to those who live in urban areas. The research of Goodrich (2009) and Blazer (1985) points out that individualism is common to urban areas due to the emphasis on individual relationships, autonomy, personal goals, and individual competence. Such characteristics are observed among students in Metro Manila since they experience highly individualized learning and an autonomous and exclusive lifestyle that moves them to exercise their faith in an individualistic manner.

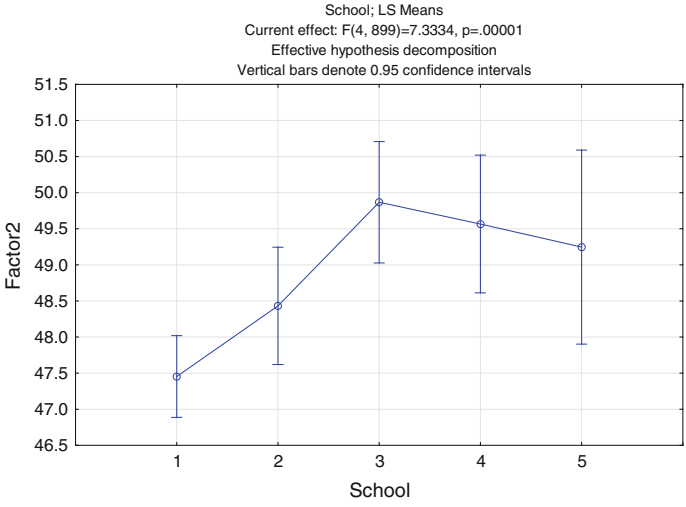


Fig. 2 Chart for Factor 2 Committed

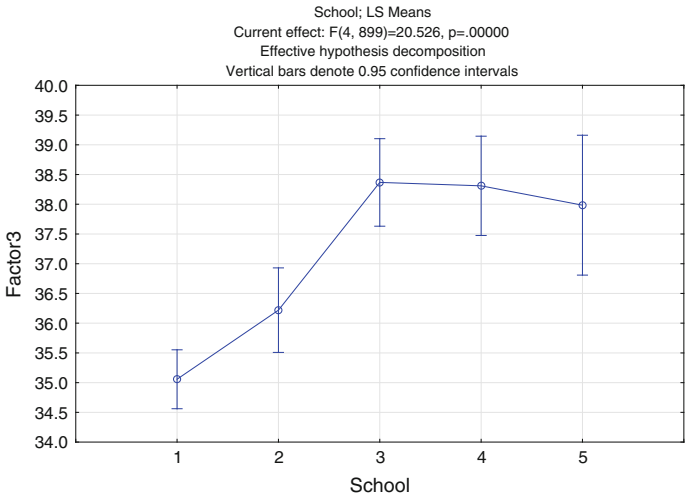


Fig. 3 Chart for Factor 3 Rooted

For the love as Rooted in God subscale, the post hoc test result of $F(4, 899) = 20.526$, $p = 0.000$ indicates that there are significant differences between schools. Table 2 shows that the mean scores among the schools are high. However, when compared to the schools from the provinces, De La Salle University, Manila ($M = 4.38$, SD = 0.12), and De La Salle University College of St. Benilde ($M = 4.53$, SD = 0.06) have lower mean scores. The chart on Fig. 3 shows the differences between the schools from Luzon and from those located in the southern part of the Philippines.

Data from the National Filipino Youth Survey or NFYS (2002) shows that the faith and religiosity of the Filipino indicate a very personal relationship with God. The NFYS also mentioned that even though the relationship of the youth with God is personal, it is also highly individualistic especially for those living in urban areas. The Episcopal Commission on Youth (2003) adds that the youth are affected by what is traditionally known as split-level Christianity. It means that faith matters little when it comes to their lifestyle.

The post hoc test result for love as Rooted in God subscale is in consonance with the results of the study of Marco and Monera in the year 2004. The researchers found out that the Lasallian youth from Metro Manila are less likely to possess better attitudes on faith life essentials. The post hoc test results and the previous study mentioned supports the low mean scores of students from Metro Manila regarding love as a response to one's faith.

Lastly, the post hoc test result on love as Integrated subscale $F(4, 899) = 3.0005$, $p = 0.0178$ shows the students of De La Salle University College of St. Benilde ($M = 4.36$, SD = 0.24), De La Salle Lipa ($M = 4.33$, SD = 0.35), University of St. La Salle Bacolod ($M = 4.33$, SD = 0.30) and La Salle University, Ozamis ($M = 4.36$, SD = 0.33), having similar perceptions on how love is expressed. On

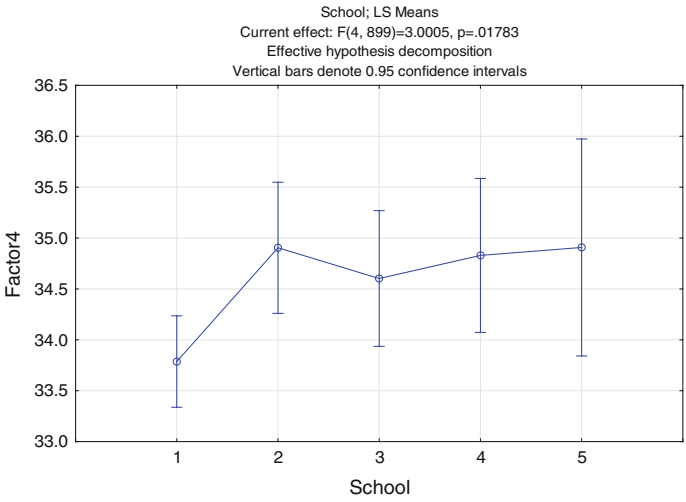


Fig. 4 Chart for Factor 4 Integrated

the other hand, De La Salle University, Manila (M = 4.22, SD = 0.34), had the lowest rating on love as Integrated among the five Lasallian institutions. Also, DLSU, Manila, showed significant difference on the Integrated subscale from the other schools. The chart on Fig. 4 shows the differences between the DLSU, Manila, and the rest.

The results of ANOVA test showed that there is significant difference among the Lasallian students' perception toward love. The post hoc tests provided specific information on the results of the LAS: (1) For the Engaged subscale, schools from Luzon got higher mean scores compared to the schools from Visayas and Mindanao. (2) In the Committed subscale, schools from Visayas and Mindanao got higher mean scores compared to those in Metro Manila. (3) The Rooted in God subscale had the same results with the Committed subscale. (4) Finally, it showed that there is significant difference among De La Salle University, Manila, students perception toward love as Integrated as opposed to students from De La Salle Lipa (DLSL), De La Salle University College of St. Benilde (CSB), University of St. La Salle Bacolod (USLS) and La Salle University, Ozamis (LSU), who have the same understanding on the matter.

5 Filipino Youth's Perceptions Toward Love: *t* test Results

t test is a statistical procedure used to decide whether two means are significantly different or not (Salkind 2009). This kind of procedure is appropriate in comparing the means of two groups especially in the analysis of the post-test two group experimental design.

Table 3 *t* test between gender

	Male		Female		*T*	Df	*P*
	M	SD	*M*	SD			
Engaged	57.87	4.92	59.05	4.65	3.60	746.63	0.000338*
Committed	47.81	5.77	49.00	5.43	3.11	743.62	0.001923*
Rooted	35.59	5.80	37.23	4.33	4.60	625.42	0.000005*
Integrated	58.86	4.71	58.39	4.18	0.77	712.61	0.441158

Note $*p < 0.05$

In this paper, the LAS was also analyzed using *t* test with two samples assuming equal variances (i.e., two samples from two population and can have different sizes). Since sample sizes in the LAS are large enough (i.e., $n > 15$), the sample size is normally distributed and can be analyzed using the *t* test.

In this section, the results of the comparison of the four factors between the identified antecedents, namely gender, family structure, and relationship background of the five selected schools will be discussed. Religion was not included anymore since 94 % of the student respondents are Christians.

In order to determine the factors that affect the beliefs of the young people on love, the researcher compared the responses between the male and the female. Also, the researcher would like ascertain if the young people's concept on love is influenced by their gender. Table 3 shows the results of *t* test with two samples assuming equal variance across gender in all five schools. There were 541 (59.8 %) female and 363 (40.2 %) male respondents to the LAS, and their average mean age is 18.2 years.

The statistical procedure revealed that there is significant difference on how males and females view of love. This is particularly true to Factor 1 Engaged wherein the males had a score of $M = 57.87$, SD = 4.92 as opposed to the females who scored $M = 59.05$, SD = 4.65. Regarding Factor 2 Committed, the males had a score of $M = 47.81$, SD = 5.77 while the females scored $M = 49.00$, SD = 5.43. On Factor 3 Rooted in God, the males had a score of $M = 35.59$, SD = 5.80 and the females scored $M = 37.23$, SD = 4.33.

The results of the *t* test show that the love perception of male and female has significant difference across regions. Female students have a higher mean score than the male students in Factor 1 Engaged, Factor 2 Committed, and Factor 3 Rooted in God. Studies about gender differences have shown that young women were more likely to have a heightened experience of love (Dion and Dion 1973), though they were often behind their partners in degree of involvement during courtship period (Rubin et al. 1981). Also, females fall in love more deeply (Dion and Dion 1973) and have a more realistic and pragmatic view of love (Lester 1985; Dion and Dion 1985; Morais and Tan 1980). In addition, females are more communicative in intimate relationships (Cancian 1987; Critelli et al. 1986; Komarovsky 1967). These research findings explain the reason for the high mean scores of the female respondents as opposed to the male respondents.

However, *t* test results for Factor 4 Integrated revealed that there is no significant difference on how males and females express their love. The males scored $M = 58.86$, SD = 4.71 while the females had $M = 58.39$, SD = 4.18. This result reflects the claim of Rubin et al. (1980), Levinger and Senn (1967), and Heiss (1991) that there are no important gender differences on how college students express their love to their significant others. Thus, male and female students have similar beliefs on how love can be expressed and becomes integrated in their lives.

The responses of the students across family structure were also compared to further shed light on the youth's beliefs on love. The researcher would like to discover whether the concept of love among the youth is influenced by circumstances like living with both parents or living with just a parent or guardian(s).

Table 4 shows the result of *t* test with two samples assuming equal variance between respondents living with both parents and respondents living with a parent or guardian(s). However, it must be noted that there are students who are living with a parent or their guardian(s) due to the following conditions: separation of the student from the family as a result of relocation (i.e., living in a condominium, apartment, or dormitory near the school), divorced or separated parents, or other family issues.

There are 521 respondents to the LAS who were living with both parents (57.6 %) while 383 (42.4 %) respondents were living with either a parent or with their guardian(s).

The *t* test results for all the factors among the respondents from the five schools show that there is no significant difference between students who are living with both parents and those who are living with one parent or their guardian(s). This result reflects the research findings of Lye et al. (1995), and Alino (2012) who discovered that the adolescents' positive well-being and views are directly affected not because of the residential or non-residential closeness with their parents but with the quality of their relationships with them. Also, adolescents who have warm relationship with their parents tend to have more mature way of understanding love (Alino 2012;). Hence, it is the quality of the relationship within the family and not proximity that affects a students' view on love.

Table 4 *t* test between family structure

	Living with both parents		Living with 1 parent or guardian		*T*	Df	*P*
	M	SD	*M*	SD			
Engaged	58.37	4.66	58.86	4.97	1.50	902	0.134263
Committed	48.77	5.73	48.34	5.41	1.13	902	0.259318
Rooted	36.77	4.82	36.29	5.31	1.41	902	0.157574
Integrated	34.23	4.4	34.59	4.3	1.23	902	0.220168

Note $*p < 0.05$

Table 5 *t* test between relationship structure

	With romantic relationship		No romantic relationship		*T*	Df	*P*
	M	SD	*M*	SD			
Engaged	58.94	4.17	57.93	5.70	2.80	516.58	0.005251*
Committed	48.79	5.38	48.04	5.96	1.92	902	0.054733
Rooted	36.64	5.09	36.44	4.95	0.57	902	0.566207
Integrated	34.74	4.27	33.75	4.56	3.26	902	0.001141*

Note $*p < 0.05$

The Engaged subscale has the highest mean score for those living with both parents ($M = 58.37$, SD = 4.66) while for those living with one parent or guardian(s) the score is $M = 58.86$, SD = 4.97. This shows that the students have a positive view on love as being engaged.

Adolescence is the time where teenagers deal a significant time of their lives in finding romantic relationship. Baber and Eccles (2003) adds, the quality of adolescent romantic relationships can have long lasting effects on their self-esteem and shape personal values regarding intimate relationships, romance, and sexuality.

In this study, the respondent's relationship background was utilized in order to gain a better understanding of how the youth perceive love. Table 5 shows the *t* test result between those who are currently in a romantic relationship or had a previous experience and those who have not experienced an exclusive–romantic relationship. The results revealed that the two groups have significant difference in the Engaged and Integrated subscales.

Barber and Eccles (2003), and Sorensen (2007) assert that healthy romantic relationships have many potential benefits for the youth. Romantic relationships can facilitate the process of youth gaining greater understanding of who they are and what they value (Sorensen 2007). Furthermore, in times when the adolescents become more autonomous from their parents and their romantic relationship increases, being with a partner becomes a source of emotional support. Also, a romantic relationship serves as a training ground for adolescents to develop interpersonal skills. Young people often refine their communication and discussion skills, develop empathy, and learn how to maintain intimate relationships (Barber and Eccles 2003). These research findings reflect the significant difference between young people who experienced romantic relationships and those who have not. Young people who have experienced romantic relationships may have discovered the other qualities of love and integrated differently from those who have not been in-love or experienced romantic love.

6 Conclusion

After analyzing the perceptions of the Filipino youth towards love based on the data gathered from the validated and factor analyzed LAS, the researcher concludes that:

1. The Filipino youth differ in their perception of love. However, they are in agreement that love is rooted in God: God wants us to be loving persons. The youth believes that we can manifest our love for God and to other people by being actively engaged to our family, friends, and neighbors. Love necessitates trust, openness, acceptance, care, loyalty, respect, giving, and honesty. The youth must first feel that they are accepted and appreciated before they can commit themselves to the community. However, the term "community" for the young still revolves among family and friends. As a result, attending a community service activity as a sign of love toward others is true for just a few young people. Furthermore, the youth view love as something that goes beyond the superficiality of saying "I love you" or being in a relationship just for the sake of being in one. Instead, love for the young people is expressed when we accept the weaknesses of other people forgiving their mistakes. Love, for them, surpasses externalities and ultimately becomes "*agapaic*" *or* authentic love.
2. There is significant difference between gender and relationship background among the Filipino youth. Female youth have a heightened experience of love, fall in love more deeply and have a more realistic and pragmatic view of love as opposed to the males. As for family structure, the study revealed that Filipino youth who live with both parents have the same attitude toward love from those who live with just one parent or their guardian(s). Hence, the quality of love among family members is more important than proximity.

Love is universal yet very personal. All men have a notion of love yet the meaning of love for every person is unique. This paper attempted to shed light on the "shared personal meaning" of love among Lasallian youth. Through the LAS administered to 904 Filipino students in Luzon, Visayas, and Mindanao, it was determined that love (according to the youth) has four factors: Engaged, Committed, Rooted in God, and Integrated.

Love as "Engaged" means being involved in the personal development of the beloved. One is also engaged when he or she is loyal, honest, respectful, supportive, and fair to other people. The second factor of love is "Committed" which is manifested through service. It involves community building and the practice of corporal and spiritual works of mercy toward other. For the Filipino youth, love is also "Rooted in God" wherein they acknowledge that the ultimate source of love is God. This love of God is manifested in the actions of Jesus Christ. Thus, Filipino youth believe that to follow the loving examples of Christ means acknowledging the love of God for man. The last factor of love is "Integrated." Love is demonstrated through external acts, but it also goes beyond physical expressions. Mercy and compassion for other people are such examples of integrated love.

This study on the perceptions of Filipino young people toward love concludes that:

1. The youth have a unique definition of love that is based on their constructs of love. Love, according to Lasallians, is engaged toward others, rooted in God, committed to service, and integrated with others.
2. The young people's perception of love corresponds to the Lasallian core values of faith, service, and communion. It means that the Lasallian core values have been integrated into the life experiences of the students.
3. The youth have a very positive view on love. They highly regard that God is the source of love. Because of this belief, they are inspired to commit themselves and manifest love toward family, friends, and neighbors.
4. The youth from Visayas and Mindanao have stronger beliefs regarding love as committed, rooted and integrated as compared to the Lasallians from Luzon. However, Lasallians from Luzon scored better on love as engaged compared to those coming from Visayas and Mindanao.
5. The gender of the Filipino young people affects their notions of love. Females tend to have a more intense experience of love compared to males.
6. Young people who have or had a romantic relationship strongly believe that love is engaged and integrated.
7. The study contradicts the result of research studies that portray the youth's notion of love as superficial.
8. The study will help the Religious Education, Campus Ministry and Service Learning personnel to better understand the young people. Religious educators as well as those who are involved in the crafting and implementation of service learning programs can benefit from the findings of this study with regard to the context of their learners.
9. A formation program must take careful consideration of the students' context and experiences in order for it to be effective.

This study also recommends the following:

For Religious Educators, Campus Ministers and People involved in Service Learning Programs or Activities

1. The results of the LAS show that the students have a strong belief on God as the source of love. Religious educators can leverage on this notion on love by providing relevant religious instruction, liturgical services, and service learning activities that will further deepen the Lasallian's love for God and move them toward Christian charity.
2. There should be conscious and deliberate collaboration among the campus ministry office, service learning office, and religious education department in tertiary institution. The collaboration is geared toward the creation and implementation of a holistic program that will contribute to the psychosocial and spiritual development of the young people. As revealed in the results of the

LAS, there is more urgency for tertiary institutions in Luzon to engage in this collaborative effort since their students are more individualized.

For the School

Create relevant and meaningful religious, campus ministry and service learning programs that would allow students to further practice love toward oneself and others.

For Parents and Educators

1. As shown in the results of this study, parents and educators can trust the youth regarding their notions on love. The four factors revealed that the youth have positive views on love.
2. Love as a notion is concretized in the eyes of the youth through the loving acts of their parents and teachers. Therefore, the notions of love among the youth (love as engaged, rooted in God, committed and integrated) should be reinforced in the home and in the school through the examples of adults.

References

Allport, G. (1929). The composition of political attitudes. *Journal of Sociology 1*, 220–238.

Alino, A. (2012). Filipino parent-adolescent relationship scale. *The International Journal of Research and Review, 9*, 59–70.

Amos, J., Hendrick, S., & Hendrick, C. (2001). Love attitudes: Similarities between parents and between parents and children. *Family Relations, 43*, 456–461.

Baber, B. and Eccles, J. (2003). Adolescent participation in organized activities. *For Indicator of Positive Development: Trends Child*.

Bain, T. (1930). Theory and measurement of attitudes, and opinion. *Journal of Psychology, 27*, 357–379.

Blazer, D., George, L., & Laderman, R. (1985). A rural and urban comparison. *Archives of General Psychiatry, 42*, 651–656.

Cancian, F. (1987). *Love in America*. New York: Cambridge University Press.

Chong, S., Shahrazad, W., & Mohd, S. (2012). Investigating the factor structure of the love attitude scale with Malaysian samples. *Asian Social Science, 8*(9), 608.

Critelli, J., Myers, E., & Loos, V. (1986). The components of love: Romantic attraction and sex role orientation. *Journal of Personality, 54*, 354–370.

De Mesa, J. & Cacho, R. (2011). *In love with God: doing theology for college students*. Manila: St. Scholastica's College.

Dion, K., & Dion, K. (1973). Correlates of romantic love. *Journal of Consulting and Clinical Psychology, 41*, 51–56.

Dion, K., & Dion, K. (1985). Personality, gender, and the phenomenology of romantic love. *Review of Personality and Social Psychology, 6*, 209–239.

Episcopal Commission on Youth-Catholic Bishops Conference of the Philippines (2002). The national filipino catholic youth survey. Manila: CBCP.

Episcopal Commission on Youth, CBCP (2003). The national filipino catholic youth survey. Manila.

Fisher, H. (2004). *Why we love: The nature and chemistry of romantic love*. USA: Holt.

Freudenburg, W. (2007). Urban-rural differences in environmental concern: A closer look. *Sociological Inquiry, 61*(2), 167–198.
Fromm, E. (2000). *The art of loving*. USA: Harper.
Goodrich, T. (2009). The relationship between individualistic, collective and transitional cultural value orientations and adolescents' autonomy and identity status. *Journal of Youth and Adolescence*, *39*, 882–893.
Hammock, G., & Richardson, D. (2011). Love attitudes and relationship experience. *The Journal of Social Psychology*, *151*(5), 608–624.
Hatfield, E. (1984). The dangers of intimacy. In V. J. Derlega (Ed.), *Communication, intimacy, and close relationships* (pp. 207–220). New York: Academic Press.
Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. USA: University of Denver.
Heiss, J. (1991). Gender and romantic-love roles. *The Sociological Quarterly, 32*, 575–591.
Hendrick, C., & Hendrick, S. (1986). A theory and method of love. *Journal of Personality and Social Psychology, 50*, 392–402.
Hendrick, S. S., Hendrick, C., & Reich, D. A. (2006). The brief sexual attitudes scale. *Journal of Sex Research, 43*(1), 76–86. Retrieved from http://dx.doi.org/10.1080/00224490609552301
Hindy, C. G., & Schwarz, J. C. (1984). *Individual differences in the tendency toward anxious romantic attachments.* Paper presented at the Second International Conference on Personal Relationships, Madison, WI. Retrieved from http://ccutrona.public.iastate.edu/psych592a/articles/Hazan_and_Shaver_1987.pdf
Hogg, M., & Vaughan, G. (2005). *Social psychology* (4th ed.). London: Prentice-Hall.
Iversen, G., & Norpoth, H. (1987). *Analysis of variance: Quantitative applications in the social sciences*. New York: Sage Publication.
John Paul II (2006). *Man and woman he created them: a theology of the body*. Translated by Michael Waldstein. Boston: Pauline Books and Media.
John Paul II (1993). *Love and responsibility*. San Francisco: Ignatius Press.
Judd, M. (1981). *Love and lifestyles*. Minnesota: Saint Mary's Press.
Katz, D., Allport, F. H., & Jenness, M. B. (1931). *Students' attitudes: A report of the Syracuse University reaction study*. Syracuse: Craftsman Press.
Komarovsky, M. (1967). *Blue-Collar Marriage*. New York: Random House.
Lasswell, T. E., & Lasswell, M. E. (1976). I love you but I'm not in love with you. *Journal of Marital and Family Therapy, 2*(3), 211–224.
Lee, J. A. (1973). *The colors of love: An exploration of the ways of loving*. Don Mills, Ontario: New Press.
Lester, D. (1985). Romantic attitudes toward love in men and women. *Psychological Reports, 56*, 662.
Levinger, G., and Senn, D. (1967). Disclosure of feelings in marriage. *Merrill-Palmer Quarterly*, *13*, 237–248.
Levy, M. B., & Davis, K. E. (1988). Lovestyles and attachments styles compared: Their relations to each other and to various relationship characteristics. *Journal of Social and Personal Relationships, 5*(4), 439–471.
Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22*, 5–55.
Lowry, R. (2008). *One way ANOVA—Independent samples*. Retrieved from http://vassarstats.net/
Lye, D., Klepinger, D., Hyle, P., & Nelson, A. (1995). Childhood living arrangements and adult children's relations with their parents. *Demography, 32*(2), 261–280.
Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*(4), 370–396. Retrieved from http://psychclassics.yorku.ca/Maslow/motivation.htm
Montgomery, M., & Sorell, G. (1997). Differences in love attitudes across family life stages. *Family Relations, 46*(1), 55–61.
Morais, R., & Tan, A. (1980). Male-female differences in conceptions of romantic love relationships. *Psychological Reports, 47*, 1221–1222.
Murphy, G., & Murphy, L. B. (1931). *Experimental social psychology*. New York: Harper.

National Service Learning of the National Youth Leadership Council of St Paul, Minnesota, USA. Retrieved from http://www.nylc.org/?gclid=CKz5gajz0bsCFaI34godNggAHg

Peck, S. M. (1978). *The road less travelled*. Great Britain: Cox & Wyman.

Rubin, Z. (1970). Measurement of romantic love. *Journal of Personality and Social Psychology*, *16*, 265–273. Retrieved from http://dx.doi.org/10.1037/h0029841

Rahner, K. (1967). The unity of love of God and love of neighbour. *Theology Digest 15*, Summer 1967, 87–93.

Rubin, Z. (1973). *Liking and loving: An invitation to social psychology*. New York: Holt, Rinehart & Winston.

Rubin, Z., Hill, C., & Peplau, L. (1981). Loving and leaving: Sex differences in romantic attachments. *Sex Roles, 7*, 821–835.

Rubin, Z., Hill, C., Peplau, L., & Dunkel-Schetter, C. (1980). Self-Disclosure in dating couples: Sex roles and the ethic of openness. *Journal of Marriage and the Family, 42*, 305–318.

Salkind, N. (2009). *Exploring research*, 7th edition. New Jersey: Pearson Education Inc.

Shaver, P., Hazan, C., & Bradshaw, D. (1986). Love as attachment: The integration of three behavioural systems. In R. Sternberg & M. Barnes (Eds.), *The anatomy of love*. New Haven, CT: Yale University Press.

Sorensen, S. (2007). Adolescent romantic relationship. USA: University of Rochester.

Springer, K. (2010). *Educational research: A contextual approach*. USA: John Wiley & Sons Inc.

Sternberg, R. J. (1986). A triangular theory of love. *Psychological Review*, *93*(2), 119–135. Retrieved from doi:10.1037/0033-295X.93.2

Sternberg, R. J. (1998). *Love is a story: A new theory of relationships*. USA: Oxford University Press.

Sternberg, R. (2011). *Cognitive psychology*. USA: Wadsworth Cengage Learning.

Stevens, J. (1999). *Intermediate statistics: A modern approach*. New Jersey: Lawrence Earlbaum Associates Publishers.

Stoodley, B. (1959). Normative attitudes of Filipino youth compared with German and American youth. *American Sociological Review, 22*(5), 553–561.

Thurstone, L. (1946). The measurement of attitude. *American Journal of Sociology 52*: 39–50.

Tzeng, O. C. S. (1993). *Measurement of love and intimate relations: Theories, scales and applications for love development, maintenance and dissolution*. Westport, CT: Praeger.

Underwood, L. (2009). *The science of compassionate love: Theory, research and applications*. Wiley-Blackwell: Wiley.

Vetter, G. B. (1930). The measurement of social and political attitudes and the related personality factors. *Social Psychology 25*, 49–89.

An Approach for the Assessment and Allocation of Individual Grades for Group-Based, Collaborative Work by Combining Self-assessment, Group-Evaluation and Wiki Data: Results from an Action Research

Simon John Williams

Abstract Increasingly, project-based and group work is being utilised in our classrooms. In order to keep pace with this shift in pedagogy, the assessment process needs to be updated so that it complements the tasks learners are being expected to complete. Project work is complex and requires a number of learners working together in order to complete a task. While this allows for active learning, it can also promote free riders (learners who do not contribute). Simply allocating a summative grade at the end of the project is no longer sufficient, as does not take into account the process and new skills the learners gain along the way. A two-cycle action research approach was taken in order to discover how individual grades can be assigned to collaborative group work based online and in the classroom. The researcher, also the teacher, used his class in a private university in Malaysia to investigate and develop an approach for assessment. The study resulted in an approach for the assessment of collaborative work using multiple sources of data including self-assessment, group-evaluation all of which was facilitated by a Wiki.

Keywords Collaborative assessment framework · Alternative assessment · Project-based learning · Allocation of individual grades

1 Background

Project-based learning (PBL) can provide opportunities for learners to engage in active learning as well as the opportunity to learn new soft skills such as collaboration, communication and negotiation. Nevertheless, PBL can also create stress and disagreement in regards to grade allocation. While observing my class and

S.J. Williams (✉)
Taylor's University, Subang Jaya, Malaysia
e-mail: Simonjohn.williams@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_29

engaging with them in PBL, I noticed that assigning a single grade for a group did not take into consideration the notion that not all members contributed equally. This caused the learners in my class to lose motivation to complete the project, as they were aware that certain members would get a grade for simply standing on the sideline. Therefore, concerned by this, I set out to study how a single grade could be allocated for each group member.

2 Literature Review

The benefits of group work, collaborative and cooperative learning have been well theorised and researched (Vygotsky 1978; Slavin 1991, 1996; Gokhale 1995; Li 2002). A review of the literature by Lai (2011) claims that collaboration can have positive effects on student learning, for instance increasing critical thinking (Gokhale 1995), informative retention (Johnson and Johnson 1986), and the opportunity to become active learners (Hew and Cheung 2008).

Since then, to further develop this type of learning methodology, task-, project-, and problem-based learning have attempted to give teachers a framework in order to create effective group learning environments, which follow a pedagogical framework. The work of Willis (1996) uses an approach whereby students are put in groups to complete a pre-designed task, which can be done collaboratively. However, any teacher attempting to implement this in their classroom may not only be faced with the challenge of using a new methodology, but trying to design a system to assess it (Elliott 2008).

A number of possible issues could arise when a collaborative project is created; the first being free riders (James et al. 2002): a name given to a group member who does not equally contribute to the overall group process. The free rider, then, may create a frustrating and stressful environment for other members of the group, as they are left to take on the extra workload (Strauss and Alice 2007). And due to this, a single grade awarded to the group would not accurately reflect the work contributed by each member (Gibbs 2009).

2.1 Alternative Assessment

To address concerns with grading collaborative work, alternative assessment procedures, such as self-assessment (SA) and group-evaluations (GE), provide the teacher with additional information on the group process. Group work may take a number of weeks to complete, so much of it is often conducted outside of the classroom, leaving the learners to be the only ones who know who is contributing to the project (Race 2001). One problem, though, with the use of SA and GE is the question over its validity (Hughes 2001; Kennedy 2005), due to the finding by Boud and Falchikov (1989) concerning over- and underrating: the former refers to

learners who give themselves higher grades than they may deserve, whereas the latter refers to learners who may give themselves lower grades than they may deserve. This finding has also been substantiated by later studies that find that weaker learners tend to overrate while stronger learners tend to underrate (Lejik and Wyvill 2001; Wilmot and Crawford 2005). Even though questions concerning over- and underrating have been raised, research has shown that these alternative forms of assessment can be reliable and valid. Nevertheless, convincing teachers and education administrators may prove to be the true challenge, unless further corroborating evidence and a clear approach can be found.

Assessing collaborative group work can be complex. This is due, in part, to the number of variables that are attributed with it and the sheer amount of systems which have been tested in order to find an effective model. A number of studies have investigated collaborative assessment (Brooks and Ammons 2003; Hartford 2005; Gibbs 2009; Exley 2010; Caple and Bogle 2011; De Wever et al. 2011) and in each study, different methods were used in order to assess collaborative work, such as reflections, SA, GE, teacher observations, and Wiki history logs. Many of the studies used quantitative or qualitative methods with none of the aforementioned studies taking an action research (AR) approach. Therefore, an AR approach may allow the researcher to investigate the complexities of collaborative assessment from the point of view of a teacher and his students.

2.2 Research Questions

1. How can individual grades be allocated for each student working collaboratively within a group?
2. What evidence can teachers collect in order to substantiate individual grades when segments of the work are completed outside of the classroom and/or online?

3 Methodology

This study adapted an action research framework from Kemmis and McTaggart (2005). The model uses a two-cycle method, whereby once the first cycle is completed, data are analysed and reviewed. Then, suggestions are made and the second cycle commences. When discussing action research, Kemmis and McTaggart say that "*…action research involves the investigation of actual practices and not abstract practices*" (2005, p. 277). McNiff (2002) also goes on to state that empirical researchers carry out research on others whereas action researchers carry out research on themselves (Fig. 1).

There are six stages to the framework in total. Stage one asks the researcher to identify what needs to be researched. Stage two requires the researcher to plan an

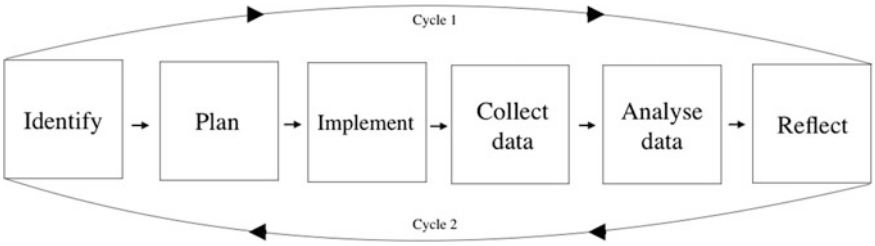


Fig. 1 Action research framework

intervention, which is subsequently implemented in stage three. Data is collected in stage four in the form of observations, questionnaires or live data from online tools. The data are then analysed in stage five with the final stage remaining for reflection.

3.1 Data Collection

Data were collected in the form of observations from the researcher (also the teacher). A questionnaire was adapted from an instrument by Fisher et al. (2005), which attempts to gauge students' perceptions of assessment. This was given to the participants at the end of each cycle to ascertain their satisfaction and feedback with the assessment system. The researcher reflected upon the completion of each cycle. To critically assess these reflections, the study established a "Critical Friend" as suggested by McNiff (2002) to provide someone to act as a critical conduit for reflective discussion. In addition to this data, the researcher also collected and tabulated the results from self-assessments, group-evaluations and Wiki logs. Each data point provided the study with a method for triangulation.

3.2 Participants, Course and Location

Ten participants in total joined both cycles of the research. The participants were aged between 17 and 24 and came from Vietnam (1), Korea (2), China (3) Kazakhstan (1), Malaysia (2), and Yemen (1). Each participant had come to Malaysia to study English before entering into the main university foundation and degree programmes. The study was conducted at a private university in Malaysia during a course that was six weeks in duration. Participants attended the course each day from 8 am to 1 pm and studied aspects related to English language learning such as writing, speaking, reading and listening.

3.3 The Project

The teacher employed a project-based approach to learning, and the participants were required to complete two projects in groups: the first was one week in length, and the second two weeks. The first project centred on culture, whereby each group had to choose a country that they had no prior knowledge of and then research. In groups, the information was then compiled on to a Wiki page. In the second project, the participants were asked to redesign their classroom. All information, such as meeting minutes, task lists, pictures and designs, was also uploaded and shared on a Wiki. The project was hosted on a Wiki website, which allowed the students to collaborate both in and outside the classroom. Each student signed into wikispaces.com and created an account and user name. Each time a student signed in and saved work, the Wiki recorded this and made the data available to the Wiki administrator via Wiki logs. Wikis form the backbone of this study as they provide a learning environment that is online and can facilitate the collaborative process by allowing users to display information, collaborate, monitor progress and add feedback (for a more in-depth discussion see Williams 2014). The log feature is especially useful as it allows group members and the teacher to track who contributes what by time stamping each occasion a user logs in with their individual user I.D. (Caple and Bogle 2011).

3.4 Threats to Validity

As a purely qualitative study with a researcher who adopts a constructivist mindset to research, positivist notions of validity do not necessarily apply to a study such as this (Maxwell 2013). Maxwell (2013) affirms that to a study that contemplates the existence of multiple realities, a more accurate concept of validity would be one that seeks to validate conclusions with careful attention paid to researcher bias and reactivity. As the teacher and researcher of his own practice, the study is biased by the simple fact that results cannot be corroborated by an external control group. Therefore, the placement of a critical friend is vital as it provides an impartial viewpoint to the interpretations made by the researcher. Finally, to address issues of reactivity as the notion that the researcher can influence his participants, a number of data points are correlated to provide rich data from a number of sources.

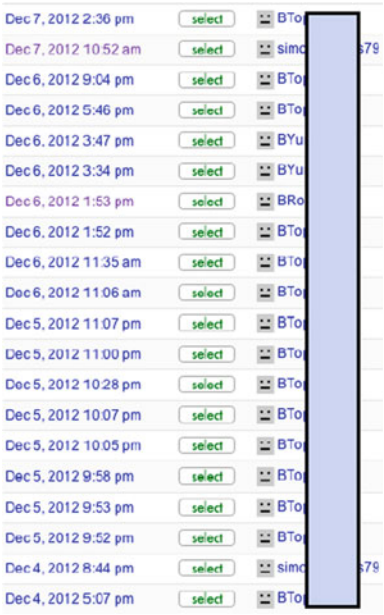
4 Data Analysis: Cycle 1—Implementation

After researching alternative assessment techniques, it was decided that self-assessment (SA) and group-evaluation (GE) would be incorporated into the assessment process (Fig. 2). The SA (Appendix 1) contains six assessed areas: research, sharing, completing tasks, contribution, listening to others and cooperation. The aim of these items was to show the participants that the process was of importance,

and therefore, would be assessed alongside the final product. The GE (Appendix 2), then, contained two simple questions: (1) "*Who worked hard in your group and why did you choose this person*?" (2) "*Who needs to work harder next time, and why*?" In line with research on over- and underrating, the use of a GE (Appendix 2) provided the researcher with corroborating data on who was contributing and who was not. Moreover, GE allows the participants to voice what occurred during the group process. As this is largely achieved out of the sight of the teacher, the GE provides a glimpse into the inner workings of the group. The final product is subsequently assessed solely by the teacher where marks are awarded for content, structure, grammar and vocabulary. The weightage was split with 25 % (process) of the grade decided by the students via the SA, and the remaining 75 % (product) decided by the teacher.

4.1 Data Analysis: Cycle One

Cycle one proposed a number of interventions in order to assess the group collaborative process. With the introduction of the self-assessment, a group-evaluation was also implemented in order to corroborate results (Fig. 2). The results were tabulated in order for the researcher/teacher to note any patterns or inconsistencies. Each student was able to express their views on the inner workings of the group, even with their limited English. For example, group members noted that participant 5E "*need to work more because she has (sic) a difficult time then*". This is corroborated by participant 5E herself who wrote that she needed to work harder next time "*because she must spend more time to (sic) group work*."



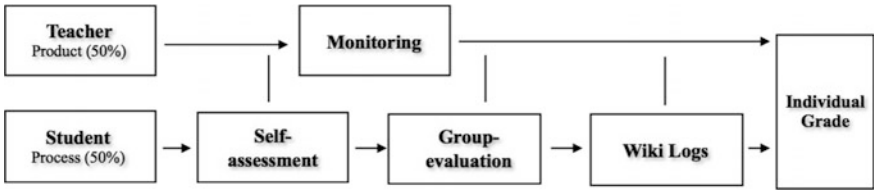


Fig. 2 Framework for the assessment of individual grades during project-based work

Contrary to the previous comments, some participants saw the group process differently from other group members. Participant 1A commented by stating that: "*all of us worked hard*". However, this statement was challenged by the remaining group members who claimed that participant 1A "*had (sic) absent our group meeting for two times*," and "*sometime she forgot the deadline to finish the job*." Aside from conflicting comments on contribution, some participants did indeed have their efforts recognised. Group members noted that participant 6F "*…did the most work*" and "*… he checked the project many times*."

In addition to the information noted in the group-evaluations, the researcher also turned to the Wiki logs in order to seek further confirmation of contribution to the group process. The figure on the left is a screen shot taken from one of the group's Wiki pages. The illustration shows students' Wiki login names, as well as the researcher (simonwilliams79). It is clear from the log that BTop31 had made the most contributions to the Wiki page, whereas Byumi31 and Bron31 made only few contributions overall. The logs were also further corroborated by the group-evaluation where BYumi31 noted that "*I didn't help top. Next time we need more help (sic) each other*."

4.2 Cycle Two: Identify and Implement

The data from cycle one provided the researcher with evidence that the students were able to self-assess and evaluate other group members for their contribution. Moreover, the research obtained corroborating data from the Wiki logs to support some of the assertions made by the students. For cycle two, though, a number of new initiatives were implemented: (1) The SA and GE were administered during and at the end of the project. (2) The GE included a section where the project

members could quantify how much work was completed by each member. (3) After a conversation with the critical friend on subjectivity of the assessment due to it being created solely by the teacher, it was suggested that the assessment rubrics should be co-created with the students involved in the process.

5 Discussion

Through the course of the research, a number of initiatives were implemented in order to investigate methods in which to assign individual grades for group work. Firstly, a self-assessment and group-evaluation were implemented. In cycle one, the SA and GE were administered once; however, in the second cycle, they were administered twice for a number of reasons: (1) increased training in the use of SA and GE, and multiple uses have been known to improve reliability (Ross 2006; Chang et al. 2013) (2) Administering the assessment twice allows learners to understand the importance of it for their learning. Students who then completed the assessment twice, and were less involved previously, were observed to be more engaged in the project afterwards. This also provides the teacher with valuable information as to who is working and who is not. No definitive conclusions can be made to explain this result, but it is possible that by completing the first round of assessment the learners realised first-hand how it would affect the outcome of their final grade (see Williams and Lim 2015 for further discussion). Therefore, by training and administering the assessment twice, the learners could be more informed as to what was expected of them, thereby reducing any mystery surrounding the assessment process. It then means that the teacher can make early interventions to discover why the group process is not functioning well and hopefully reduce any frustration caused by free riders as reported by Strauss and Alice (2007).

Wikis were a valuable tool in this study, as they allowed the teacher/researcher to be a silent partner in the group process. Through the use of the logs, which provided an extra data source on who was doing what and when, the teacher was able to intervene when groups were falling behind, or when some group members were not contributing. This finding is substantiated by Caple and Bogle (2011), who also discovered that Wiki logs could help to make group work grading fairer. Though the Wiki logs do not provide conclusive proof to a learner's contribution, they do offer an additional data point so that the teacher can be more informed as to what is happening within the group.

There is, though, a limitation that the research identified later on. In both cycles, one and two, the SA was worth only 25 % of the final grade, with the remaining 75 % left to the teacher to decide. On one occasion, it was observed that a participant received 0 from a possible 5 points for her/his SA due to low contribution. Nevertheless, as the remaining group members received 12 points out of a possible

15 for their group project, plus an additional 3.5 points for their SA bringing their total score to 15.5. Because of the 75/25 % teacher/student split, the free rider still received a satisfactory score of 12/20, while the participant's other team members received 15.5/20. Meaning that the free rider was awarded a passing grade for doing very little. In the second cycle, the weightage was shifted to 50 % process and 50 % product, which rectified the limitation slightly.

6 Conclusions

The study set out to investigate two research questions: (1) How can individual grades be allocated for each student working collaboratively within a group? (2) What evidence can teachers collect in order to substantiate individual grades when segments of the work are completed outside of the classroom and/or online? To address question one, the study created a system that assessed both process and product with 50 % if the grade coming from a self-assessment score of the process, and the remaining 50 % coming from an assessment of the final product by the teacher. These grades were then combined to get the final grade.

Question two, then, sought to find evidence to corroborate the scores from the SA. To do this, the study created a group-evaluation, which asked each participant to state who had worked hard, and who needed to work harder. This, plus information from Wiki logs, helped to provide the teacher with more information on the group process, as well as corroborating grades from each self-assessment.

Assessment by its very nature is subjective, and there may not be a flawless method in which to do it. What teachers can do, though, is to collect more information on the process of the project and use a number of assessment measures such as self-assessment, peer-evaluation and online tools. By collecting data from various sources and participants, the teacher can paint a more detailed picture of the group process which, in turn, can allow the teacher to make more informed decisions when it comes to assigning grades.

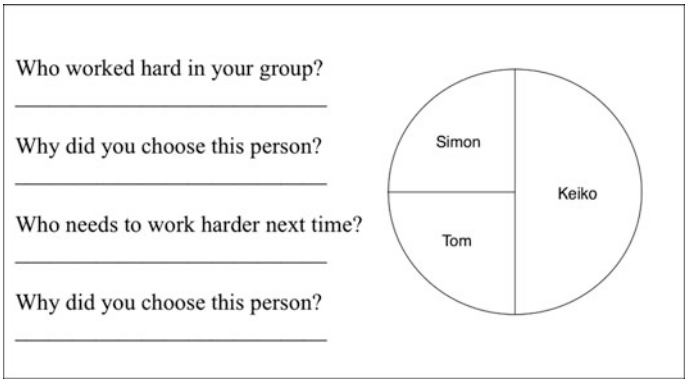
Finally, the group process is a complex one, as it requires a number of people collaborating together in order to produce a piece of work. This process can be made more effective and less stressful by creating a transparent assessment system whereby the stakeholders (the learners) can voice how they think they should be assessed. Not only does this provide a fairer system of grading, it also allows the learners to become more aware of the assessment process and what is actually expected of them.

This study was conducted in a small class setting; therefore, further research could investigate a large scale implementation of this process over different subject areas to test its effectiveness and acceptance by both learner and teacher.

Appendix 1: *Self-assessment Rubrics*

	Very Good	Good	Developing
Research and collecting information 6	I collected *lots of* information from various places, such as books, the internet etc.	I collected *some* information from a few places.	I only collected *a little* information from few places.
Sharing 8	I *always* shared my information or ideas with *all* my team members.	I *sometimes* shared information or ideas with my team members.	I shared *little* information or ideas with my team members.
Completing tasks 8	I met *all* deadlines and I was not late for meetings or to complete work.	I met *most* deadlines and was only late for *some* meetings and to complete work.	I missed *many* deadlines and was *often* late for meetings or to complete work.
Contribution 8	I *always* helped *every* team member with all tasks, such as gathering information, editing work.	I helped *some* of my team members, but not all to gather information and edit work.	I *didn't* help my team mates to gather information, edit work etc.
Listening to other group members 5	I *always* listened to the ideas and suggestions from my team.	I *sometimes* listened to ideas and suggestions from my team.	I *didn't* listen to my other team members. I often did it my own way.
Co-operating with my team 5	I *never* argued with my team members. I *always* talked about ideas and got everyone's opinion.	I *sometimes* argued with my team. I *sometimes* talked about ideas and thought about some opinions.	I *often* argued with my team mates. I *never* listened to their ideas and didn't think about their opinions.

Appendix 2: *Group-Evaluation Form*



References

Boud, D., & Falchikov, N. (1989). The role of self-assessment in student grading. *Assessment and Evaluation in Higher Education, 15*(1), 101–111.

Brooks, C. M., & Ammons, J. L. (2003). Free riding in group projects and the effects of timing, frequency and specificity of criteria in peer assessments. *Journal of Education for Business, 78* (5), 268–272.

Caple, H., & Bogle, M. (2011). Making group assessment transparent: What Wikis can contribute to collaborative projects. *Assessment & Evaluation in Higher Education., 38*(2), 198–210.

Chang, C., Liang, C., & Chen, Y. H. (2013). Is learner self-assessment reliable and valid in a web-based portfolio environment for high school students? *Computers & Education, 60*(1), 325–334.

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2011). Assessing collaboration in a Wiki: The reliability of university students' peer assessment. *Internet and Higher Education, 14*(4), 201–206.

Elliott, B. (2008). *Online collaborative assessment*. Scottish Qualifications Authority

Exley, K. (2010). *Managing and assessing students working in groups*. Centre for Education Development. Queen's University. Belfast. Accessed at http://www.qub.ac.uk/directorates/AcademicStudentAffairs/CentreforEducationalDevelopmentFilestoreDONOTDELETE/Filetoupload,196874,en.pdf

Fisher, D. L., Waldrip, B. G., & Dorman, J. P. (2005). *Student perceptions of assessment: Development and validation of a questionnaire*. Paper Presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Gibbs, G. (2009). *The assessment of group work: Lessons from the literature. Assessment standards knowledge exchange*. Oxford, Centre for Excellence in Teaching and Learning in Higher Education, The business School, Oxford Brookes University. Accessed at: https://www.brookes.ac.uk/aske/documents/Brookes%20groupwork%20Gibbs%20Dec%2009.pdf

Gokhale, A. A. (1995) Collaborative learning enhances critical thinking. *Journal of Technology Education 7*.

Hartford, T. (2005). *Facilitation and assessment of group work using web-based tools*. BEE-j 5.

Hew, K. F., & Cheung, W. S. (2008). Attracting student participation in asynchronous online discussions: a case study of peer facilitation. *Computers & Education, 51*(3), 1111–1124.

Hughes, I. E. (2001). But isn't this what you're paid for? The pros and cons of peer and self-assessment. *Planet, 2*, 20–23.

James, R., McInnes, C., & Devlin, M. (2002). *Assessing learning in Australian Universities: Ideas, strategies and resources for quality in student assessment*. Melbourne: Centre for the study of Higher Education, University of Melbourne.

Johnson, R. T., & Johnson, D. W. (1986). Action research: Cooperative learning in the science classroom. *Science and Children, 24*, 31–32.

Kemmis, S., & McTaggart, R. (2005). *Participatory action research: Communicative action and the public sphere*. Thousand Oaks, CA: Sage Publications.

Kennedy, G. J. (2005). Peer assessment in group projects: Is it worth it? *Australian Computing Education Conference on Research in Practice in Information Technology, Australia* (Vol. 42).

Lai, E. R. (2011). *Collaboration: A literature review*. Pearson.

Lejik, M., & Wyvill, M. (2001). The effect of the inclusion of self-assessment with peer assessment of contributions to a group project: A quantitative study of secret and agreed assessments. *Assessment and Evaluation in Higher Education, 26*(6), 551–561.

Li, Q. (2002). Exploration of collaborative learning and communication in an educational environment using computer-mediated communication. *Journal of Research on Technology in Education, 34*(4), 503–516.

Maxwell, J. A. (2013). *Qualitative Research Design* (3rd Ed.). Thousand Oaks, CA London, UK: Sage Publications.

McNiff, J. (2002). Action research for professional development: Conciseadvice for new action researchers. Jean McNiff. 3.

Race, P. (2001). *A briefing on self, peer and group assessment*. Assessment Series No. 9: LTSN Generic Centre. Retrieved from http://internt.iha.dk/paedagogik/seminarer/Chris%20Rust/ASS009PhilRa

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment Research & Evaluation, 11*(10).

Slavin, R. E. (1991). *Student team learning: A practical guide to cooperative learning*. Washington, DC: National Education Association.

Slavin, R. E. (1996). Research for the future—Research on cooperative learning and achievement: What we know, What we need to know. *Contemporary Educational Psychology, 21*, 43–69.

Strauss, P., & Alice, U. (2007). Group assessment: Dilemmas facing lecturers in multicultural tertiary classrooms. *Higher Education Research & Development, 26*(2), 147–161.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.

Williams, S. J., & Lim, C.H. (2015). Holistic assessment: Creating assessment with students. In S. F. Tang & L. Logonnathan (Eds.), *Taylor's 7th q Teaching and Learning Conference Proceedings* (SE 36, pp. 389–397). Springer, Singapore. Accessed at: http://www.springer.com/education+%26+language/learning+%26+instruction/book/978-981-287-398-9

Williams, S. J. (2014). Using Wikis to carry out project-based learning. *Journal of Interdisciplinary Research in Education (JIRE), 4*(1), 1–10.

Willis, J. (1996). *A framework for task-based learning*. Addison Wesley Longman Limited: Canale.

Wilmot, P., & Crawford, A. (2005). Validating the assessment of individualswithin undergraduate teams. In *Proceedings of International Conference on Engineering Educations*, Gliwice, Poland. Accessed: http://webpa.lboro.ac.uk/tutors/support/why_use/icee2005_pa_paper.pdf

Students' Pre-assessment Analysis to Improve Academic Performance and Thinking Skills

Seng Yue Wong and Wee Jing Tee

Abstract The student-centered focus of graduate capabilities which demands application of specialized knowledge and skills targeted to meeting student-specific needs warrants that the teaching and learning strategies used for graduates must empower with the ability to use higher-level academic skills and critical thinking skills effectively in the future. This study is conducted to assess and evaluate students' feedback to determine the student success factors in their learning, in order to improve their academic performance and thinking skills on their overall learning experience. A student pre-assessment survey analysis is used to provide a graphical representation of student success profile, which covered five main categories: academic skills development, study and thinking skills, personal issues, planning for the future, and resource needs. This survey is carried out among 18 students from foundation course at School of Computing and Information Technology, Taylor's University. The results have shown that conflict resolution, study skills, and time management are three factors that affect students' success in their learning. This analysis and findings of the survey are used to provide useful information to prepare students toward their success via effective teaching and learning strategies with their graduate capabilities in the future.

Keywords Pre-assessment · Critical thinking skills · Academic skills · Graduate capabilities

S.Y. Wong (✉) · W.J. Tee
School of Computing and IT, Taylor's University, Lakeside Campus, Subang Jaya, Malaysia
e-mail: SengYue.Wong@taylors.edu.my

W.J. Tee
e-mail: WeeJing.Tee@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_30

1 Introduction

The momentous shift of the practice paradigm from teacher-centered to student-centered that emphasizes students' direct involvement in acquire knowledge and their graduate capabilities in the future. Graduate capabilities are critical factors in the adoption of curriculum for the training of students. Graduates are expected to be highly skilled at the provision of effective and efficient services or knowledge in collaboration with other professionals. Critical thinking and problem solving, communication, cultural adaption, creativity and innovation, and collaboration are the five factors that should be embedded in graduates after they finish their studies. Critical thinking harnesses the inbuilt brainpower to have a deep understanding of a concept, identify and analyze the relationship between the various components within a concept, apply the concept, and assess the impact of the application in the real world of practice (Yusuff 2015).

Recent education changes emphasis on developing skills and attitudes that are needed for lifelong learning. Lifelong learning is one of the graduate capabilities that should be equipped with. Students are able to locate, extract, synthesize, and utilize information effectively and relate their learned knowledge to their daily life. Active and collaborative learning is utilized to complement lectures in many classrooms. Students are encouraged to take responsibility and play a greater role in their learning. Students learn concepts and skills in greater depth, but the time-intensive nature of the inquiry-based approach often means that less content is covered in the classroom, and some contents are came solely from textbooks and readings. There is no conversation in lecture-based learning and most online tutorials. The information flow is one way, from the teacher to the student, until the time of quiz at the point neither instruction nor the learning can be modified for improvement (Broussard 2014). Therefore, students are required to be active participants throughout the learning process which will improve student performance, engagement, and motivation (Broussard 2014).

The concept of education and teaching differs from what it has been during the past decades. After teaching and learning processes, students' roles are to receive the information and store them in their memory and apply them in their future life. The academic achievement is of great importance that makes the individual feels the excellence and success, improves self-confidence, and raises ambition level, whereas success makes the individual feel proud and confident of his potentials and abilities and his ability to succeed and make achievement and such impact is a mutual one (El-Daw and Hammoud 2014). Students can be taught thinking skills that allow them to process information efficiently, think productively, and generate and evaluate ideas (Kobzeva 2015).

Therefore, there is a need to conduct student-led assessment and pre-assessment analysis to enable students to understand and apply quality criteria to their work, freeing them from teacher dependence, including graduate capabilities. This paper will present a student pre-assessment analysis to determine students' success factors in their learning. From the results of analysis, we will present a graphical

representation of student success profile, which covered up academic skills, study and thinking skills, personal issues, planning for the future, and resource needs. Finally, we hope this survey can contribute to improve students' academic performance and thinking skills in the assessment field.

2 Literature Review

Pre-assessment analysis is a tool to measure students' capability prior to instruction and help to assist teachers effectively match instruction with the needs of the students. This covers teacher's decisions about teaching and learning content, pacing, learning materials, grouping activities, and specific learning activities. Pre-assessment results enable students to engage in their studies which suit their interest, learning styles, and prior knowledge. Students are challenged at an appropriate level based on their prior knowledge and skills. Indirectly, students are more likely to demonstrate continuous progress and growth, including academic performance and thinking skills.

Academic skills development is one of the main points in students' pre-assessment survey. Reading, writing, speaking, and listening skills are four main academic skills need to be developed well to ensure students success when they graduate later. Thinking is the ability to attack a problem or task and solve it rationally and then provide a reasonable explanation for the solution (Ngang et al. 2014). By developing student's thinking skills, student can make achievements; can become successful; can shine in social life; and can attain emotional, social, and economic maturity (Ngang et al. 2014). Thinking skills assist a person to select the ideal strategy to solve the problem that is being faced. Some researchers also suggested that thinking skills should not be taught separately but be incorporated in the teaching and learning processes in subjects or programmes (Ngang et al. 2014).

Besides that, problem-solving skills need to be emphasized in education system, thus producing intellectual citizens and enabling our country to achieve developed nation status in the world. Researchers also emphasized that intellectual development among students is necessary in order to get them used to the existing realities, enabling them to prepare themselves for their future life and achieving success (Ngang et al. 2014).

Self-esteem is defined as a belief and self-confidence in your own ability and value. Self-esteem is how you feel toward yourself. This feeling about yourself is come from convictions about yourself as a capable, competent person having worth. Feeling capable of worth means having self-respect, viewing yourself as able to scope effectively with life's challenges (Kobzeva 2015). Individuals who have high self-esteem generally demonstrate a high degree of acceptance of themselves and others and environment. Social skills will be affected by self-esteem, and the more the students feel self-esteem, the more they feel enjoy social skills to satisfy their basic needs (El-Daw and Hammoud 2015).

As mentioned above, the pre-assessment survey is useful to determine flexible groupings for students' learning activities and to engage underachieving students using their interests, learning styles, and prior experience. Other studies also identified that numerous personality traits will collate with self-efficacy, which includes self-esteem, optimism, health, as well as academic and career performance (Azizli et al. 2015).

3 Methodology

The student pre-assessment survey is conducted in School of Computing and Information Technology, Taylor's University. There are 18 foundation course students from this school involved in this survey. Sample size ranges from ten to twenty are allowed for this type of survey although sample size ranges between 30 and 500 are suitable for most studies (Sekaran 2000). Since the population for this school is small, 18 foundation course students are good enough to represent the school for this survey. This pre-assessment survey is also known as 'student success kit' (Ellis 1998), which contains 144 items under five main categories: academic skills development, study and thinking skills, personal issues, planning for the future, and resource needs. Each subcategory contains 8 items.

Reading, writing, speaking, and listening skills will be tested under academic skills development. Study and thinking skills cover four subcategories which are learning styles, memory, study skills, and creative and critical thinking skills. Personal issues include five subcategories too: motivation, self-esteem, personal relationships, conflict resolution, and health. On the other hand, time management, money management, personal purpose, and career planning are covered under planning for the future category. However, resource needs only cover one subcategory, support resources.

Five-point scale is utilized to score each of the items on the questionnaire. The details about five-point scale are shown as below:

5 The statement is always or almost always true for you (around 100 %).
4 The statement is often true for you (around 75 %).
3 The statement is sometimes true for you (around 50 %).
2 The statement is rarely true for you (around 25 %).
1 The statement is never or almost never true for you (around 0 %).

There are 8 items for all 18 subcategories, and the highest possible score for each subcategory is 40. If scores have shown below 25, the subcategory is probable trouble areas for students' learning. Potential areas for improvement for those scored between 25 and 30. Scoring 30–40 indicates probable success areas for student success.

4 Results and Discussion

The pre-assessment survey is one of the methods to measure students' capabilities directly. It is not enough to assess students' knowledge and skills at the end of the course or program if we want to gauge how much students have learned. This explains why we need to find out what students know coming in so that we can identify more specifically the knowledge and skills they have gained during the course or program. Pre-assessment results provide a way for teacher to gather key information (probable trouble areas, potential areas for improvement, and probable success areas) about what students know and are able to do prior to instruction, as well as what tasks to do.

Pre-assessments analysis results are shown in Table 1. Table 1 shows the amount of students who follow the subcategory under three types of scores: below 25, 25–30, and 30–40. From Table 1, we noticed that 15 of 18 students agree that creative and critical thinking skills are probable success areas. There are 7 of 18 students who agree that reading is probable trouble areas for students' learning. Graph bar to represent the pre-assessment results also is illustrated in Fig. 1. The overall results have indicated that creative and critical thinking skills, learning styles, personal purpose, and carrier planning are four probable success areas, which achieve at least 75 % students to score at 30–40. Study skills, conflict resolution, and time management are three probable areas that need to be improved. 40–50 % of students score at 25–30 for these three factors. Reading is the probable trouble area for students in the pre-assessment results, although only 38.9 % students score below 25.

For academic skills development, speaking and listening skills have indicated as probable success areas for students. However, reading and writing are probable trouble areas from the pre-assessment results. If comparing students amount among these four subcategories, reading is the most probable trouble area in students' learning. For this, reading skill needs to be emphasized in teacher's instruction.

For study and thinking skills, learning style and creative and critical thinking skills are two indicated success areas for students. Memory and study skills are two probable areas that need improvement to increase students' success in their achievement. Memory and study skills need to be improved for students' teaching and learning. Extra learning activities or e-learning tools have to be designed to increase students' memory and study skills.

In personal issues, self-esteem and personal relationships are two subcategories that indicate success areas for students. On the other hand, conflict resolution is a probable area that needs improvement. Therefore, teacher's instruction should emphasize problem-solving skills so that students can improve their skills to resolve conflicts that happened during teaching and learning.

In planning for the future, students' graduate capabilities are vital. Pre-assessment analysis results have shown that money management, personal purpose, and career planning are three indicated success areas for students. However, time management is the factor that needs improvement, which shows 44.4 % of students score at

Table 1 Students' score for pre-assessment analysis

Category	Academic skills development				Study and thinking skills			
Subcategory/score	Reading	Writing	Speaking skills	Listening skills	Learning styles	Memory	Study skills	Creative and critical thinking skills
30–40	7	7	9	10	14	5	5	15
25–30	4	5	4	5	2	7	9	1
Below 25	7	6	5	3	2	6	4	2

Category	Personal issues					Planning for the future				Resource needs
Subcategory/score	Motivation	Self-esteem	Personal relationships	Conflict resolution	Health	Time management	Money management	Personal purpose	Career planning	Support resources
30–40	8	12	10	7	8	7	12	14	14	9
25–30	7	0	4	8	6	8	3	2	2	6
Below 25	3	6	4	3	4	3	3	2	2	3

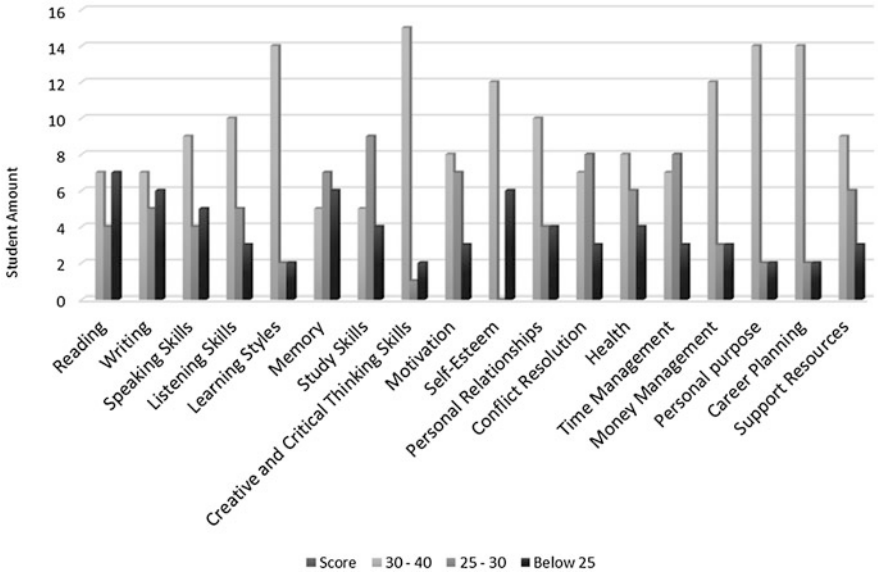


Fig. 1 Students' score for pre-assessment analysis in graph bar

25–30. Some soft skills, which include time management, may need to be embedded in teaching and learning. For resource needs, support resource would not affect students' success in their learning, which has shown 50 % students give high score, i.e., 30–40.

5 Conclusion

This study examined how pre-assessment results will improve academic and thinking skills among students. How students view their competencies and capabilities is indicative their life satisfaction and their consideration of future consequences. By identifying predictors of probable areas need for improvement and offering programs and learning activities that enhance self-efficacy beliefs should be of practical value. Creative and critical thinking skills, learning styles, personal purpose, and career planning are four probable areas that indicated success. This means that students perceived that these four areas are their success factors in their future life. Future study should focus on how we plan and design teaching and learning based on the pre-assessment results and increase the sample of pre-assessment survey.

References

Azizli, N., Atkinson, B. E., Baughman, H. M., & Giammarco, E. A. (2015). Relationships between general self-efficacy, planning for the future and life satisfaction. *Personality and Individual Differences, 82*, 58–60.

Broussard, M. J. S. (2014). Using games to make formative assessment fun in the academic library. *The Journal of Academic Librarianship, 40*, 35–42.

El-Daw, B., & Hammoud, H. (2015). The effect of building up self-esteem traning on students' social and academic skills. *Procedia Social and Behavioral Sciences*, *190*, 146–155.

Ellis, D. (1998). *Becoming a master student* (8th ed.). Boston: Houghton Mifflin Co.

Kobzeva, N. (2015). Scrabble as a tool for engineering students' critical thinking skills development. *Procedia Social and Behavioral Sciences*, *182*, 369–374.

Ngang, T. K., Nair, S., & PRachak, B. (2014). Developing instruments to measure thinking skills and problem solving skills among Malaysian primary school pupils. *Procedia Social and Behavioral Sciences*, *116*, 3760–3764.

Sekaran, U. (2000). *Research methods for business*. New York: Wiley.

Yusuff, K. B. (2015). Does self-reflection and peer assessment improve Saudi pharmacy students' academic performance and metacognitive skills? *Pharmaceutical Journal, 23*(3), 266–275.

A Comparative Analysis Between Teacher Assessment and Peer Assessment in Online Assessment Environment for Foundation Students

Seng Yue Wong, Wee Jing Tee and Wei Wei Goh

Abstract This study explores the comparative analysis of teacher assessment and peer assessment in the context of online-based assessment. The comparative study will help to improve grading accuracy in our assessment system. Therefore, this research is about the feasibility of implementation of online assessment by integration with learning activity management system (LAMS) for students by conducting a qualitative survey. In this survey, 18 students from foundation in computing course were asked on problems and disadvantages of paper-based assessment, feasibility study of acceptance for implementation of online assessment in tutorial and examination, suggestions, and recommendations. This paper will present the literature review on types and modes of assessment for students in higher learning institution. The qualitative survey on the comparison of these two modes of assessment is conducted and its findings are presented. From there, a conceptual model that integrated online assessment with learning activity management system (LAMS) into Taylor's Integrated Moodle e-Learning System (TIMeS) is proposed in this paper.

Keywords Teacher assessment · Peer assessment · Online assessment · Comparative analysis

S.Y. Wong (✉) · W.J. Tee · W.W. Goh
School of Computing and IT, Taylor's University, Lakeside Campus, Subang Jaya, Malaysia
e-mail: SengYue.Wong@taylors.edu.my

W.J. Tee
e-mail: WeeJing.Tee@taylors.edu.my

W.W. Goh
e-mail: WeiWei.Goh@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_31

1 Introduction

Internet is widely used by people daily for variety of purposes in many areas, such as e-commerce, e-learning, online banking, e-government, online shopping, social networking, and even in a quick search of information. The use of e-learning system has shown exponential growth from past to present. New technology provides more exciting and motivating environment to improve learning. Effective and fruitful usage of it by proper planning, implementation, and evaluation is critical for educational goal attainment (Gul et al. 2015).

In the age of technology, computers have become attitude objects. Attitude is a cognitive, affective, and behavioral response to any psychological objects or attitude objects in the environment or to social issues, organized by individuals based on their experience, motivation, and knowledge (Gul et al. 2015). Thus, online applications have also led to significant developments in education, including assessment. Studies on online applications specifically in measurement and evaluation applications, an indispensable part of teaching, have shown the significance of assessment (Gul et al. 2015).

Assessment is happening everywhere in higher education. We need to assess students' satisfaction and learning since it is no longer acceptable to assume students are learning what is taught in classroom (Broussard 2014). At some higher education institutions, they adopt an online assessment system where electronic profiles are rated based on instructional design, global learning object, growth assessment, hyperlink, utilization of applications, and application of coursework (Chang et al. 2012). Some campuses implement assessment rubric which comprises four aspects: learning objective, reflection, content quality, and overall performance (Chang et al. 2012).

Online assessment can be used for the evaluation of academic achievement too. Online assessment provides students with prompt feedback on learning information or concepts, and it significantly contributes to learning (Gul et al. 2015). Some researcher also suggested that the number of applications in higher education where online assessment is effectively utilized is rather restricted. Online assessment for students' performance and familiarization and guidance will make both assessment applications more economical and reporting process faster. Furthermore, confidentiality of the assessment will be controlled well (Gul et al. 2015).

Eventually, online assessment is able to be conducted since research has shown that there was no significance between the test scores obtained from paper-based assessment and online-based assessment to measure intelligence and abilities and it also shows high correlations between these two types of assessment mode (Gul et al. 2015). However, students believe assessment mode still essential as it enables them to use online applications effectively. Currently, it has been observed that technology was largely employed in most educational studies. The aim of this study is to determine the efficient assessment method for students' learning. In this paper, a comparative study between teacher assessment and peer assessment in online assessment is done. Some literature review on this two assessment modes is

discussed. The qualitative analysis will be presented in table form. Finally, a conceptual framework that shows integration of online assessment into our learning activity management system and Taylor's Integrated Moodle e-Learning System (TIMeS) is proposed.

2 Literature Review

Teaching can be assessed via multiple sources, which include self-reflection, teacher, and peer. These are three most common methods used in Web-based assessment, with which teachers may have options to select which method to adopt in order to serve their pedagogical needs (Chang et al. 2012). Classroom assessment also can be formative, summative, and possibly mixed of both (Cox et al. 2013). Each type of assessment has its advantages and disadvantages as a teaching evaluation source.

Teacher assessment is the most used approach, but the sole dependence on which might overlook students' real thoughts. Teachers are likely to use strict scoring criteria (Chang et al. 2012). Generally, teacher assessment is a traditional assessment option in most of the education institutions and can be summative, formative or both. Teachers still play a vital role as they retain considerable control over what classroom assessment practices are utilized and how these are implemented (Harris and Brown 2013).

Teacher-led assessment is practicable and dominates the classroom assessment. Teachers examine all students' performance based on their deeper understanding of learning goals, rubric assessment, and the criteria for the quality of learning tasks. Since teachers are providers of feedback, they are qualified as professional and suitable for all assessments, either traditional paper-based assessment or online assessment. However, student-led assessment also starts to be implemented so that students can understand their work and their involvement may increase their learning since they are involving to assess their own works.

Peer assessment enables students to observe other student's work, which brings new ideas and encourages self-reflection and improvement (Chang et al. 2012). Peer assessment also benefits students by offering opportunity to observe and compare peers' portfolio, emulate the strength, and prevent the weakness of other students. Students are likely to obtain more ideas or inspirations from peer comments from those who have in-depth understanding of them.

Some studies also indicated that the utilization of peer assessment and self-reflection is a good combination of learning strategies that fasten the process of development of metacognitive skills (Yusuff 2015) and self-regulation (Harris and Brown 2013) among students. Peer assessment often leverages on social pressure associated with students that do not want to lose face in front of their peers. Thus, students are motivated to focus on the process of self-reflection and self-regulation to prevent embarrassment of having their peers openly identify the probable gaps in their learning among their peers (Yusuff 2015). Student-led assessment processes

are designed to enable students to understand and apply quality criteria to their work, freeing them from teacher dependence (Harris and Brown 2013).

3 Methodology

Qualitative study is selected as the researchers need to listen to the views of participants for the study and obtain detailed information as well as seek to understand the participants' experiences. A qualitative survey is conducted to compare teacher assessment and peer assessment in the context of online-based assessment as well as suggestions on our future assessment system in School of Computing and Information Technology, Taylor's University. There are 18 students from this school as samples for this survey. Sample size between 10 and 20 is allowed for this type of survey although sample size between 30 and 500 is suitable for most studies (Sekaran 2000; Creswell 2002).

This survey consists of open-ended questions and provides qualitative data to obtain in-depth information about participants' thoughts, beliefs, knowledge, and reasoning and feelings about assessment's type, mode, and suggestions for assessment method. The questionnaire comprises of three parts: type of assessment, mode of assessment, and suggestions. The first two parts contain three questions, and last part contains one question only. The survey questions are shown as below:

A. ***Type of assessment***

A1. In your point of view, what kind of assessment that you prefer? Paper-based assessment or online assessment?
A2. If you prefer paper-based assessment, why?
A3. If you prefer online assessment, why?

B. ***Mode of assessment***

B1. Do you prefer peer assessment or teacher assessment? Why?
B2. Which type of assessment that make you earn much? Why?
B3. Do you agree that peer assessment is going to be implemented in university? Why?

C. ***Suggestions***

C1. What are your suggestions in assessment mode?

4 Results and Discussion

The comparative survey analysis is done between two types and two modes of assessment. Part A is comparison between paper-based and online assessment. On the other hand, Part B is comparison between peer assessment and teacher

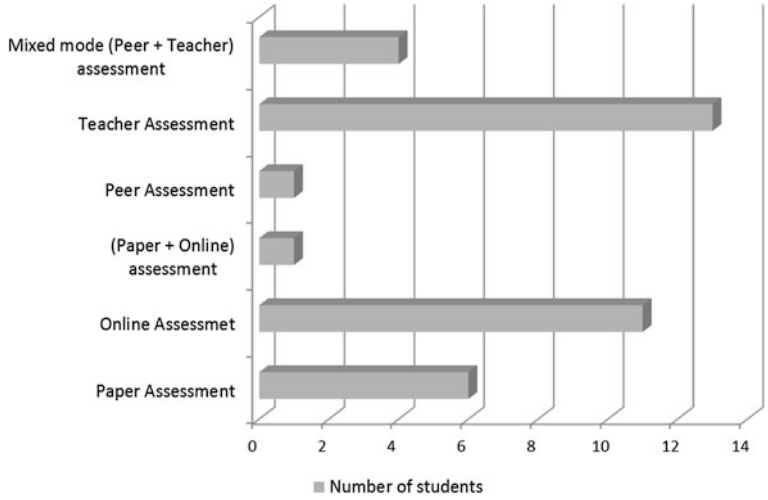


Fig. 1 Graph bar that shows the number of students according to their preferable types and modes of assessment

assessment. Figure 1 illustrates the graph bar that shows the number of students followed by their preferable types and modes of assessment.

The graph bar shows that most of the students prefer teacher assessment and online assessment, which indicates that thirteen students prefer teacher assessment and eleven students prefer online assessment. Table 1 shows the reasons why students choose online assessment as their preferable assessment in their studies.

Table 1 Reasons why students prefer online assessment

This way I can outsource more information and do better for my assessment
It saves trees
Online assessment is much faster and easier than paper-based assessment. There is no problem with handwriting, and the answers will be more organized and clear
It is a fun and interesting way to learn as it also captures my attention. It increases my IT skills as well
Because I do not like to learn things by reading words, instead going online to watch video and related article
I like it because it is interactive and very knowledgeable. It can be done easier and fast. Resources are available online
I prefer online assessment as we have the tools to do online research and learn much more. Humans are social creatures, and the traditional education system is examination based and makes people selfish and competitive. This sort of mind-set does not work in an organization where people are expected and required to work together to achieve a common goal. The true sign of success is when people learn to collaborate together and balance out each other's weaknesses
Online assessment, because it saves papers and easily done and be submitted

Table 2 Reasons why students prefer paper-based assessment

This is because paper is a media to facilitate the ideas effectively, and primary as well as secondary schools are all based on papers for 100 %
Paper-based assessment is better because traditional method is always the better
I prefer paper-based assessment because students can keep it well. When the examination is coming, students can just refer to the paper and make revision
I prefer paper-based assessment because paper-based assessments are eventually discussed in class in detail, while online-based assessments are sometimes done without discussion later
Its fine, but I feel bored of such type of assignment
In my point of view, I prefer paper-based assessment than online assessment, because online assessment using computer which will affect our eye when we stare long through the question. We will also feel sleepy when we stare long enough to the computer screen. Next, through online assessment, student can use autocorrect features. This will affect the English vocabulary skills among students. This is my reason why I prefer paper-based assessment
Students are less likely able to cheat during examinations

Their reasons are online assessment is much faster, easier, better, fun, interactive, and interesting and can save papers. On the other hand, Table 2 indicates the reasons why students prefer paper-based assessment. These indicated reasons are as follows: It is effective for primary and secondary school, they can do revision easily and can be discussed in detailed in class, and health problem could occur when using too much computers and less cheating in paper-based assessment. Although both types of assessment have their supported reasons, there is one student who prefers mixed-type assessment. This student prefers a mix of both assessments because it allows students that are sick to be able to do their assignments online while allowing those that attend the class to receive an advantage as they were punctual.

For comparison between peer assessment and teacher assessment, only one student prefers peer assessment with the reason that peers can help each other with their studies or personal life and they can understand their feeling better. Table 3 shows the reasons why students prefer teacher assessment or lecturer assessment. Most of the students still prefer teacher assessment or lecturer assessment because lecturers are more professional, qualified, experienced, less bias, and more accurate. There are four students who prefer a mix of both modes of assessment. Their given reasons are as follows: Peer assessment helps to know how all classmates think how the answers should be, whereas for lecturers, it is much more accurate as they are the ones who mark our papers at the end of the day; lecturer assessment, as they have more experience on the particular subject, whereas peers are able to relate more to students and understand them better, in a sense that they can correct students' mistakes.

Table 4 indicates students' suggestions on mode of assessment. Most of the suggestions prefer mixed mode of assessment because both roles are also crucial in assessment as mentioned in Table 4. Lecturer assessment is still the main choice of their preferable assessment mode. However, they will understand their learning tasks via peer assessment with the lecturer's monitoring.

Table 3 Reasons why students prefer teacher assessment or lecturer assessment

Lecturer assessment is better because it is less bias compared to peer assessment
Lecturer assessment, because it is more professional Lecturer. Because a lecturer is qualified No. Because they are not experienced
Only lecturer assessment is the best way to learn and get honest reviews No, this is a doom day if that is going be implemented, and honesty and justice issues are going to be compromised
Lecturer assessment, because a lecturer can be more accurate when judging and assigning the marks accordingly Group assignments, because it has many advantages such as decision making, understanding team members' opinions. We can learn from each other as a team No, because some students might not be sure about what they have learned and this will affect the partner
Lecturer assessment. Through lecturer assessment, students will more understand Lecturer assessment. Because lecturer assessment can let me more understanding and remember Agree, because students can try to found more information themselves
I prefer lecturer assessments because I do not trust the knowledge of someone who is taking the same level of education as me to assess my work I believe lecturer assessment will benefit me the most because the lecturer knows more and will be able to point out my mistakes in information so that I can improve more than being assessed by a student I agree of the peer assessment because it is something new and new things are vital for researching and improving the current state of education
Lecturer assessment because the lecturer will have the most experience in giving marks Paper-based assessment because writing and reading are much better way to learn even though I prefer online assessment Yes, because it can train peers to mark and learn more as well
Assessment that I prefer is lecturer assessment because lecturer is good and kind to us!
Lecturer assessment, because the lecturer is trained to provide proper assessment and feedback Lecturer assessment, because I am able to easily trust the feedback given to me Yes, because this trains students to share knowledge

Figure 2 illustrates a conceptual model which is integrating mixed mode of assessment in online assessment for LAMS in TIMeS. This conceptual model for TIMeS basically comprises three parts. As usual, teacher designs learning materials and contents for the module. Next, teachers will upload and share to students in TIMeS. Besides that, teachers also design individual or group learning activities including games for students' learning. Teacher can invite them by e-mail, or chat. After that, assessment tools can be designed and they are question and answer (Q&A), assignment, and quiz. After assessment activity, teacher can set the assessment setting for students to conduct peer assessment or self-assessment in TIMeS. Finally, teachers will do teacher assessment for students' activity after they update their grade for peer assessment. Teachers will do justification for the final assessment at the end of the class.

Table 4 Students' suggestions in assessment mode

I suggest that lecturers are given the whole priorities to assess a student, and it is an important message that peers are not able to assess another student as they are still in the learning process as everyone else. Hence, it is impossible for peers to judge peers even to give suggestions because most of them have inner various intentions that will be practiced if they are allowed to do assessment on other students
I suggest that a mix of peer-based and lecturer-based assessment be used in order to effectively gauge the student in terms of their understanding of the subject
Forms of assessment that help people improve their skills and critical thinking
No, because some students might not be sure about what they have learned and this will affect the partner
I suggest paper-based and lecturer assessment. It is still the useful matter for students to learn well
I want lecturers to supervise these assessments so that no mishap and improper assessments are done to harm the students' assessments
Team-based work should be graded individually
In my opinion, lecturer assessment and peer assessment should be done hand in hand for all subjects. Although it is difficult to achieve, I believe that in the long run it will be extremely beneficial to students especially those who are weak in a particular subject
The lecturers or peers must be able to provide useful and constructive feedback in order to ensure that the assessment is helpful to the student
Let the student interact with the lecturers in the class to gain more knowledge than old school learning system

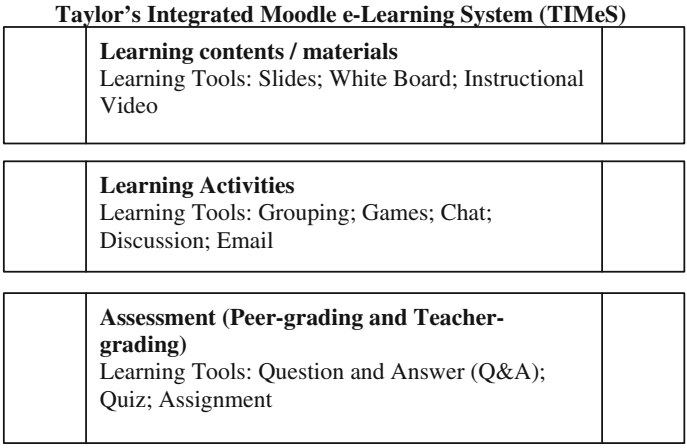


Fig. 2 Conceptual model for TIMeS to integrate peer assessment and teacher assessment

5 Conclusion

This comparative study explores teacher assessment and peer assessment in online assessment context so that it can be integrated with the suitable assessment mode in our e-learning system, TIMeS. The results indicate that most students prefer online assessment and they also prefer teacher assessment as their main choice for their performance evaluation and assessment. However, most of the students' suggestions are inclined toward mixed mode of both assessments (peer assessment and teacher assessment). Recent studies also promote peer assessment and self-assessment as strategies that actively involve students in the process of assessment (Harris and Brown 2013). Hence, combination of peer assessment and teacher assessment is an undeniable strategy to improve students' learning and develop students' self-regulation as well as metacognitive skills.

References

Broussard, M. J. S. (2014). Using games to make formative assessment fun in the academic library. *The Journal of Academic Librarianship, 40*, 35–42.

Chang, C.-C., Tseng, K.-H., & Lou, S.-J. (2012). A comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a web-based portfolio assessment environment for high school students. *Computers & Education, 58*, 303–320.

Cox, C. D., Peeters, M. J., Stanford, B. L., & Seifert, C. F. (2013). Pilot of peer assessment within experiential teaching and learning. *Currents in Phamarcy Teaching and Learning, 5*, 311–320.

Creswell, J. W. (2002). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (2nd ed.). Boston: Pearson Merrill Prentice Hall.

Gul, E., Cokluk, O., & Gul, C. D. (2015). Development of an attitudes scale toward online assessment. *Procedia Social and Behavioral Sciences*, *174*, 529–536.

Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education, 36*, 101–111.

Sekaran, U. (2000). *Research methods for business*. New York: Wiley.

Yusuff, K. B. (2015). Does self-reflection and peer assessment improve Saudi pharmacy students' academic performance and metacognitive skills? *Saudi Pharmaceutical Journal, 23*(3), 266–275.

Comparison Study on End-of-Posting OSCE Versus MCQ Scores in Anaesthesiology Posting, Taylor's University School of Medicine

Phutane Gnandev, Keng Yin Loh, M.I. Nurjahan, S. Roland and A.R. Noor

Abstract Objective structured clinical examination (OSCE) is a well-established assessment tool in clinical settings, used by many medical schools worldwide for clinical assessment. At Taylor's University School of Medicine, OSCE is the main assessment tool for clinical skills at the end-of-posting (EOP) and end-of-semester (EOS) examinations. Multiple-choice question (MCQ) is the major mode of assessment for theory, testing the cognitive domain of learning; OSCE assesses the clinical skills which are mainly in the psychomotor domain. The aim of the study is to ascertain the correlation between MCQ scores and OSCE performance scores in anaesthesiology posting in Taylor's University School of Medicine. A retrospective analysis was done on the EOP examination for anaesthesiology posting in 2015. A total of 58 students had sat for the examination. Their MCQ scores and OSCE scores are tabulated in Excel and analysed with SPSS version 17.0. The mean scores are MCQ (64.1 + 16.6) and OSCE (64.8 + 10.05). The Pearson's correlation of MCQ and OSCE is 0.257, which shows minimal correlation. MCQ has a near-normal distribution of bell curve, but OSCE shows bimodal presentation where 62.0 and 72.0 are the two modes noted. These results demonstrated the student who performed well in theory may not perform well in OSCE and vice versa. The lack of correlation between the results of different modes of assessment in the posting indicates that more guidance needs to be given to bridge the gap between the theory and clinical skills. One reason why there is a low correlation may be that the two tools test different domains. Deficiency in either component would indicate incomplete achievement of student outcomes.

Keywords OSCE · MCQ · Comparison · Anaesthesiology

P. Gnandev (✉) · K.Y. Loh · M.I. Nurjahan · S. Roland · A.R. Noor
Taylor's University School of Medicine, Sungai Buloh, Selangor, Malaysia
e-mail: Gnandev.phutane@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_32

1 Introduction

Anaesthesiology as a subspecialty posting in Taylor's University to undergraduates in the year four and year five is to introduce the basics of patient care coming under surgery and intensive care. They are given both theory and practical skill training using the anaesthesia teaching module. Adequate objectives are earmarked in various teaching modules in both theory and clinical skills to help them learn properly.

Summative assessment of theory knowledge or cognitive domain of learning is done using multiple-choice questions (MCQ), while assessment of basic practical skills or psychomotor skill is done using objective structured clinical examination (OSCE) at the end of semester (EOS). Appropriately constructed MCQ examinations are efficient, objective, and capable of discrimination and can be combined with other assessment strategies to contribute to a comprehensive student assessment strategy.

Objective structured clinical examination (OSCE) is a well-established assessment tool used by many medical schools worldwide for clinical assessment, which measures the aspects of clinical competence and psychomotor skills as well as theoretical knowledge. It has many advantages over traditional methods and should therefore be considered for inclusion in the assessment of medical students (Coovadia and Moosa 1985).

Our concern is whether the students are learning properly and are able to perform their task well within our acceptable limits. Assessment of their knowledge and skills and correlating them would give us an input to further our training programme to achieve a high standard.

The aim of this study is to ascertain the correlation between MCQ and OSCE performance scores at the end of posting in anaesthesiology at Taylor's University School of Medicine.

2 Methodology

A retrospective analysis of anaesthesiology using end-of-semester (EOS) assessment in 2015 was carried out. A total of 58 students had completed the EOS assessment which comprised of both MCQ and OSCE for clinical assessment. The total number of MCQ questions assessed was twenty. The total OSCE stations were 12 in which one is anaesthesiology OSCE.

The scores were retrieved from academic registry after approval from programme director and dean of Taylor's University School of Medicine.

The data were tabulated using Excel and analysed using SPSS version 17.0. This study would not in any way affect the actual results of the students. The results from this study may be able to give some feedback on their performance and provide steps to improve the structure of undergraduate anaesthesiology training.

Table 1 Mean scores of MCQ and OSCE

Central tendency	MCQ	OSCE
Mean	64.1	64.8
Median	62.5	63.8
Mode	55.0	62.0 and 72.0
Std. deviation	16.6	10.0

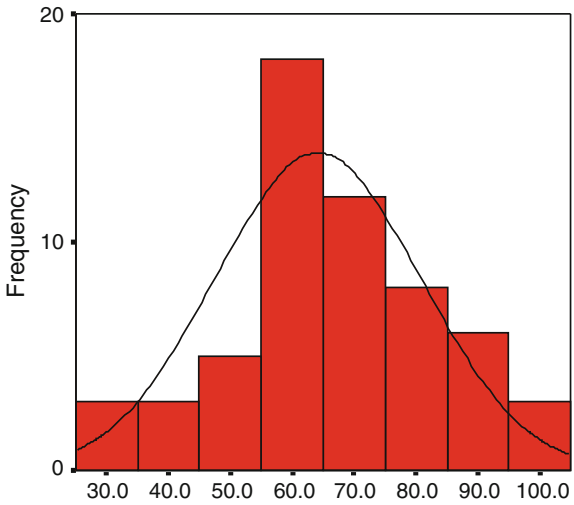


Fig. 1 Bar chart of MCQ scores

3 Results

The mean scores for MCQ were 64.1 + 16.6 and for OSCE were 64.8 + 10.05. There was no difference in the mean score of MCQ compared to OSCE ($p > 0.05$) (Table 1). The mean score for MCQ showed a normal distribution (Fig. 1), but the mean score for OSCE showed a bimodal presentation, where 62.0 and 72.0 are two modes noted (Fig. 2).

The Pearson's correlation coefficient of MCQ and OSCE is 0.257, which showed minimal correlation. These results demonstrated the student who performed well in theory may not perform well in OSCE and vice versa (Fig. 3).

4 Discussion

Teaching is an art and science done with dedication, patience, and perseverance to get the students do their best in a given area to excel. There are different methods and modules to assess how effective was the teaching. To select a method or combination of methods in assessment is always challenging. Assessment is essential to get to know our strength and weakness.

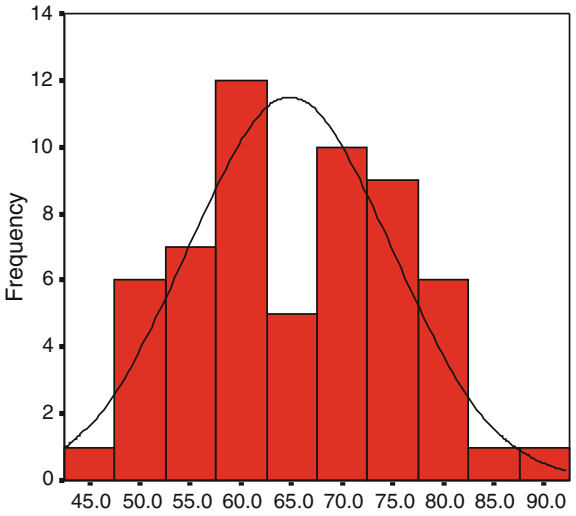


Fig. 2 Bar chart of OSCE scores

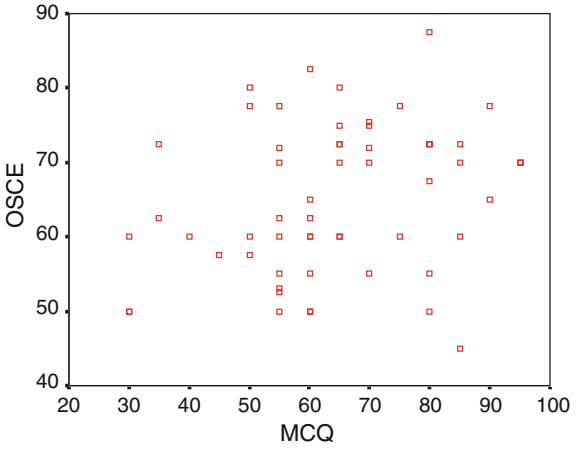


Fig. 3 Scatterplot showing the correlation between MCQ and OSCE scores ($R = 0.257$)

Assessment is essential to give feedback to students and teachers, reduce our teaching weakness, and improve our strength. Assessment if done preteaching and post-teaching would see the amount of effective learning the student has undergone.

In Taylor's University School of Medicine, "single best answer" type of MCQ is used as the major tool of assessment for theory component. It covers a wide range of clinical conditions which were taught within the postings. On the other aspect, OSCE is the main tool for assessing clinical skills. Wide domains of clinical skills are assessed in OSCE, such as performing a clinical procedure, examining patient, and communication skill.

In this cohort of study, the mean scores of both MCQ and OSCE were almost similar and there was no statistical significance between these two modes of

assessment. However, the Pearson's correlation between OSCE and MCQ is 0.257, which is considered poor. The possible reason was that MCQ assessed mainly theoretical knowledge and OSCE assessed the actual clinical skills in psychomotor domain. Those students who scored well in theory may not be doing well in clinical component. Nevertheless, in clinical school, a good competency of the students must be in both the theory and clinical skills.

This cohort only has 58 students which constituted a small sample size. More studies need to be conducted to evaluate further in other semesters and across all other subjects. But the above results did bring us the attention that remedial measures need to be initiated for candidates who are weak in either component.

5 Conclusion

The overall mean score of MCQ and OSCE did not show any significant difference. The performance of candidates in the theory domain and clinical aspects generally is almost similar. However, lack of correlation between MCQ and OSCE indicates that there is a gap between theory knowledge and application of practical skills of individual student. Remedial measures with more guidance to bridge the gap between the theory and clinical skills are necessary. Deficiency in either component would indicate incomplete achievement of student outcomes. Proper intermix of the group and group interaction is essential.

6 Recommendation

Future study using a larger number of candidates across all semesters is recommended. A comparison across all the different clinical subjects is recommended to ascertain any inadequacy in theory (MCQ) versus clinical psychomotor skills (OSCE).

Reference

Coovadia, H. M., & Moosa, A. (1985). A comparison of traditional assessment with the objective structured clinical examination (OSCE). *South African Medical Journal, 67*(20), 810–812.

The Art of Communicating Statistics: Why Simulation and Graphical Presentations Are Important Tools for Understanding Statistical Concepts

Aishah Mohd Noor

Abstract Applications of statistical methods require a deeper understanding of its fundamental concepts. Hence, teaching statistics to engineering students is very challenging. This is because the statistical concepts introduced are too abstract. In addition, engineering students focus on the statistical procedures rather than the understanding of the concepts. Thus, it is difficult to introduce statistical reasoning and thinking if the educators stress only on the statistical procedures. Moreover, the traditional teaching approach is lack of experiencing statistics in real applications. The approach failed to stimulate students' self-exploration, lack of self-thinking and self-evaluation. Therefore, they tend to be isolate learner and most of them trap with statistics is about calculation. In this paper, two teaching tools based on graphical representation and simulation activity will be demonstrated. The main concern is to improve students learning process and more importantly experiencing and appreciate statistics. Moreover, the advantage of simulation activity is it allows them to explore two essential questions that is the 'what if' and the 'why.' Such questions are crucial in order to foster thinking-based learning environment. The scopes of our discussions in this paper are topics selected from Engineering Statistics (EQT 271) syllabus offered by Institut Matematik Kejuruteraan, Universiti Malaysia Perlis (UniMAP).

Keywords Probability distribution · Statistical inference · Engineering education · Thinking-based learning

1 Introduction

There are two important questions for educators; the first is "How well will students learn?" and the second is "How good is instructor or teacher preparation?" Both questions are essential in order to stretch out students' potential and also in

A.M. Noor (✉)
Universiti Malaysia Perlis, Perlis, Malaysia
e-mail: aishah.mn@unimap.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_33

preparing them to the real world. For example, the industrial demands are constantly change and require us to be more innovative and creative. Reinventing the traditional approach is mandatory to allow innovations in teaching statistics for engineering students. The pen–paper-based approach no longer supports students' needs and industrial demands.

Tishkovskaya and Lancaster (2012) presented the major directions of the statistics education reform movement. The reformation includes (a) pedagogical reforms toward development of conceptual understanding and teaching to statistical thinking and reasoning; (b) changes in the content of statistics courses, especially introductory level courses; (c) improving the instructional techniques used in statistics courses; and (d) integration of technology and computer-based methods into teaching statistics as an important tool for effective delivery of teaching and essential part of effective pedagogy.

Thus, it is important for educators to communicate among each other and develop awareness to step out of the traditional way. However, we believe that the first step is always challenging. In this paper, we share our experience in fulfilling the role as service course to engineering students at Universiti Malaysia Perlis (UniMAP). The main objective is to improve students learning experience through the philosophy that exploration is more important than calculation.

Teaching statistics to engineering students is a very challenging. This is because students are not able to see the concepts in real. In engineering classes, students are able to experience the concepts in real applications during laboratory sessions. On the other hand, statistics as a subject is too abstract compared with engineering subject. Hence, students struggle with self-motivation while in statistics classroom. This is the first challenge for statistics educators.

Thus, the traditional approach should be integrated with technology touch in order to improve students learning environment in a way that they are not focusing on the statistical procedure but should clearly understand when to use statistical tools and how to interpret statistical result. The use of technology will help learning of statistics, for example, simulation activities can be a teaching tool for delivering the statistical concepts. Technology has also expanded the range of graphical and visualization techniques to provide powerful new ways to assist students in exploring and analyzing data and thinking about statistical ideas, allowing them to focus on the interpretation of results and understanding concepts rather than on computational mechanics (Tishkovskaya and Lancaster 2012).

The traditional approach is lack of experiencing statistics in real applications. The approach failed to stimulate students' self-exploration, lack of self-thinking and self-evaluation. Therefore, they tend to be isolate learner and most of them trap with statistics is about calculation. The learning environment should also embrace students' ability to go beyond the well-posed text book problems. This is the second challenge.

Simulation is widely recognized as a useful tool for teaching statistics (Wood 2005; Novak 2013; Mills 2002). Cobb (1994) noted that teaching laboratories are increasingly incorporating computer simulations to illustrate important statistical concepts and to allow students to discover important principles themselves. Mills

(2002) provide a comprehensive literature review on using computer simulation methods in statistics in order to guide and help students understanding of difficult concepts. Mills (2002) summarized and critically evaluate the literature review and although researchers are encourage to using computer simulation methods to enhance students understanding of statistical concepts, the literature reveals very little empirical research to support the recommendations. Hagtvedt et al. (2008) discussed teaching the topic of confidence interval using simulation in order to help students abandon their misconception and achieve understanding. Mills (2002) developed a simulation tool by using Excel spread sheets that encourages experimentation with multiple confidence intervals derived from the same population. Chance and Rossman (2006) provide examples of the integration of simulation to enhance topics throughout an introductory statistics course through a combination of Minitab macros and specifically designed applet. They also highlight how simulation can motivate students to learn the more mathematical derivation.

The third challenge would be in encouraging students to appreciate statistics. Statistics is not about calculation or delivering concepts as a fact but require practitioners to be proficiently using statistics in the process of decision making. In other words, practitioners should be able to develop statistical thinking and reasoning while using or working with statistical methods. Once they have developed the appropriate statistical skills, they are able to communicate the statistical results in a meaningful way. Viles (2008) presented a practical activity to improve statistical thinking among engineering students. Miltiadis et al. (2005) discussed the role of statistical thinking in industrial engineers and managers in twenty-first century. The paper provides the key principles of statistical thinking and then discusses the possible reasons for lack of statistical thinking in modern organizations.

Christine and Johannes (2005) discussed some challenges in teaching statistics with Excel. Several examples such as measures of central tendency and variability are introduced by using Excel. They noted that the use of statistical software in replace to manual calculation will be the future challenge in teaching statistics. However, the progress in computing technology has resulted in changes in the way in which students learn statistics and the way in which instructors teach statistics.

Thus, based on the above literature reviews and industrial feedbacks, there is a need to revise, reevaluate and reform the teaching of statistics to engineering students. For this purpose, our first step in introducing statistical software and Excel spread sheets would be beneficial for our future engineer. This is a part of our contribution in statistics education for engineering students. In this paper, we highlight the issues of why reformation is a must in Sect. 2. To what follows, we provide examples of simulation activities in order to help students understand the statistical concepts. The exploration activities using the ‘what if’ questions are further discussed in Sect. 4. In the last section, we present our further direction for venturing into new ideas in teaching statistics for engineering students.

2 Why Reformation Is a Must

Integrating statistical software has had a major impact on the teaching and learning of statistics worldwide. The development of user-friendly statistics packages such as SAS®, SPSS®, Excel®, and MINITAB® should increase students' motivation to learn statistics. However, the current culture and awareness specifically among engineering community, the applications of statistical software for solving engineering problems are still considered to be at a low level of awareness.

An advantage of the use of statistical packages in the statistics classroom is that it allows students to understand statistical concepts using simulation activities. Moreover, 'what if' questions can further open the door for self-exploration and consequently self-thinking. Education should provide space and time for students to experience their thinking. Thus, simulation activities allow instructors to guide students to statistical thinking and reasoning.

The statistical package not only operates as a tool for exploration activities, but it can also help students to accomplish computational tasks more effective, efficient, and quick. This process is very important since in real applications, engineers are required to make decision based on the data analysis as soon as possible.

3 Examples of Simulation Activities

3.1 Probability Distributions

In this course, there are three major parts. The first is the descriptive statistics, followed by probability concepts and lastly the statistical inference. It would be difficult to let them see the difference between probability concepts and to what follow is why the study of statistical inference. For engineering students, notations are, for example, $X \sim \text{Bin}(n, p)$, $X \sim \text{Poi}(\lambda t)$ and $X \sim N(\mu, \sigma^2)$ are too abstract. In order to help them, we introduce simulation and graphical of probability histogram and explain the characteristic of the histogram based on the parameter of the random variable under study. They may not able see it as description about the distribution of random variable, X.

Binomial distribution is an important discrete probability distribution that students in introductory statistics courses may encounter. Many educators have reported that students never seem to explore the features of the binomial variable but instead focus on calculating probabilities. Focusing on the calculation only is not enough to increase student's motivation to learn probability.

In order to explore the role of parameters in a certain probability distribution such as the sample size, n, and the probability of success, p, in a binomial probability distribution, we illustrate the characteristics of those parameters using Microsoft Excel. Using graphical representation and with the help of Microsoft

Excel function, 'what if' questions are now further discussed. For example, students are asked the following:

1. How does the sample size change the shape of probability distribution?
2. For large sample size and small probability of success, show that the shape of the original probability distribution is approximately normal?

Figures 1 and 2 are two examples of graphical representation of binomial probability distribution using Minitab package for a different probability of success, *p,* while keeping the sample size unchanged. However, while in class we demonstrate using Microsoft Excel to allow students to animate changes in the shape of binomial probability distribution by replacing any value of corresponding parameters. Although, those examples are discussed in many statistics textbooks, statistical instructors do not demonstrate it in a practical way instead they just focus on the calculation of probability of an event of interest.

Hence, it is important to engage the used of any statistical software or the simplest tool using the Microsoft Excel function in order to illustrate the characteristics of a

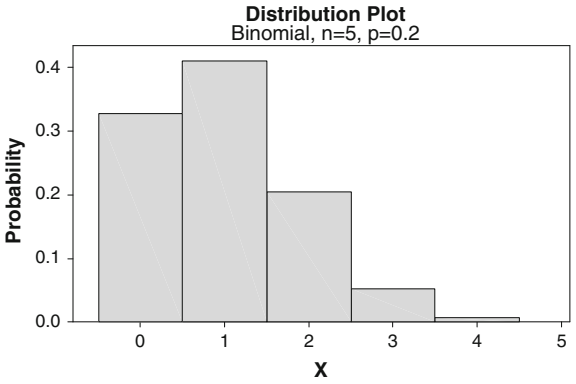


Fig. 1 Distribution plot for $X \sim \text{Bin}(5, 0.2)$

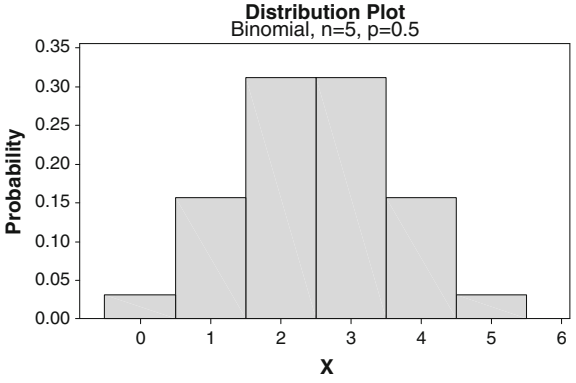


Fig. 2 Distribution plot for $X \sim \text{Bin}(5, 0.5)$

probability distribution. Furthermore, students are exposed to simulating data with certain parameters of interest and later are able to see the behavior of the data set simulated. For example, in this learning environment, students are able to apply the use of descriptive statistics such as the variance of a data set simulated from normal probability distribution and how does the sample size contribute to the spread of a normal curve. Thus, the topic of probability distribution is not all about calculating the probability of an event but should be further explored by engaging students to the characteristics of the random variable and its probability distribution.

3.2 Sampling Distributions

Researchers and educators have found that statistical ideas are often misunderstood by students and professionals (delMas et al. 1999). A deeper understanding of fundamental concepts is essential in order to develop statistical reasoning skills among students. Thus, the introduction to sampling distribution should be prepared by means of statistical experiments. The sampling distributions laboratory sheet is developed in order to guide student self-exploration and self-discovery.

Sampling distribution is defined as distribution of a sample statistic. Instructor should be sensitive with the terms "sampling" and "statistic." One of the common misunderstandings among students is the different between parameter and statistic. Thus, it reflects difficulty to further illustrate the concept of sampling distribution. For this reason, we discuss by means of example in order to foster students motivation to grasp the idea behind the sampling distribution concept.

We illustrate one example to deliver the concept of sampling distribution: electrical firm manufacturer light bulbs that have a length of life that is approximately normally distributed with mean equal to 800 h and standard deviation of 40 h. Find the probability that a random sample of 16 bulbs will have an average of life less than 775 h? This is a common example and in practice, we focus on the calculation of $P(\bar{x} > 775 \text{ h})$. However, we choose to start with simulating data from $X \sim N(800, 40^2)$. The purpose is to let student see the process of data collection. Although it is not real, we create situation through imagination of collecting n light bulbs from a large population and measure the lifetime of each light bulbs. The data are given as in the statistical package and example as in Table 1.

Using the simulated data, students are asked to calculate the sample means from each samples' data simulated and, consequently, draw the histogram of sample means (see Fig. 3). From this activity, we had proved that the theoretical sample means are normally distributed that is $\bar{X} \sim N\left(800, \frac{40^2}{25}\right)$. Therefore, the concept of sampling distribution is easily presented. Students are not only absorbing the concept but are also exposed to data collections and representations.

The sampling distribution is a distribution of a sample statistic. While the concept of a distribution of a set of numbers is intuitive for most students, the concept of a distribution of a set of statistics is not. Therefore, we introduce

Table 1 Simulated data from $X \sim N(800, 40^2)$

Sample	1st item	2nd item	3rd item	4th item	…	25th item
1	821	743	826	817	…	772
2	819	785	778	747	…	761
3	794	835	836	813	…	784
4	812	795	822	815	…	788
5	782	782	739	807	…	848
6	804	832	741	834	…	831
7	762	750	785	811	…	747
8	799	795	851	792	…	796
9	725	886	776	815	…	859
⋮	⋮	⋮	⋮	⋮	⋮	⋮
100	923	762	748	746	…	747

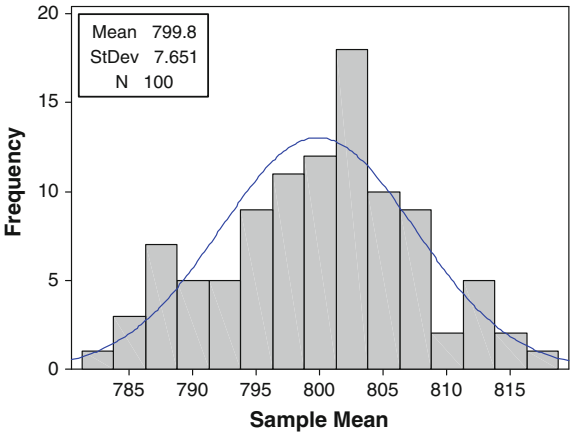


Fig. 3 Histogram for sample mean based on 100 samples

simulation activity in order to help students gain understanding about a statistic such as the sample mean in above example and its distribution. Interestingly, we found that the process of simulating data and proving the theoretical distribution property of a statistic using simulated data is very helpful when explaining the concept of statistical inference.

3.3 Estimation Using Confidence Interval

Statistical inference is the heart of statistics. It is about how to make valid conclusions based on data from random samples. If instructor able to demonstrate the idea behind the inference and how does the sampling distribution plays it roles and as the key to enter the idea of statistical inference, the remaining topics, for

example, inference about regression model, can be explained and understood by the students easily.

There are two major parts in statistical inference. The first is the estimation of the population parameter and the second is testing the hypothesis about population parameter of interest. How to start the discussion about confidence interval (CI)? Promoting the concept of CI is preferable by using simulation activity. This is because we will be exposed students to the process of decision making. Students are asked to simulate data, construct CI based on simulated data, and finally provide the interpretation based on the graphical representation of those CI constructed. The steps are as follows:

1. Simulate data of size 25 from normal distribution, $X \sim N(800, 40^2)$.
2. Calculate sample mean and sample standard deviation.
3. Obtain confidence interval.
4. Repeat step 1–3 for 100 intervals.
5. Record the number of intervals do not include the population mean, $\mu = 800$.

Figure 4 shows the first fifty confidence intervals for the estimation of population mean. From this figure, students are able to see which intervals out of 100 intervals do not include the population mean $\mu = 800$. They are asked to repeat the process ten times and calculate the average number of intervals do not contain the population mean. Later on, students are able to formulate the conclusion that for a 95 % CI approximately there will be five intervals do not contain the population mean. One may follow the same process apply for 90 % CI and 99 % CI.

In order to cultivate the exploration activity, students are asked to:

1. Compare the width of intervals based on 90 % CI, 95 % CI, and 99 % CI using the interval plot from Minitab package. How do you explain the difference?
2. What if sample size is changed to 15, 20, 25, 30, 50? How does sample size give effect to the interval?

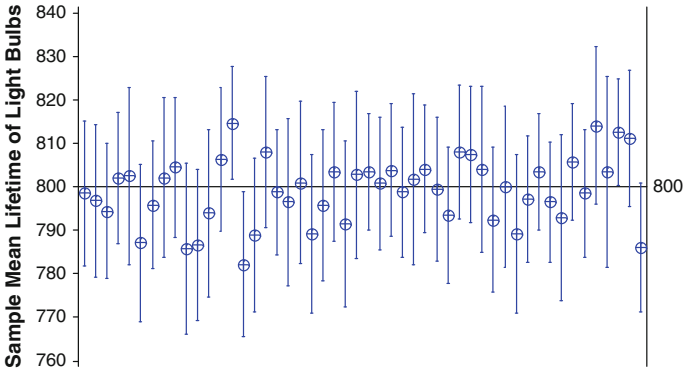


Fig. 4 95 % confidence interval for population mean lifetime of light bulb

Thus, such computation and exploration would be carried out using a statistical package instead of pen–paper–calculator-based process of learning. In the above exploration activities, the focus is not about calculation of the interval at certain level of confidence but we go in-depth by looking at the comparison between level of confidence and its confidence interval.

4 Students Evaluation to Simulation Activities and Integrating Statistical Software

Teaching transformation through integrating statistical software is a new experience for our students. Specifically in our educational system, engineering students are being taught in a traditional way for mathematics and statistics courses. As a result, they do not see the importance of both courses in enriching their skills and supporting the engineering courses. Thus, we conduct the first transition here at UniMAP in order to provide evidence to the engineering community. Such evidence is essential for us to change the perception about Statistics course. It is a common view among engineering community that the role of statistics is as a service course. Thus, at the end of the course short survey is conducted to obtain students feedback. Table 2 gives the summary of students' feedbacks. This is the first initial study and material presented in this article is regarded as our preliminary study and will be improved in the future. The following are survey questionnaire delivered to our students:

1. A simulation activity helps me to understand the sampling distribution and confidence interval concepts.
2. The use of statistical software motives me to learn statistics.
3. Graphical representation is helpful and useful in looking beyond the values or numbers.
4. Statistics should be taught using statistical software.
5. Microsoft Excel helps in computation instead of using calculator and encourages self-discovery.
6. The course had increased my knowledge and skills in statistics as compared to traditional pen–paper-based experience in matriculation/diploma level.

Table 2 Students' responses

Question	Response		
	Yes (%)	No (%)	No opinion (%)
1	61	21	18
2	63	17	20
3	57	23	20
4	62	21	17
5	58	24	16
6	60	13	17

Table 2 gives a signal that almost 30 % of students may not able to see the advantages of new approaches in statistics delivery for engineering students at UniMAP. Thus, further plan in teaching, learning, and evaluation process can be improved in order to increase the awareness and helps them to see the advantages. One of the reason is students are focused on examination-based system. If there is no assessment given for simulation problems or data analysis problems which require them using the software for examination purpose, they would prefer a pen–paper-based approach. However, further statistical tools to validate students preferences are still ongoing process since this is our first step and initial findings.

5 Rooms for Improvements

This paper discusses three examples of simulation activities in introducing statistical concepts. The motivation is to increase students' interest to learn statistics. We believed that the first step for education reform is always challenging. To our experience, students' preparations for simulation thinking-based learning and skills should be enhanced during the first year of undergraduate study. Moreover, education system at university level should increase awareness among Statistics instructor about the importance of using technology for a basic Statistics course. This is to facilitate students learning and develop in-depth understanding on the statistical concepts. Embracing students to such new exploration environment require encouragement from instructors. Although there is no further study about the effectiveness of such learning medium, we believed that the efforts to step out from the traditional way of teaching Statistics will be fruitful in long term.

Finally, four issues highlighted based on our first experience introducing simulation activities using statistical software are as follows:

1. Asking students to anticipate the simulation results before interacting with it appears to be an effective instructional technique.
2. However, we received a passive respond since students are trap with calculation-based mind-set and this is a new experience to them.
3. Instructor attitude in terms of underestimate students' potential by keeping them at comfort zone. Let them explore the statistical concepts although they are engineering students.

Require a well-structured materials for teaching simulation in a way that instructor is able to guide and help students to see the advantages and experience the statistical thinking and reasoning in making decision.

6 Conclusion

Bridging real applications with statistical concepts will encourage and motivate engineering students to appreciate statistics. Engineering students should be exposed to basic statistical data analysis using Microsoft Excel or any other statistical software. If those tools are being integrated while in the classroom, it helps to embrace students in experiencing statistical thinking and reasoning. Although statistics course in engineering program play the role as a service course, the demand for statistical skills should be seen as important literacy among engineering graduates. In future, we plan to develop a teaching tool to guide young instructors to teach statistics with skills required not so much focusing on calculation per se but statistical thinking skills and ability to communicate statistical results as the product of in-depth understanding on the statistical concepts.

References

Chance, B., & Rossman, A. (2006). Using simulation to teach and learn statistics. In *Proceedings of the Seventh International Conference on Teaching Statistics*, Bahia, Brazil. Retrieved March 27, 2014 from http://www.ime.usp.br/∼abe/ICOTS7/Proceedings/PDFs/InvitedPapers/7E1_CHAN.pdf

Christine, D., & Johannes, K. (2005). Teaching statistics with Excel—A big challenge for students and lecturers. *Austrian Journal of Statistics*, *37*(2), 195–206.

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher, 23*(7), 13–20.

delMas, R. C., Garfield, J., & Chance, B. L. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education*. Retrieved May 29, 2014, from www.amstat.org/publications/jse/secure//v7n3/delmas.cfm

Hagtvedt, R., Jones, G. T., & Jones, K. (2008). Teaching confidence interval using simulation. *Teaching Statistics, 30*(2), 53–56.

Mills, J. D. (2002). Using computer simulation methods to teach statistics: A review of the literature. *Journal of Statistics Education*, *10*(1). Retrieved January 28, 2014, from www.amstat.org/publications/jse/v10n1/mills.html

Miltiadis, M., Antony, J., Antony, F., & Kumar, M. (2005). Statistical thinking and its role for industrial engineers and managers in the 21st century. *Managerial Auditing Journal, 20*(4), 354–363.

Novak, E. (2013). Effects of simulation-based learning on students' statistical factual, conceptual and application knowledge. *Journal of Computer Assisted learning, 30*(2), 148–158.

Romero, R., Ferrer, A., Capilla, C., Zunica, L., Balasch, S., Serra, V., et al. (1995). Teaching statistics to engineers: An innovative pedagogical experience. *Journal of Statistics Education*. [Online], Retrieved March 23, 2014, from http://www.amstat.org/publications/jse/v1n1/garfield.html

Tishkovskaya, S., & Lancaster, G. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education*, *20*(2). Retrieved January 28, 2014, from www.amstat.org/publications/jse/v20n2/tishkovskaya.pdf

Viles, E. (2008). Paper-clip case: A practical activity to improve statistical thinking. *Teaching Statistics, 30*(2), 57–60.

Wood, M. (2005). The role of simulation approaches in statistics. *Journal of Statistics Education*, *13*(3). Retrieved February 2, 2014, from http://www.amstat.org/publications/jse/v13n3/wood.html

An Evaluation of the Relationship Between Student Engagement, Academic Achievement, and Satisfaction

Kok Choy Cheong and Bessie Ong

Abstract Classroom and other campus activities are both important parts of student engagement in institutions of higher learning. Involvement in these activities gives them a sense of belonging and identity with the institution besides fostering academic competencies and self-development. This study seeks to identify the range of activities that students are engaged in a liberal arts program at a private institution of higher learning. A self-reported questionnaire was used for cross-sectional sampling of students who were in the program for at least three semesters. Data reduction was done by exploratory factor analysis. Using the ten factors extracted, the extent to which engagement explains students' academic achievement and with satisfaction was assessed with Pearson's correlation. Engagement in clubs and organizations was found to be positively correlated with both students' cumulative grade point averages and satisfaction.

Keywords Campus engagement · Liberal arts · Academic achievement · Satisfaction

1 Introduction

Student engagement in educationally purposeful activities can take two forms—routine classroom participation and out-of-the-class activities. Such activities can include good study habits, regular class attendance, interest and enjoyment of academic challenge, frequent interaction with faculty and other institutional personnel, and participation in sports and clubs. Astin (1999) associates the quantity and the psychological experiences in these activities to the amount of student

K.C. Cheong (✉) · B. Ong
School of Liberal Arts and Sciences, Taylor's University Lakeside Campus,
47500 Subang Jaya, Malaysia
e-mail: cheong.kokchoy@taylors.edu.my

B. Ong
e-mail: ong.bessie@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_34

learning and personal development. Studies have consistently shown that such involvement is positively correlated with academic performance (Kuh 2008), student satisfaction (Pascarella and Smart 1991; Pennington et al. 1989), and persistence (Bean 2005; Tinto 1993, 2005).

The institutional policy and that of the American Degree Transfer Program (ADP) in particular, where this research is conducted, are driven toward creating enriching student experiences at the campus. The ADP practices a liberal arts education curriculum and prepares students to transfer credits after five semesters to undergraduate programs in North American Universities. A liberal arts program is generally noted for its broad-based curriculum which promotes a transformative undergraduate experience (Pascarella and Blaich 2013). In order to evaluate the engagement of students in a liberal arts program in Malaysia, this study is done with two aims:

1. To determine the range of collegiate activities that students in a two-year undergraduate transfer program are engaged in and outside the classroom and
2. To examine how these self-reported experiences correlate with academic achievement, as represented by students' cumulative grade point averages (CGPA) with our students' satisfaction with ADP.

2 Literature Review

Student engagement refers to the time and effort that a student spends on participation in classroom and out-of-the-class activities on campus, which Kuh et al. (2008) describe as educationally purposeful activities. Many institutions of higher learning are now paying attention to optimize students' experience on campus for a wholesome student development.

The theoretical framework of student engagement originates from Alexander Astin's (1984), theory of involvement, and Pace's concept of quality of effort (1980). Both theories basically explain that students change and develop as a result of attending institutions of higher learning. The core concepts of Astin's theory are related to three elements. A student's demographics and previous experiences provide the inputs, while the environment in higher education (also known as college) accounts for all his experiences, which will translate into outcomes covering a student's knowledge, attitudes, beliefs, and values upon graduation (Astin 1993).

Following Astin's work, there is now a multitude of efforts by others who have researched on the various applications of this theory in higher education. In particular, engagement is generally associated with positive student outcomes such as better academic achievements, retention, and satisfaction with the institution. Zehner (2011) who draw on data from the National Survey of Student Engagement (NSSE) survey reported that students in Purdue University who are active in cocurricular activities earned more credit hours and achieved grade point average

(GPA) of 3.00 and above. Extracurricular activities such as sports, clubs, and organizations are established to have other effects on students including better behavior, improved grades, and positive attitudes toward studies (Massoni 2011). Kuh et al. (2008) found that engagement in educationally purposeful activities is positively correlated with first-year student grades and persistence to second year in the university. Similarly, Tinto (1993) concludes that students who are more engaged are better integrated intellectually and socially with the institution of learning than those who are less engaged. Because participation builds connections to the campus community, i.e., a sense of belonging, students who are engaged feel that they are accepted, supported and involved with other people in the setting of a campus. This gives them a sense of belonging and identity with the institution; thus, they are less likely to leave. In contrast, commuter students who are generally less engaged also demonstrated the highest attrition rates. Participation in campus activities such as student organizations and clubs also leads to commitment and positive perception of experiences, which are correlated with greater satisfaction (Montelongo 2002; Abrahamowicz 1988). Higher levels of engagement with faculty, staffs, and students together with effort contribute to not only a higher cumulative grade point average (GPA) but also perception of satisfaction with one's entire academic experience (Weber et al. 2013).

Active participation in the classroom and campus activities not only contributes to students' success, but also does build skills that enable students to manage change and behave ethically and morally (Weber et al. 2013; Pascarella and Terenzini 2005; Pascarella 1985). Kuh (2009) views this as value-added which reflects an institutional quality. Thus, it should be an education institutional policy and ethos to encourage students' involvement in classroom and outside the classroom activities for them to have an enriching experience.

3 Method

3.1 Data Collection

The data for this study originated from our previous work (Cheong and Ong 2014). We used the College Student Experiences Questionnaire (CSEQ, Pace & Kuh, 1998). The CSEQ is chosen for this initial investigation as it contains a comprehensive inventory of the student experiences. The survey collects information about the student's background (e.g., age, sex, class, race and ethnicity, residency, major, and parent's education level) and asks questions about the student's experience with the institution in three areas: (a) college activities, (b) the college environment, and (c) estimate of gains. This paper focuses on experiences in college activities or "quality of efforts." There are eleven scales in this dimension comprising 150 questions. We selected 111 questions which are relevant to our program.

A convenient sample was obtained over several days with the help of four student volunteers. We collected completed questionnaires from 168 full-time students who had been in ADP for at least four semesters. We believed these students would have a wider range of experiences to report on.

3.2 Data Analysis

An exploratory factor analysis (EFA) with SPSS (version 21) was used for data reduction to handle all the questions. Principal component factors were extracted by the varimax rotation method and Kaiser normalization with 25 iterations. An eigenvalue of 1.0 was used as the cutoff, and only items loading at 0.5 or higher were retained as a factor. Reliability of the samples was analyzed by Cronbach's alpha, of which a value of at least 0.7 generally suggested that the items in the group are measuring the same thing (Hair et al. 2006). Mean Likert scores were calculated for all the factors.

Pearson's correlation analysis was used to observe relations between the engagement factors with academic achievement as measured by CGPA and satisfaction. Satisfaction is defined in this study as the willingness to return to ADP.

4 Results

Exploratory factor analysis (EFA) retained 67 variables which are extracted into ten factors that are considered to be reliable (Cronbach's alpha reliability coefficient $\geq$ 0.7, Table 1). These ten factors together explained almost 60 % of the variance. Writing experience is not retained as a distinct factor.

Out of the ten factors extracted, students reported only two to have a positive impact on their CGPAs (Table 2). Participation in clubs and organizations (Factor 6), which explained 25 % of the variability in CGPA, showed the strongest correlation. Variables retained for this factor relate to involvement as an ordinary member, committee member, a leader as well as regular meetings with faculty and club advisor to discuss activities of a campus sports club, society, or student government group.

Experience in the sciences, mathematics, and other quantitative courses (Factor 9) is the second factor that is found to be positively correlated with CGPA. Students reported that experiences such as routine problem-solving skills in mathematics, acquiring laboratory skills and experimentation, explaining theoretical concepts related to scientific, environmental, and global issues, and reading scientific articles for an assignment did help to boost their academic achievement.

When each of the ten factors was correlated with the willingness to return to ADP, an indicator of satisfaction with ADP, students reported experiences in course learning (Factor 3) and participation in clubs and societies (Factor 6) as significant

Table 1 Exploratory factor analysis of student experiences at ADP

Factor	Cronbach's alpha
1. Library experiences (7)	0.791
2. Computer and ICT (8)	0.869
3. Course learning (3)	0.718
4. Experiences with faculty (9)	0.846
5. Art, music, and theater (7)	0.885
6. Clubs and societies (7)	0.912
7. Personal experiences (5)	0.804
8. Student acquaintances (6)	0.898
9. Science and quantitative experiences (9)	0.885
10. Topics of conversation (6)	0.864
Cumulative % of total variance	59.8 %

Numbers in parenthesis show the number of variables that are retained for each factor

Table 2 Correlation of students' experiences with cumulative grade point averages

Factor component	Cumulative grade point average		Strength of association (%)
	Pearson's coefficient	*p*-value	
1. Library experience	0.054	0.436	
2. Computer and ICT	0.106	0.124	
3. Course learning	0.055	0.424	
4. Experience with faculty	0.041	0.550	
5. Arts, music, and theater	−0.016	0.821	
6. Clubs and organizations	0.247**	0.000	25
7. Personal experience	−0.073	0.288	
8. Student acquaintances	0.076	0.269	
9. Science and quantitative experiences	0.138*	0.045	14
10. Topics in conversation	−0.074	0.285	

*Correlation is significant at the 0.05 level (two-tailed)
**Correlation is significant at the 0.01 level (two-tailed)

factors for their satisfaction (Table 3). In contrast having to engage with the sciences, mathematics and other quantitative courses (Factor 9) showed a reverse impact on satisfaction. This factor explained 16.4 % of the variability that students are unwilling to return to ADP. Similarly, topics in conversation (Factor 10) also showed a negative impact accounting for 14.3 % of the variability of dissatisfaction.

Table 3 Correlation of students' experiences with satisfaction (willingness to return)

Factor component	Pearson's coefficient	*p*-value	Strength of association
1. Library experience	−0.105	0.129	
2. Computer and ICT	0.033	0.630	
3. Course learning	0.137*	0.047	14 %
4. Experience with faculty	−0.025	−716	
5. Arts, music, and theater	0.125	0.070	
6. Clubs and organizations	0.160*	0.020	16 %
7. Personal experience	0.031	0.654	
8. Student acquaintances	0.079	0.253	
9. Science and quantitative experiences	−0.164*	0.017	16.4 % (Inverse)
10. Topics in conversation	−0.143*	0.038	14.3 % (inverse)

*Correlation is significant at the 0.05 level (two-tailed)

5 Discussion

Although this study did not search for causal participation in extracurricular activities (clubs and societies) and success in ADP, it is clear that participation and academic success are strongly correlated as evidenced by participants' better CGPA and students' willingness to return, *vis-a-vis* satisfaction. The ADP has a very active and vibrant "clubs and societies" culture which offers voluntary membership in sports and athletic academic clubs, the arts (music, theater, and dance), student government, as well as community service. The findings from this study are consistent with those by Weber et al. (2013), Massoni (2011), Zehner (2011), Montelongo (2002), as well as Pascarella and Smart (1991) who found that participation in clubs, societies, and athletics has a positive effect on a student's satisfaction in college. They explained these participations build social connections with other students and the institution and enable the application of academic skills in real-world context, besides developing life skills, discipline, self-esteem, and other character development. These experiences cultivate positive influence on students' lives, making them more committed to persist and to be motivated to do well in their studies.

At ADP, all students are required to take some basic math and science courses as part of the liberal arts curriculum to cultivate quantitative reasoning. Mathematics and science subjects are perceived by many students to be more difficult, compared to other subjects, and that only a clever few can study. There is evidence that suggests student perceptions of difficulty are reflected in examination performance (Osbone and Collins 2001). Doing these courses is also hard work which requires a lot of time studying, practicing mathematical problems, and working in the laboratory. The ADP students too have the same perception and reported experiences in science and quantitative courses to be a significant antithesis to satisfaction.

Conversely, those who reported better CGPA are those who are likely to ace these subjects, hence the positive correlation of science and quantitative experiences with CGPA shown in this study.

One of the key factors that correlate with students' satisfaction with ADP is found to be the hallmark instructional strategies practiced here, as is typical in most liberal arts programs. The various course learning experiences had engaged students through classroom discussions, working collaboratively with others on class assignments, projects, and presentations and through integrating ideas from multiple sources. Such experiences help in developing habits of the mind and heart that enlarge their capacity for continuous learning and personal development. Pascarella and Blaich (2013) suggest that these typical instructional/learning environments in liberal arts college students produce consistent cognitive impacts on students and increase the probability of persistence and satisfaction.

A significant negative correlation on topics in conversation with satisfaction reflects the somewhat apolitical culture present in Malaysian private institutions. The variables in this factor cover social issues such as peace, justice, race relations, and ethical issues in science and technology, such as pollution and the military. It is not the norm to openly discuss such issues, while others might have subdued conversations only in private. Contrary to the CSEQ norms from four-year US universities (Gonyea et al. 2003), ADP students do not perceive writing experiences (such as thinking about grammar, sentence construction, word choice, seeking help from an instructor to improve writing assignments, and writing a major paper exceeding 20 pages) to be notable. This could be attributed to the fact that these students are enrolled in a two-year program to complete the general education courses at ADP. Thus, their writing assignments do not require the intense reading and preparation required by those taking courses at higher levels.

This study builds on our previous work that students at ADP have gained communication skills and the ability to work in teams (Cheong and Ong 2014) by suggesting that these skills are developed from the quality of efforts that students employ, especially in their course learning and participation in extracurricular activities. In this regard, the efforts of the faculty and the administrators at ADP are responsible in educating our students who can take their place as "global citizens." This study has shown that out-of-the-class involvement is important for both academic achievement and satisfaction with the institution. Therefore, extracurricular engagement should be encouraged to augment and enhance the students' experience.

References

Abrahamowicz, D. (1988). College involvement, perceptions, and satisfaction: A study of membership in student organizations. *Journal of College Student Development*, *29*(5).

Astin, A. W. (1984). Student Involvement: A developmental theory for higher education. *Journal of College Student Personnel., 25*, 297–308.

Astin, A. W. (1993). *What matters in college: Four critical years revisited*. San Francisco: Jossey-Bass.

Astin, A. W. (1999). Involvement in learning revisited: Lessons we have learned. *Journal of College Student Development 40*(5), 587–598.

Bean, J. P. (2005). Nine themes of college student retention. In A. Seidman (Ed.), *College student retention: Formula for student success* (pp. 215–244). Washington DC: ACE & Praeger.

Cheong, K. C., & Ong. B. (2014). *What have students gained from the college experience*. *Taylor's 7th Teaching and Learning Conference Proceedings*, 2014, pp. 341–348, Springer ISBN 978-981-287-398-9

Gonyea, R. M., Kish, K. A., Kuh, G. D., Muthiah, R. N., & Thomas, A. D. (2003). *College Student experiences questionnaire: Norms for the* (4th ed.). Bloomington, IN: Indiana University Center for Postsecondary Research, Policy, and Planning.

Hair, J. F., Black, C. B., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). New Jersey: Pearson Education Inc.

Kuh, G. D. (2008). Diagnosing why some students don't succeed. *The Chronicle of Higher Education, 55*(16), A72.

Kuh, G.D. (2009). What Student Affairs Professionals Need to Know About Student Engagement. *Journal of College Student Development*. *50*(6), 683–706.

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *Journal of Higher Education, 79*(5), 540–563.

Massoni, E. (2011). Positive effects of extra curricular activities on students. ESSAI: Vol. 9, Article 27. Available at: http://dc.cod.edu/essai/vol9/iss1/27

Montelongo, R. (2002). Student participation in college student organization: A review of literature. *Journal of the Indiana University Student Personnel Association,* 50–64. https://education.indiana.edu/…/2

Osborne, J., & Collins, S. (2001). Pupils' views of the role and value of the science curriculum. *International Journal of Science Education, 23*(5), 441–467.

Pace, C. R. (1980). Measuring the quality of student effort. *Current Issues in Higher Education, 2*, 10–16.

Pace, C.R. and Kuh, G.D. (1998). College Student Experiences Questionnaire, Fourth Edition. Indiana University Center for Postsecondary Research and Planning, *School of Education*, Bloomington.

Pascarella, E. T. (1985). College environmental influences on learning and cognitive development: A critical review and synthesis. In J. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 1, pp. 1–64). New York: Agathon.

Pascarella, E., & Blaich, C. (2013). Lessons from the Wabash National Study of Liberal Arts Education. *Change* March–April. http://www.education.uiowa.edu/centers/crue/Publications.aspx#WNSLAE

Pascarella, E., & Smart, J. (1991). Impact of intercollegiate athletic participation for African American and Caucasian men: Some further evidence. *Journal of College Student Development*, *32*(3).

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A Third decade of research*. San Francisco: Jossey-Bass.

Pennington, D., Zvonkovic, A., & Wilson, S. (1989). Changes in college satisfaction across an academic term. *Journal of College Student Development*, *30*(11).

Tinto, V. (1993) *Leaving College: Rethinking the causes and cures of student attrition* (2nd Ed.). USA: The University of Chicago Press.

Tinto, V. (2005). Moving from Theory to Action. In A. Seidman (Ed.), *College Student Retention: Formula for Student Success* (pp. 317–333). Washington DC: ACE & Praeger.

Weber, K. L., Bauer, R., & Zhang, Q. (2013). Does involvement really matter? Indicators of College student success and satisfaction. *Journal of College Student Development, 54*(6), 581–611.

Zehner, A. (2011). Co-curricular activities and student learning outcome. https://www.heacademy.ac.uk/…/GPA-report

Assessing Item Difficulty and Discrimination Indices of Teacher-Developed Multiple-Choice Tests

Ahmad Zamri Khairani and Hasni Shamsuddin

Abstract Item analysis is an important procedure to determine the quality of the items. The purpose of this study is to assess two important indices in item analysis procedure, namely (1) item difficulty (*p*) and (2) item discrimination (*D*) as well as a correlation between them. The study involves ten 40-item multiple-choice mathematics tests. A total of 1243 Form 2 students from public schools in Penang, Kedah, Perak, and Pahang are employed as sample for this study. Both indices are calculated based on classical test theory framework because of the advantage provided for teachers. On average, only 67 % (min = 50 %, max = 87.5 %) of the items are considered of good quality that can be kept for future testing. There is incomprehensible result regarding the correlation between the two indices. Implication in terms of teachers' competency in test development is also discussed.

Keywords Difficulty index · Discrimination index · Item analysis · Mathematics Form 2

1 Introduction

Multiple-choice test (MCT) is considered as one of the most important formats of testing students' performance. This is because MCT provides several advantages for test developers. Firstly, MCT is able to assess a large sample of learning outcomes. As such, it covers broad areas of important knowledge, skills, and ability associated with students' performance (Miller et al. 2009). Secondly, MCT

A.Z. Khairani (✉)
School of Educational Studies, Universiti Sains Malaysia,
11800 Gelugor, Penang, Malaysia
e-mail: ahmadzamri@usm.my

H. Shamsuddin
Sekolah Menengah Sains Kepala Batas,
11800 Kepala Batas, Penang, Malaysia
e-mail: emel_hasni@yahoo.com

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_35

provides objective scoring that ensures accurate and easy marking for teachers, which in turn provides instant feedbacks that can help both students and teachers toward improving learning and teaching (Higgins and Tatham 2003; Kuechler and Simkin 2003).

In addition, the MCT also offers evidence of higher reliability and validity compared to other modes of testing. For example, because reliability depends on the number of items (Wells and Wollack 2003), a typical 40-item MCT is likely to produce the higher reliability coefficient compared to 10-item restricted response test, or 5-item essay test. Reliability in the MCT format is also preserved based on automated scoring compared to examiner-influenced essay test.

Because MCT is able to cover many important learning outcomes, the format is considered to demonstrate high evidence of content validity (Popham 2000). Ensuring high evidence of content validity is important not only because it ensures representativeness of important aspects of the measured construct (Kerlinger 1986), it also ensures high evidence of the test's construct validity (Yaghmale 2003). In terms of practicability, an increase in class size in future classroom may lead to the increase of the need for MCT (Schrecker 2009).

Nevertheless, many quotes that MCT is difficult to develop, especially in terms of constructing plausible distractors (Burton et al. 1991). More importantly, many teachers are not able to identify difficulty levels of the item (Jandaghi 2009; Zhou 2009). There are many cases where teachers simply take items from reference books without properly addressing the issue of content representation of the curriculum. As such, information gathered from the test is vague, making the decision made from the information questionable. Therefore, the purpose of this study is to provide evidence of the quality of items for teacher-developed MCTs through investigation of item difficulty and item discrimination indices. Assessing quality of items is valuable in improving items which will be used for future testing. In addition, the study also aims at providing empirical data on a long-standing issue of the relationship between both statistics.

2 Item Analysis

According to Nitko (2004), item analysis can be defined as a process of collection, summarizing and using information from students' responses to make decisions. Item analysis always report statistics such as item difficulty index, item discrimination index, as well as item mean and variance. In addition, item analysis also includes examining the function of the item such as in distractor analysis. Besides the item-level examination, item analysis also includes test-level statistics such as reliability index as well as evidence of test validity. There are various advantages associated with item analysis. Nevertheless, with regard to assessing quality of items, item analysis provides insights by providing evidence of good and bad items. As such, good items can be retained for future testing, whereas bad items can be improved. More importantly, frequent item analysis will eventually enhance

teachers' knowledge and skills as well as competency in test development. This study, however, will only address quality of items developed based on only two important statistics, namely item difficulty and discrimination index.

Item analysis serves many purposes. Firstly, it provides evidence on whether a particular item functions as intended. Secondly, item analysis offers feedbacks to students about their performance by clarifying their incorrect choices. Item analysis also provides feedback to teachers with regard to students' difficulties in the assessment tasks. Profiling nature of students' error in the test helps teachers to identify remedial needs for his or her students. If it is found that students repeatedly perform lowly in a particular area, then teachers may need to revisit the learning outcomes or the curriculum in order to provide diagnostic assessment of the content.

2.1 Item Difficulty Index

The item difficulty index (also known as item facility index) for an item i, p_i, is calculated as a proportion of examinee who answers correctly for an item i (Miller et al. 2009). For example, if 25 out of 40 examinees answer correctly for Item 2, then $p_2 = 25/40 = 0.625$. Note that Item 3 ($p_5 = 0.30$) is more difficult than Item 4 ($p_6 = 0.90$) since only 30 % of the students are able to answer correctly Item 3, compared to 90 % for Item 4. Item difficulty is calculated for every item in the test; i.e., there are 40-item difficulty indices in a 40-item MCT. The values of probable item difficulty are between 0.00 and 1.00, with values close to 1.00 that indicate easy items. There is a wide range of acceptable item difficulty index, with the range of 0.2–0.8 which is considered as a rule of thumb. According to Crocker and Algina (1986), item difficulty index is essential because it affects all test score statistics, such as item mean, item variance, and the reliability. Bachman (1990) quotes that assessing item difficulty is a common practice in determining whether or not the test is a reliable measure of students' ability. Henning (1987) explains that a test which is too easy or too difficult lost its ability to separate students, thus weakening the reliability.

2.2 Item Discrimination Index

The item discrimination index for item i, D_i, is calculated as difference between proportion higher ability students and lower ability students who answered correctly for an item i (Miller et al. 2009).

$$D_i = \text{pH} - \text{pL}, \tag{1}$$

where
pH proportion of high ability students who answered correctly item *i*
pL proportion of low ability students who answered correctly item *i*.

Item discrimination is important statistics in terms of assessing quality of items because tests are intended to provide information about individual differences in the ability that the tests aim to measure. There are various ways of identifying high and low ability students. A common practice is to take the top 30 % of students as the cutoff score for high ability students and the bottom 30 % as lower ability students (Miller et al. 2009). For small sample size such as at classroom analysis, the 50 % cutoff score can be applied. The values of D range from −1.00 to +1.00. Item 3 ($D_3 = 0.6$) is less able to discriminate between high and low ability students compared with Item 4 ($D_4 = 0.85$). As a rule of thumb, $D_i \geq 0.2$ is considered productive in discriminating students of different abilities. As in the item difficulty index, item discrimination is also calculated for every item in the test.

3 Classical Test Theory

One important shortcoming of using the item difficulty and item discrimination indices as an indicator of item quality is that the indices are influenced by characteristics of the group. For example, an easy item in high ability group may become moderate or difficult for lower ability students and vice versa. Item I might have low discriminating power for homogenous ability students compared to group with various abilities. The shortcoming may be related to classical test theory, measurement framework used to calculate the indices. Thus, it is important to describe characteristics of the theory in order to provide better understanding of the indices.

Classical test theory (CTT) helps to explain the unobservable trait (T) and the observable trait (X) of whatever constructs being measured in a simple linear equation which is as follows (Crocker and Algina 1986):

$$T = X + E \tag{2}$$

where
T true score of the trait;
X observed score (from the test); and
E error score (from the measurement).

In the example of item difficulty index above, the unobservable trait is the test difficulty statistic, while the observable trait is the number of students scored correctly on the item. Error score, meanwhile, includes unexpected responses such as guessing or carelessness that influence respondent's score that in turn influence the observed score of the item. CTT deals with raw scores to provide the explanation

between both traits. As mentioned earlier, this is considered problematic since raw scores are sample dependent. The problem is addressed in the development of other measurement theory called item response theory (IRT). That is, in IRT, test parameters, such as item difficulty index, are not influenced by the group from which sample is drawn. The description of IRT is beyond the scope of this study, but more information can be found in the work of Hambleton and Swaminathan (1985) as well as van der Linden and Hambleton (1997). Nevertheless, calculating item difficulty using IRT is complex and requires knowledge about item modeling as well model-data expectation. Therefore, even though CTT provides less accuracy in the estimation of the parameter as well as calculation is sample dependent, it is easier to work with since it does not require complex calculation. School teachers can easily calculate the item difficulty index based on CTT framework.

4 Methodology

4.1 Sample

The sample for this study consists of 1243 Form 2 students (14 years old) from 10 public schools in the state of Penang, Kedah, Perak, and Pahang. The schools are selected based on purposive sampling, and approximately 120 students are selected from each school. The number of students selected is considered appropriate for calculating the difficulty index of each item in a particular test based on CTT (Cappelleri et al. 2015).

4.2 Instrument

The study examines 10 teacher-made tests, that is, one test for every school. Each test is a 40-item MCT that covers three strands, namely (1) number, (2) shape and space, and (3) relations that are considered as important strands in students' mathematical ability construct (Curriculum Development Division 2000). Each correct answer is scored 1, while no mark is given if it is answered incorrectly. While each teacher may have their own way of developing the tests, every test should be developed based on the learning outcomes specified in the Form 2 mathematics curriculum specifications (Curriculum Development Division 2002). The researchers have no control over test development. However, in a normal practice, after a test is constructed, it is given to experience teachers such as head of panel in order to ascertain suitability of the items. All tests are administered as school's year-end examinations.

4.3 Data Analysis

Both item difficulty and item discrimination indices for each of the 400 items are calculated, and number of items that fit the characteristics of good item is reported. In this study, acceptable value of the item difficulty index is 0.26–0.75, while good items should also have more than 20 % discrimination power (Ong and Mohamad 2014). Correlation between the indices for every test is calculated using Pearson correlation using IBM SPSS 22.0.

5 Findings

Table 1 shows the number of items and its percentage for each acceptable item difficulty and item discrimination indices as well as correlation (r) between item difficulty and item discrimination indices. For Test *A*, 34 of the items (85.0 %) demonstrate acceptable item difficulty index, while 35 items (87.5 %) have good discrimination power. 28 items (70.0 %) from Test *A* are of good quality and can be retained for future testing. Meanwhile, the correlation between item difficulty and item discrimination indices for Test *A* is low but significant at $p < 0.05$. Overall, 300 of the 400 items (75.0 %) are within the acceptable item difficulty range. Test *G* contains the most acceptable items with 38 items (95.0 %), while Test *D* provides the least with only 19 items (47.50 %). With regard to item discrimination index, 318 items (79.50 %) are in the range of acceptable values. As high as 39 items (97.5 %) in Test *E* demonstrate acceptable item discrimination index compared to only 24 items (60.0 %) in Test *B*. It can also be seen that most teachers are able to develop item with satisfactory discrimination compared to items with right difficulty.

Table 1 Difficulty index, discrimination index, and the correlation

Test	$0.26 \leq p \leq 0.75$	$D \geq 0.20$	$0.26 \leq p \leq 0.75$ and $D \geq 0.20$	r
A	34 (85.0 %)	35 (87.5 %)	28 (70.0 %)	0.319*
B	27 (67.50 %)	24 (60.0 %)	22 (55.0 %)	−0.117
C	33 (82.50 %)	35 (87.5 %)	31 (77.5 %)	0.041
D	19 (47.50 %)	34 (85.0 %)	21 (52.5 %)	−0.747**
E	35 (87.5 %)	39 (97.5 %)	35 (87.5 %)	0.086
F	35 (87.5 %)	30 (75.0 %)	28 (70.0 %)	−0.213
G	38 (95.0 %)	33 (82.5 %)	32 (80.0 %)	0.361*
H	28 (70.0 %)	29 (72.5 %)	25 (62.5 %)	−0.157
I	24 (60.0 %)	30 (75.0 %)	20 (50.0 %)	−0.289
J	27 (67.50 %)	29 (72.5 %)	26 (65.0 %)	−0.582*

*$p < 0.05$, **$p < 0.01$

However, result shows that only 67 % (min = 50 %, max = 87.5 %) of the items are considered of good quality that can be kept for future testing based on the combined criteria of $0.26 \leq p \leq 0.75$ and $D \geq 0.20$. Test *E* has the most items that can be retained, with 35 items (87.5 %) while as many as half of the items from Test *I* need improvement. For Test *A*, even though the teacher is able to develop many items with acceptable item difficulty (34 items) or item discrimination indices (35 items), the number of items that can be retained is considered low (28 items).

With regard to the correlation between the indices, the findings show inconclusive results. Test *A* and Test *G* show significant weak correlations, while moderate, weak, and negative correlations are reported for Test *D* and Test *J*. Another observation is that negative correlation between the indices can be traced back to tests which have the least number of good items.

6 Discussions

The purpose of this study is to assess the item difficulty and item discrimination indices of a teacher-developed MCTs as well as a correlation between them. Results show that while there are some teachers who are able to develop a good test, there are a considerable number of teachers who have lack of competency in doing so. For example, in Test *I*, only half of the items are considered of having good quality. As such, one might wonder whether information provided by the test is enough to make useful decision regarding students' mathematical ability. Note that since this is a year-end test, information from the test is used to make many important decisions about the students. For example, the score of the test is used as the Form 3 *take-off-value* (TOV)—baseline data for monitoring progress of the students. Teachers usually use the TOV for tracking purposes—where students are sorted into courses, groups, classes, or schools, according to achievement and ability (Callahan 2005). Researchers such as Steven and Vermeersch (2010), Turner (2007), and Carbonaro (2005) warn that tracking can have detrimental effects if it is not conducted appropriately. Tracking without accurate information may pose a risk with regard to students' progress.

The present study reports mixed results in terms of correlation between both indices. The result is consistent with positive, but weak correlations reported from the study of Sim and Rasiah (2006) as well as Pande et al. (2013) and negative correlation reported by Mitra et al. (2009) and Sabri (2013). As such, the study provides another indication that test developers need to have knowledge and skills in both indices in order to develop a good item. Apparently, knowledge and skill on the conceptualization of item difficulty may be found in Bloom's taxonomy, but no guidelines are presented for developing items with proper discrimination power. As propose by Haladyna (2004) and McDonald (2007), the ability of an item to discriminate students of different ability depends heavily on the distractors. Nevertheless, writing plausible distractors itself are also considered as difficult (DiBattista and Kurzawa 2011). For example, each option for the item should be

selected by at least 5 % of the students (Ware and Vik 2009). In short, as observe by Downing (2005) as well as Tarrant and Ware (2008), weakness in the development of individual item will eventually jeopardize the whole test. Perhaps, it is about time more studies to be conducted in term of characteristics of an item with good discrimination power.

It is also interesting to know more about the procedure used by the teachers to develop their tests and quality of items produced using the procedure since there is a possibility that many teachers did not have the necessary skills and time to develop their own test items. Rather, they assemble published test item from reference books. They, therefore, abandoned important steps in test development, such as identifying test parameters, or conceptualization of items' difficulties. In a more worrying situation, teachers do not construct the table of specifications because of the similar reasons—thus putting the content validity of the test questionable. Therefore, it is about time that research on assessment practice among teachers to be conducted so that teachers are made aware of the importance of developing a good test.

It limits notwithstanding, and the present study demonstrates that even with careful considerations in test development, many items are still deemed not suitable. As such, the researchers would like to suggest that each teacher develop his or her own item bank to facilitate test development. Item bank refers to suitable test items that are coded with various pertinent item information such as item difficulty and discriminating indices (Gronlund 1998). Researchers such as Gronlund (1998) and Umar (1999) agree that item bank helps inexperienced teachers to develop good test that provide accurate information regarding the measured construct. It is made possible because item information has been predetermined and can be used to tailor the purpose of the test.

References

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Burton, S. J., Sudweeks, R. R., Merrill, P. F., & Wood, B. (1991). How to prepare better multiple-choice test items: Guidelines for university faculty. Retrieved 3 July 2015 from https://testing.byu.edu/handbooks/betteritems.pdf

Callahan, R. M. (2005). Tracking and high school english learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305–328.

Cappelleri, J. C., Lundy, J. J., & Hays, R. D. (2015). Overview of classical test theory and item response theory for quantitative assessment of items in developing patient-reported outcome measures. *Clinical Therapeutics, 36*(5), 648–662.

Carbonaro, W. (2005). Tracking, students' effort and academic achievement. *Sociology of Education, 78*(1), 27–49.

Crocker, L., & Algina, A. (1986). *Introduction to classical and modern test theory*. Orlando: Holt, Rinehart and Winston Inc.

Curriculum Development Division. (2000). *Sukatan pelajaran Matematik menengah rendah*. Kuala Lumpur: Ministry of Education.

Curriculum Development Division. (2002). *Curriculum specifications for mathematics form 2*. Kuala Lumpur: Ministry of Education.

DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. Retrieved from http://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=1061&context=cjsotl_rcacea

Downing, S. M. (2005). The effects of violating standard item writing principles on tests and students: the consequences of using flawed test items on achievement examinations in medical education. *Advances in Health Sciences Education, 10*, 133–143.

Gronlund, N. E. (1998). *Assessment of student achievement* (6th ed.). Needham Heights: Allyn and Bacon.

Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Springer Science and Business Media.

Henning, G. A. (1987). *A guide to language testing—development, evaluation, research*. London: Newbury House Publisher.

Higgins, E., & Tatham, L. (2003). Exploring the potentials of multiple-choice questions in assessment. *Learning and Teaching in Action, 2*(1), 1–12.

Jandaghi, G. (2009). Assessment of validity, reliability and difficulty indices for teacher-built physics exam questions in first year high school. *Arts and Social Sciences Journal, 11*, 1–4.

Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). New York: CBS Publishing.

Kuechler, W. L., & Simkin, M. (2003). How well do multiple choice tests evaluate student understanding in computer programming classes? *Journal of Information System Education, 14* (4), 389–400.

McDonald, M. E. (2007). *The nurse educator's guide to assessing learning outcomes*. Sudbury, MA: Jones and Bartlett.

Miller, M. D., Linn, R. L., & Grounlund, N. E. (2009). *Measurement and assessment in teaching*. New Jersey: Pearson International.

Mitra, N. K., Nagaraja, H. S., Ponnudurai, G., & Judson, J. P. (2009). The levels of difficulty and discrimination indices in type A multiple choice question of pre-clinical Semester 1 multidisciplinary summative tests. *IeJSME, 3*(1), 2–7.

Nitko, A. J. (2004). *Educational assessment of students* (2nd ed.). Englewood Cliffs, NJ: Merrill.

Ong, E. T., & Mohamad, M. A. (2014). Pembinaan dan penentusahan instrumenn kemahiran proses sains untuk sekolah menengah. *Jurnal Teknologi, 66*(1), 7–29.

Pande, S. S., Pande, S. R., Parate, V. R., Nikam, A. P., & Agrekar, S. H. (2013). Correlation between difficulty and discrimination indices of MCQs in formative exam in physiology. *South-East Asian Journal of Medical Education, 7*(1), 45–50.

Popham, W. J. (2000). *Modern educational measurement: Practical guidelines for educational leaders*. Boston: Allyn and Bacon.

Sabri, S. (2013). Item analysis of student comprehensive test for research in teaching beginner string ensemble using model based teaching among music students in public universities. *International Journal of Education and Research*, *l*(12), 1–14.

Schrecker, E. (2009). The bad old days. *Chronicle of Higher Education, 55*(40), 31.

Sim, S., & Rasiah, R. I. (2006). Relationship between item difficulty and discrimination indices in true/false type multiple choice questions of a para-clinical multidisciplinary paper. *Annals of the Academy of Medicine, Singapore, 35*, 67–71.

Steven, P. A., & Vermeersch, H. (2010). Streaming in Flemish Secondary Schools: Exploring teachers' perceptions of and adaptations to students in different streams. *Oxford Review of Education, 36*(3), 267–284.

Tarrant, M., & Ware, J. (2008). Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Medical Education, 42*, 198–206.

Turner, P. (2007). Reflections on numeracy and streaming in Mathematics education. *Australian Mathematics Teacher*, 63 (2), 28–33.

Umar, J. (1999). Item Banking. In G. N. Masters & J. P. Keeves (Eds.), *Advances in measurement in educational research and assessment*. New York: Pergamon Press.

van der Linden, W. J., & Hambleton, R. K. (1997). *Handbook of modern Item response theory*. Boston: Springer Science and Business Media.

Ware, J., & Vik, T. (2009). Quality assurance of item writing: During the introduction of multiple choice questions in medicine for high stakes examinations. *Medical Teacher, 31*, 238–243.

Wells, C. S., & Wollack, J. A. (2003) An Instructor's Guide to Understanding Test Retrieved 21 June 2015 from Reliability. https://testing.wisc.edu/Reliability.pdf

Yaghmale, F. (2003). Content validity and its estimation. *Journal of Medical Education, 3*(1), 25–27.

Zhou, W. (2009). Teachers' estimation of item difficulty: What contributes their accuracy? In S. L. Swars, D. W. Stinson, & S. Lemons-Smith (Eds.), *Proceedings of the 31st annual meeting of the North American Chapter of the International Group for Psychology of Mathematics Education* Atalanta, GA: Georgia State University. Retrieved 21 July 2015 from http://www.academia.edu/640866/Proceedings_of_the_31st_Annual_Meeting_of_the_North_American_Chapter_of_the_International_Group_for_the_Psychology_of_Mathematics_Education

Application of Multiple Intelligence Theory in the Assessment for Learning

How-Yee Lai and San-Lim Yap

Abstract In this study, multiple intelligence theory was applied in the learning of chemistry at higher education level with the objective to assess how cognitive skills, in terms of understanding, remembering, applying and expanding knowledge, can be enhanced. In a class consisting of big group of students with diverse intelligences and preferred learning styles, it is often a challenge to educators to enhance students learning in an equally effective manner to reach the desired learning outcome. The effectiveness of teaching–learning process may be optimized if educators have knowledge on the multiple intelligence (MI) profiles of their students and apply appropriate strategies in their mode of instruction which may result in improved learning outcome. At the start of the study, students taking the organic chemistry course participated in the MI assessment based on Dr Howard Gardner's theory of MI and were then surveyed via a questionnaire on their learning ability of the subject area. Thereafter, various strategies were used throughout the teaching of the course. The strategies included lecture, question-and-answer, problem-solving, video, collaborative, mind map, assignment presentation, individual work, e-learning and hands-on experimentation. At the end of the course, students were surveyed again in a second questionnaire to investigate responses on the learning outcomes. The trend of the MI profile of students from 2011 to 2015 was also investigated. Findings on the study indicated that application of MI theory enabled a more structured strategy that enhances students' cognitive skills. The trend of MI profile showed a shift from intrapersonal dominance in the previous years 2011–2013 to interpersonal, kinaesthetic and naturalistic intelligences in current years 2014–2015.

Keywords Multiple intelligence · Chemistry · Teaching–learning · Higher education

H.-Y. Lai (✉) · S.-L. Yap
Taylor's University, Selangor, Malaysia
e-mail: HowYee.Lai@taylors.edu.my

S.-L. Yap
e-mail: SanLim.Yap@taylors.edu.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_36

1 Introduction

Multiple intelligence (MI) theory has received much attention and has been adapted into teaching and learning of various curricula, since it was first coined by Gardner (1983). Gardner defines intelligence as 'the capacity to solve problems or to fashion products that are valued in one or more cultural setting". (Gardner and Hatch 1989). The theory is based on the philosophy that people are intelligent in different areas and every individual has a combination of a few intelligences (Goodnough 2001). There are eight intelligences, namely visual/spatial, verbal/linguistic, musical/rhythmic, logical/mathematical, bodily/kinaesthetic, interpersonal, intrapersonal and naturalistic. It has been assessed that the strength and development of the intelligences vary with each individual as a result of unique background in education, social and culture (Standford 2003). All intelligences are needed to function productively in society (Brualdi 1996).

Traditional education has focused on the verbal linguistic and logical–mathematical intelligences, for example teaching using lectures and assessments based on written examinations and assignments. Students who are weak in either of these intelligences are usually disadvantaged in class. Hence, educators should broaden their instructional and assessment strategies to draw on many intelligence (Goodnough 2001). As students do not learn in the same way, they should not be taught and assessed in a uniform manner. It is important that a teacher realizes the 'intelligence profile' of his students in class that will enable him to make more informed decisions on how to teach, present information and evaluate the students' progress (Lazear 1992). It has been observed that unsuccessful, unmotivated students have experienced academic growth when exposed to multifaceted interventions and techniques principled by MI theory (Stanford 2003). In essence, MI theory offers educators a common sense framework with which to explore about their students' ability in learning and consequently structure learning experiences with the ultimate aim to achieve more effective outcomes. Similarly, it motivates students to take charge of their learning which may lead to enhanced rate of achievement.

This action research study is conducted to test the following hypothesis:

1. Application of MI theory will enhance cognitive skills in understanding, remembering, applying and expanding the knowledge in a tertiary organic chemistry course
2. There is a shift in the MI profile of undergraduate students over a period of 5 years.

Hence, the objectives are to assess the development of students' cognitive skills with the incorporation of MI theory at the tertiary level and to investigate the trend in the dominance of specific intelligences of students enrolled in science-based programmes over a period of 5 years.

2 Method

In the first study to assess the development of students' cognitive skills with the incorporation of MI theory, the students enrolled in organic chemistry were selected. A total of 99 students (87 students from class of 2012 and 12 students from class of 2011) participated in this study. The students generated their MI profile using the online test via Website http://www.bgfl.org/bgfl/custom/resources_ftp/client_ftp/ks3/ict/multiple_int/index.htm (Birmingham City Council 2002–2011) (Fig. 1).

Subsequently, a first questionnaire was given to each student at the 3rd week of the semester with the objective to understand the students' academic background and attitude towards learning of organic chemistry. Students were asked to respond whether they 'strongly agree', 'agree', 'disagree' or 'strongly disagree' to the following questions:

I generally put a lot of effort into my studying.
I often have difficulty in remembering concepts and subject content.
I often find difficult to understand parts of the subject topic.
I often have difficulty when working on problem-solving questions.

During the 1st to 3rd week of the semester, teaching strategies were conducted in traditional lectures styles and occasional problem-solving questions during tutorials. Thereafter, various strategies were used throughout the teaching, beginning from week 3 until week 13. The strategies included lecture, question–answer sessions, problem-solving sessions, video, collaborative discussion in small groups, mind map, assignment presentation, individual work, e-learning (online forum, lecture capture) and hands-on experimentation. At the end of the course, students were surveyed again in a second questionnaire to investigate their responses towards the

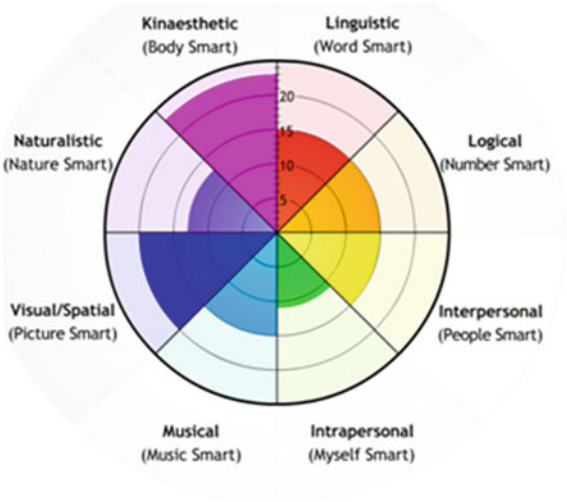


Fig. 1 Example of a result of a multiple intelligence (MI) profile of student

uses of various teaching strategies in the enhancement of learning. In the second questionnaire, students were asked to list the most obvious strategy that has helped them in the following cognitive skills:

(A) I can understand the topic better through this approach (understanding).
(B) I can remember the facts of the topic better after this approach (remembering).
(C) I have improved on applying concepts (applying).
(D) This approach has expanded my knowledge to a bigger frame that enabled to appreciate the topic and its application in the real world (expanding).

In the second study to investigate the trend of the MI profile of students, data on the MI profiles were taken from a total of 457 first semester students enrolled in a few science-related programmes from Taylor's University. Over the five-year period, the number of students that contributed to the data on MI profile was 127 (2011), 87 (2012), 67 (2013), 57 (2014) and 119 (2015).

3 Results and Discussion

In this study, multiple intelligence (MI) theory was used as the framework to design and structure teaching–learning strategies for first-year undergraduate students in an organic chemistry course. The MI profile data were a compilation of each student's highest ranking intelligences. In cases where there were two equally high ranking intelligences, both intelligences were recorded. Results of the MI profile as shown in Fig. 2 were obtained from 87 students in the class of 2012 and 12 students from class of 2011. Students from both classes showed very different profiles. The MI assessment for 2012 class showed a fair distribution of all eight intelligences (10.3–17.3 %) where the highest was observed for intrapersonal (17.3 %) followed by visual spatial (14.7 %). On the other hand, students of class 2011 showed

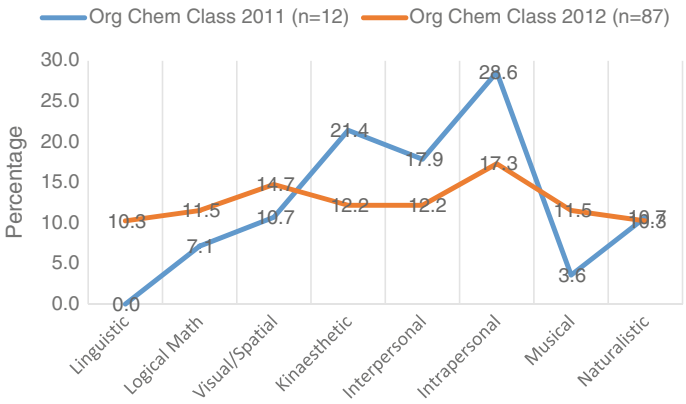


Fig. 2 Multiple intelligence (MI) profile of students in organic chemistry classes

significantly higher intrapersonal (28.6 %) > kinaesthetic (21.4 %) > interpersonal (17.9 %) > visual/spatial ≈ naturalistic 10.7 % > logical/mathematics (7.1 %) > musical (3.6 %) > linguistic (0).

The author has normally taught chemistry by the traditional lecture style method with occasional problem-solving questions during tutorial sessions. In this study, the author expanded the teaching strategies by addressing the students' MI profile to associate with some assumed preferred styles of learning. The targeted outcome for the teaching–learning strategies was to attain enhanced cognitive and knowledge in the subject. The strategies used were as follows: (1) lecture, (2) question-and-answer sessions with questions phrased to provoke thinking, e.g. 'What is produced …' 'How does it react …' 'Why does this occur?', (3) problem-solving exercises, (4) relevant video movies and audio with voice narration or music, (5) group collaborative work in 3–4 per group, (6) mind-mapping, (7) assignment with presentation, (8) individual work, (9) e-learning (online forum, lecture capture) and (10) hands-on experimentation.

An initial survey was carried out to understand the students' learning ability in studying the subject. Results showed 81.8 % of the students agreed (and strongly agreed) that they put in a lot of effort in studying the subject (Fig. 3). However, the students were mixed in agreeing or disagreeing whether they have difficulty in remembering the concepts, understanding and on their ability to work on problem-solving (45.5–54.6 % in total agreement or disagreement).

Figure 4 illustrates the responses of the students with different dominance of MI on the preferred teaching strategies in the enhancement of knowledge, based on their responses to the questions in the second questionnaire.

It is apparent from Fig. 4 that 'problem-solving exercise' was unanimously selected by all eight dominance as the most obvious strategy to enhance knowledge in all four manners, i.e. understanding, remembering, applying and expanding. In addition to problem-solving exercises, other strategies preferred by each dominance seemed to correlate with the MI dominance. For example, it was anticipated

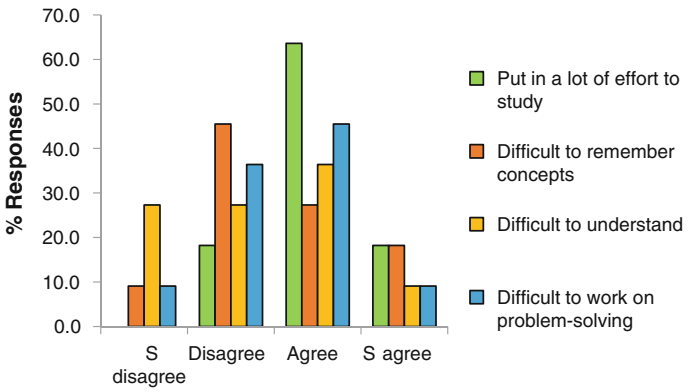


Fig. 3 Initial survey on students' learning ability in subject area

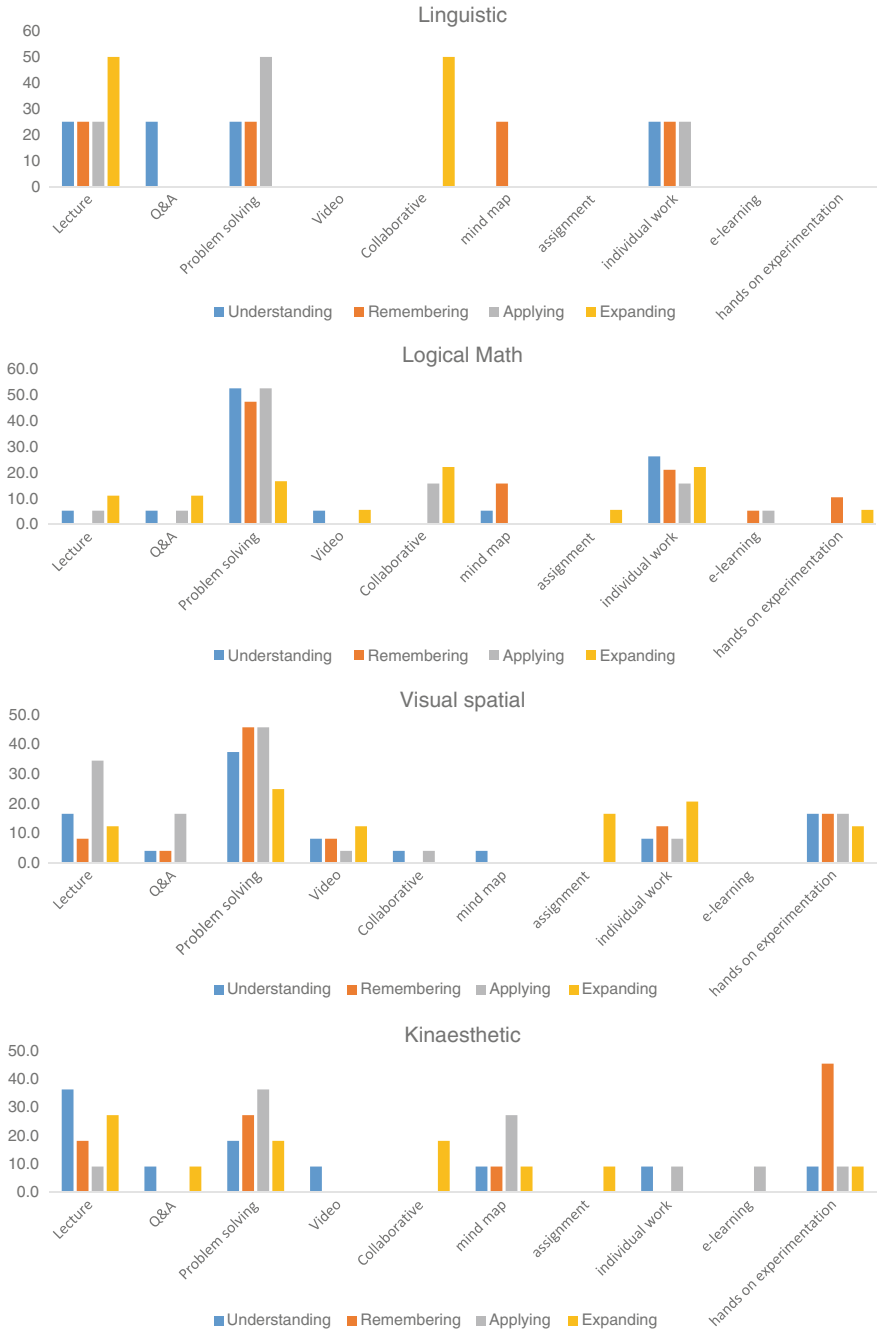


Fig. 4 Responses of students of various intelligence dominance on the preferred strategies for understanding, remembering, applying and expanding knowledge

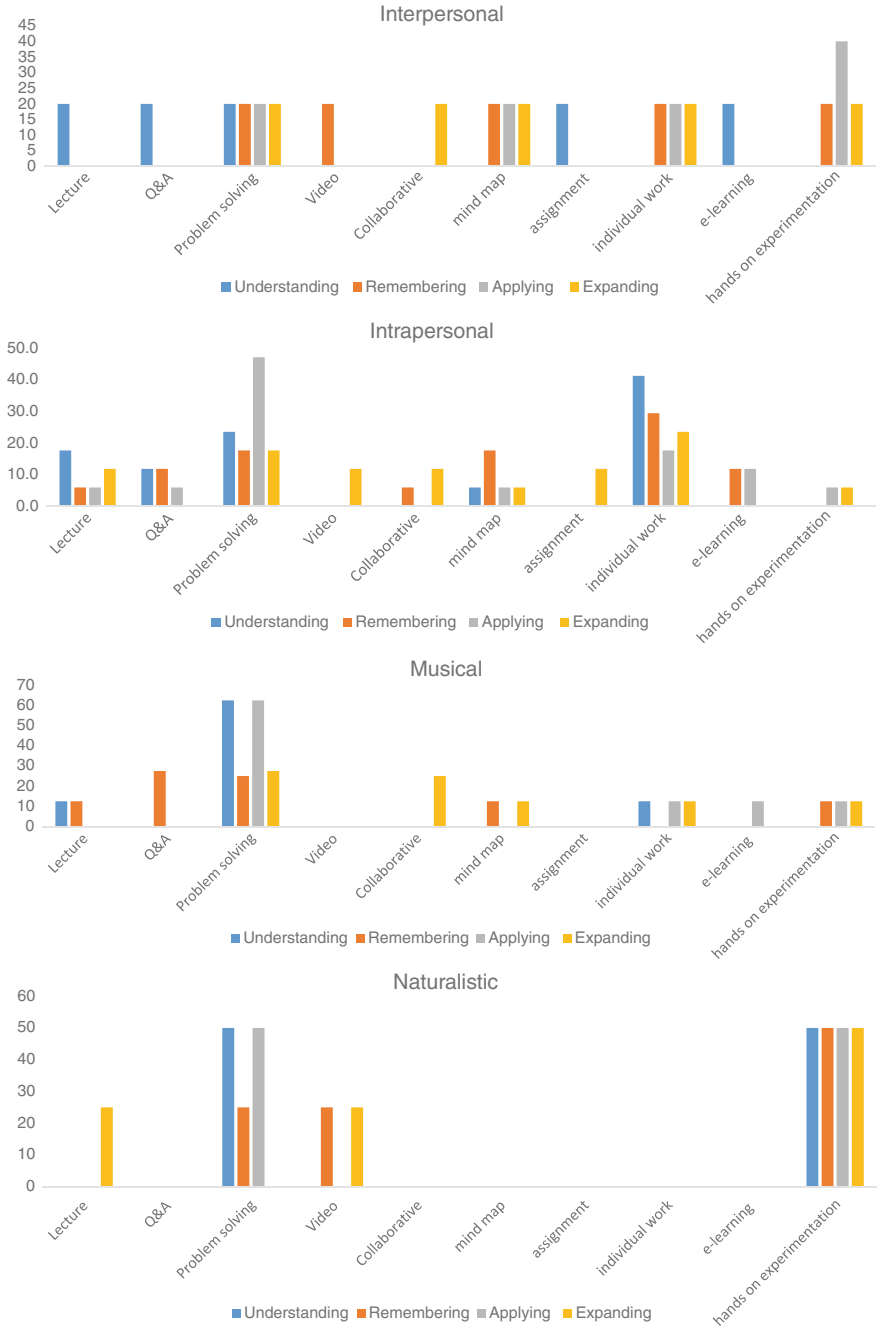


Fig. 4 (continued)

intrapersonal students should work well through self-reflection activities and individual works. The results from Fig. 4 (intrapersonal) revealed that 29–41 % of the intrapersonal students felt that the individual work strategy enabled them to remember and understand, while 47 % of the students responded to problem-solving strategy that enabled them to apply. The other strategies responded by these students were lectures, mind-mapping and e-learning (Fig. 4—intrapersonal). Conversely, students with linguistic dominance (Fig. 4—linguistic) responded well to 'lectures' (25–50 %), kinaesthetic dominance (Fig. 4—kinaesthetic) responded to hands-on experimentation in helping them to remember (46 %), and naturalistic dominance prefers hand-on experimentation work to enable them to understand, remember, apply and expand (25–50 %).

The previous studies and reviews on MI theory have placed much emphasis into the practice by primary and secondary schools, i.e. at elementary education (Stanford 2003; Sharifi 2008; Kelly and Tangney 2006), and the impact on mathematics (Isik and Tarum 2009; Sulaiman et al. 2010), science (Goodnough 2001) and e-learning (Krishnasamy et al. 2003). On the other hand, there is relatively limited report on the incorporation of MI theory into teachings of higher education despite the increase in student population and diversity in the tertiary education (Barrington 2004). Compared to the teaching pedagogies at the elementary level, teaching and learning in tertiary institution are relatively still conservative and teacher-centred which focused on language as the major mode of instruction such as lectures and assignment-based pedagogies. The focus on language as the major mode of communication meant that other modes, such as visual, bodily kinaesthetic or logical mathematics modes, were neglected. Some lecturers of tertiary institutions, although found MI to be relevant to higher education, have reservations to incorporate the theory in their teaching (Barrington 2004). This may be due to the limited time and resources available for university instructors to develop MI into each module. Besides, some modules have a set curriculum which does not allow flexibility for instructors to design their own curriculum and mode of instruction (Barrington 2004). In addition, the philosophy of higher education in terms of the mode of instruction is primarily based on facilitation of knowledge to matured students. Therefore, with the mindset that matured students should practise independent learning, university lecturers are less inclined to carry out multiple activities during the classes to cater for the diversity in intelligence amongst students.

It has been widely acknowledged that students learn differently as they process and present knowledge in different ways (Kelly and Tangney 2006). Findings from this study indicated that there may be correlations between the students' MI profile with preferred learning strategies. Previous studies have also suggested that students learn more effectively when instruction is adapted to the way they learn (Rasmussen and Davidson-Shivers 1998; Krishnasamy et al. 2003; Sulaiman et al. 2010). Hence, MI theory provides opportunities to educators to arrange appropriate strategies that value the different levels of intelligence and individual learning characteristics of students (Isik and Tarum 2009). However, it is noted here that the quantum of some strategies used may vary depending on the MI profile of the class. Intrapersonal intelligent students tend to prefer learning through self-reflective

thinking processes as necessitated during lectures, problem-solving and questioning–answering and individual exercises. Likewise, classes with students with kinaesthetic intelligence may be given more activities, e.g. mind-mapping and hands-on experimentation works. Collectively, these strategies would empower students with stronger grasp of knowledge and cognitive skills. Nevertheless, problem-solving exercises' strategy does not seem to correlate with any MI dominance as it is preferred by all dominance. Hence, this strategy should be used in all classes irrespective of the profile.

Subsequent investigations on the MI profiles of science-based students from 2011 to 2015 are shown in Fig. 5. It is clearly shown that all classes have various profiles which necessitate the educator to apply varying strategies to each class for effective enhancement of knowledge. Nevertheless, there seems to be an apparent change in trend from high intrapersonal dominance (17–29 %) in the previous years of 2011–2013 to a reduced 14 % in the recent years 2014–2015. However, there are obvious increases in the MI profile for interpersonal, visual spatial and naturalistic dominances from those in 2011–2013 to the current years 2014 and 2015. The change in trend of MI profile over the years may likely be attributed to the rise in the use of gaming and internet resources as the generation goes into the digital era. This finding may necessitate the change in the strategies used by instructor, perhaps to various forms that engage these intelligences, for effective learning.

In conclusion, knowledge of MI profile of students enables educators to effectively enhance students learning. Incorporation of multiple intelligence theory into the teaching of a course enables a more student-centred approach that caters to specific students' learning needs, hence effecting positive outcomes on students' learning towards the subject. There may be a shift in the trend of the MI profile from intrapersonal to interpersonal, kinaesthetic and naturalistic intelligences in recent years which will necessitate a shift to related strategies in the current and near future.

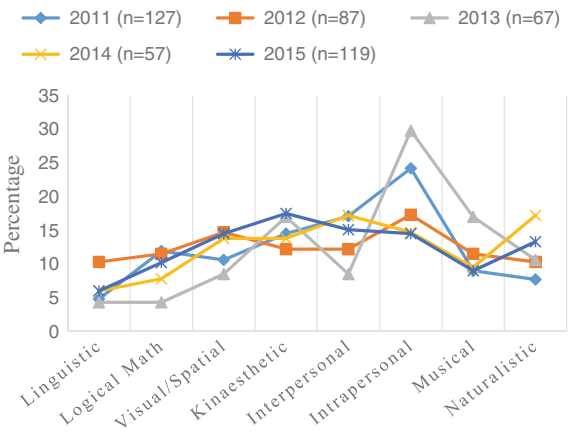


Fig. 5 Multiple intelligence profile of students enrolled in science-related programmes over a five-year period from 2011 to 2015

References

Barrington, E. (2004). Teaching to student diversity in higher education: How multiple intelligence theory can help. *Teaching in Higher Education, 9*(4), 421–434.

Birmingham City Council. (2002–2011). Multiple intelligence. Website: http://www.bgfl.org/bgfl/custom/resources_ftp/client_ftp/ks3/ict/multiple_int/index.htm. Accessed 10 May 2011.

Brualdi, A. C. (1996). Multiple intelligences: Gardner's theory. In *ERIC clearinghouse on assessment and evaluation* (pp. 1–6). Washington, DC: ERIC Digest.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books Inc.

Gardner, H., & Hatch, H. (1989). Multiple intelligence go to schools: Educational implications of theory of multiple intelligences. *Educational Researcher, 18*(8), 4–9.

Goodnough, K. (2001). Multiple intelligences theory: A framework for personalizing science curricula. *School of Science and Mathematics, 101*(4), 180–193.

Kaya, O. N. (2008). How is a science lesson developed and implemented based on multiple intelligences theory? *Hacettepe University Journal of Education*, *34*, 155–167.

Isik, D., & Tarum, K. (2009). The effects of the cooperative learning method supported by multiple intelligence theory on Turkish elementary students' mathematics achievement. *Asia Pacific Education Review, 10*, 465–474.

Kelly, D., & Tangney, B. (2006). Adapting to intelligence profile in an adaptive educational system. *Interacting with Computers, 18*, 385–409.

Krishnasamy, K., Lee, S. P., & Palaniappan, S. (2003). Effective design of E-learning application incorporating multiple intelligences. *Lecture Notes in Computer Science, 2911*, 693–694.

Lazear, D. (1992). Teaching for multiple intelligences. Fastback 342 Bloomington. In *Phi Delta Kappan Educational Foundation*.

Rasmussen, K. L., & Davidson-Shivers, G. V. (1998). Hypermedia and learning styles: Can performance be influenced? *Journal of Multimedia and Hypermedia, 7*(4), 291–308.

Sharifi, H. P. (2008). The introductory study of Gardner's multiple intelligence theory in the field of lesson subjects and the students' compatibility. *Quarterly Journal of Education Innovations, 24*, 11–20.

Stanford, P. (2003). Multiple intelligence for every classroom. *Intervention in School and Clinic, 39*(2), 80–85.

Sulaiman, T., Abdurahman, A. R., & Rahim, S. S. A. (2010). Teaching strategies based on multiple intelligence theory among science and mathematics secondary school teachers. *Procedia-Social and Behavorial Sciences, 8*, 512–518.

Enhancing Oral Presentation Skills of ESL Students: The Use of Oral Communication Strategies

Hairuzila Idrus

Abstract Oral presentation skill is an asset for graduates. A graduate with good oral presentation skills has a better chance in career advancement and promotion rather than one who does not. Thus, it is vital for university students to have exposure to oral communication strategies for them to become good presenters. The purpose of this study was to identify the communication strategies used by undergraduate students while delivering oral presentations in an oral communication skills course. The strategies used were mapped against the grades they obtained for the course to gauge the strategies used by students who scored an A (good presenters) versus those who scored a B (average presenters). Presenting to an audience may generate anxiety and concern that the presentation might not be delivered effectively. Good presenters have found ways to overcome or cope with this concern. Thus, the Oral Communication Strategies Inventory (OCSI) developed by Nakatani (Mod Lang J 90:151–168, 2006) was used to access strategies that the students used for coping while delivering their oral presentations. The OCS inventory included eight categories of strategies for coping while speaking and seven categories of strategies for coping while listening. However, for the purpose of this paper, only speaking strategies were analyzed. The respondents were 100 engineering students enrolled in a communication skills course. Results of the data analysis revealed that the good presenters employed more oral communication strategies than the average presenters. The better students had more frequently used social affective, fluency-oriented, and nonverbal strategies for coping with speaking problems. The average students, on the other hand, do utilize social affective and nonverbal strategies but on a much lesser degree. The findings suggest that average presenters should be made aware of the mentioned skills and be accustomed to them in order to enhance their oral presentation skills. As such, introducing the oral communication strategies to the students could help them improve their oral communication delivery.

Keywords Communication strategies · Oral presentation skills · ESL learners

H. Idrus (✉)
Department of Management & Humanities, Universiti Teknologi PETRONAS, Seri Iskandar, Perak, Malaysia
e-mail: hairuzi@petronas.com.my

© Springer Science+Business Media Singapore 2016

S.F. Tang and L. Logonnathan (eds.), *Assessment for Learning Within and Beyond the Classroom*, DOI 10.1007/978-981-10-0908-2_37

1 Introduction

The importance of oral communication skills in mastering a second language is emphasized by Ellis (2003) when he pointed out that second and foreign language acquisition involves the ability to use the sound and grammar systems to communicate meanings. As emphasized by Lazaraton (2002), one of the main goals in learning a language, particularly a second or foreign language, is to be able to communicate effectively in the language. Oral communication means communicating orally in a manner which is clear, fluent, and to the point, and which holds the audience attention, both in group and in one-to-one situations. In second language learning, one of the biggest difficulties for the learners to improve is the lack of self-confidence in using the target language. This is especially so for ESL learners from Asia (Chuanchaisit and Prapphal 2009; Metcalfe and Noom-Ura 2013; Nakatani 2012; Norafini and Sarjit 2014; Peng 2007). They tend to be very reticent when it comes to communicating or expressing their thoughts and ideas in the target language. This study focuses on ESL learners in Malaysia. Oral communication skills are essential for professionals like engineers and technologists who aspire to carry out professional practice in the global arena.

The inability to communicate well in English has been named as one of the causes of unemployment among university graduates. In August 2004, a Malaysian leading newspaper reported the government's concern about the increasing number of unemployed graduates, many of who lacked communication skills (Sibet 2005). Sibet also points out that a survey by JobStreet, a Malaysian employment agency, found that weak in English as the most prominent factors (56 %) for graduates' unemployment in Malaysia. Another leading Malaysian newspaper stated that a government survey has revealed that of nearly 60,000 unemployed graduates, many could not get jobs largely, due to poor English and communication skills. Thus, there is ample evidence that these graduates lack the required standard of oral communication skills, particularly when compared to the needs of industry internationally (Riemer 2002). This indicates the importance of the ability to communicate effectively in English in Malaysian as well as global context. One of the communication abilities that graduates are required to possess is oral presentation skills. A study conducted by Darling and Dannels (2003) on the role of oral communication at workplace found that the ability to give formal presentations was critical to the job.

In view of the oral communication problems, this study is conducted to find out the oral communication strategies used by ESL undergraduates when delivering technical and persuasive presentations in professional communication skills (PCS) course. It is important to know the strategies used particularly by the high-ability students so that the other students could be exposed to the effective strategies as well.

2 Studies on Communication Strategies

The use of communication strategies in communication by second or foreign language learners was suggested by a well-known linguist, Selinker (1972), as one of the five central processes for interlanguage. Interlanguage is a term he introduced to refer to a situation where language learners, in their attempt to communicate in meaningful situations, would produce imperfect language which contains features of their native and target languages. Thus, to assist language learners to develop their communicative competence, Canale and Swaine (1980) suggested the introduction of communication strategies. Varadi (1983) conducted a systematic analysis of communication strategies by introducing several taxonomies and terms used in research in communication strategies. Researchers in 1980s, 1999s, as well as in the 2000s, generally see communication strategies as systematic, communicative devices employed to handle difficulties in communication and to avoid communication breakdown (Canale 1983; Corder 1981; Dornyei 1995; Long 1993; Nakatani, 2006; Yule and Tarone 1991).

Communication strategies have then been categorized into two main types which are achievement or compensatory strategies and reduction or avoidance strategies (Dornyei and Scott 1997; Nakatani 2006). These two broad categorizations echo the underlying behavior of the learners when they are faced with communication problems. When in this critical situation where they need to make fast decisions to achieve their communicative goals, learners may choose to address the problem directly, resulting in the achievement strategies, or to avoid the problem which leads to the use of reduction strategies. In using the achievement strategies, the communicative goals remain the same, only that the learner will formulate an altered plan to achieve the goals. On the other hand, if the learner chooses to use the reduction strategies, it involves the changing of the goal or possibly abandoning it totally.

Nakatani (2006) has developed an inventory to assess the communication strategies of language learners or coping with speaking problems and listening problems based on these two categories. This inventory was recently developed; thus, it is used in this study. Furthermore, this inventory was originally developed with Asian students as the respondents which are of similar background to the majority of respondents in this study.

The descriptions of the strategies for coping with speaking problems are as follows:

1. Social affective strategies involve learners' affective factors in social contexts.
2. Fluency-oriented strategies are related to fluency of communication.
3. Negotiation for meaning while speaking strategies are relevant to the participants' attempts to negotiate with their interlocutors.
4. Accuracy-oriented strategies are concerned with a desire to speak English accurately.

5. Nonverbal strategies while speaking require using eye contact, gestures, or facial expressions to give hints and to help the listener guess the intended meaning.
6. Attempt to think in English strategies involves thinking as much as possible in the foreign language during actual communication.
7. Message reduction and alteration strategies involve avoiding a communication breakdown by reducing an original message, simplifying utterances, or using similar expressions that can be confidently used.
8. Message abandonment strategies are associated with message abandonment by learners in communication.

Studies on the oral communication strategies employed by Malaysian students are still lacking. Most of the studies are on communication strategies used in communication activities in general. For instance, Ting and Phan (2008) conducted a study on communication strategies and proficiency, while Raed et al. (2013) studied the communication strategies of Malaysian students at tertiary level. Using this similar instrument, OCSI, Norafini and Sarjit (2014) conducted a study on the strategies used by diploma students in an institution in Malaysia in coping with speaking and listening problems. Halizah et al. (2012) carried out an interesting study on the communication strategies used by the ESL learners in an online discussion via Facebook. Tan et al. (2012) conducted a study on communication strategies in Malaysia, but the focus was on EFL students, specifically Arabic students, studying in a Malaysian university. This study on the oral communication strategies of ESL learners in oral presentations is to add to the body of the literature on communication strategies in the Malaysian context and on Malaysian students.

3 Research Questions

This study is guided by two research questions as follows:

1. What are the types of oral communication strategies used by the students when delivering oral presentations?
2. What are the differences in the types of communication strategies used by the students of different abilities in oral presentation?

4 Research Instrument

The Oral Communication Strategies Inventory (OCSI) developed by Nakatani (2006) has been used in this study which focuses on communication strategies for listening and speaking. The OCSI is a set of questionnaires that focus on two parts: first, on the strategies in coping with speaking problems and second, strategies on

coping with listening problems. The former part was used for the current research purpose.

Nakatani (2006) listed eight strategies in coping with speaking problems which are social affective, fluency-oriented, negotiation for meaning, accuracy-oriented, nonverbal, attempt to think in English, message reduction and alteration, and message abandonment. The questionnaire consists of 32 items that address these coping strategies. Each item in the questionnaire is followed by a 5-point Likert scale whereby 1 indicates "never or almost never true of me," while 5 indicates "always or almost always true of me."

Cronbach's alpha was used to measure the reliability coefficient that assessed the consistency of the questionnaire. The alpha level for the items was 0.89. According to Hair et al. (1998), the generally acceptable lower limit for Cronbach's alpha is 0.70. As such, the consistency of the questionnaire is highly acceptable. The questionnaire responses were analyzed using the Statistical Package for Social Sciences (SPSS) version 20.0.

5 Research Setting and Respondents

The respondents for this study were 100 students from a private university in Malaysia, enrolled in a communication course named professional communication skills (PCS). The students were in their second year ($n = 63$) and third year ($n = 37$) of study and majoring in engineering and technology. The respondents were all ESL learners, 83 from Malaysia, while 17 were international students. About 64 % of the respondents were males and 34 % were females. The respondents completed the survey questionnaires in class voluntarily. They were asked to indicate the extent to which they felt that each item/statement described the ways of coping with speaking problems when they were delivering their presentations. It is worth noting here that the respondents have never been taught oral communication strategies in this course; however, they were taught effective delivering strategies. In PCS course, the students were required to do two individual presentations and two group presentations. In the questionnaire, the respondents were required to indicate their names so that the grade that they received for the course at the end of the semester could be mapped against their responses. This is to get the answer for research question 2.

6 Results

This section is divided into two sections, according to the research questions addressed in this study.

6.1 Types of Oral Communication Strategies Used

In answering the first research question, what are the types of oral communication strategies used by the students when delivering oral presentations, the scores were obtained through calculating the mean score for the items in the scale. The results reveal that social affective strategies have the highest mean score (M = 3.86), which indicates that this strategy is used most by the students. The next highest mean score is nonverbal strategies (M = 3.85), followed by fluency-oriented strategies (M = 3.69), accuracy-oriented strategies (M = 3.67), negotiation of meaning strategies (M = 3.58), message reduction and alteration strategies (M = 3.57), attempt to think in English (M = 3.16), and lastly message abandonment strategies (M = 2.71). In analyzing the mean scores, the interpretation was adapted from Landell (1997) whereby a mean of 1.00–2.33 signifies low usage of strategy, 2.34–3.67 average usage, and 3.68–5.00 high usage. The result is summarized in Table 1.

6.2 Types of Strategies Used Based on Grades

The second research question attempts to find out the differences in the types of communication strategies used by the students of different abilities in oral presentation. The results found that the grades obtained by the respondents in this study ranged from level high to average, where high grades are those who scored "A" and "A−," while average are the ones who scored "B+" and "B" in the course. These grades are the grades obtained by majority of the students enrolled in this course as the medium of instruction in this private university is English. Thus, the students are familiar with English language. Out of the 100 respondents, 56 are in the category of high grades and 44 in the average grades (refer to Table 2).

An analysis on the oral communication strategies employed by the students while delivering oral presentations in the PCS class based on the grades obtained reveals that there was not much difference in terms of the communication strategies

Table 1 Mean, standard deviation, and rank of strategy used in coping with speaking problems

Strategies	Mean	Std. dev	Rank	Level of usage
Social affective strategies	3.86	0.62	1	High
Nonverbal strategies	3.85	0.88	2	High
Fluency-oriented strategies	3.69	0.59	3	High
Accuracy-oriented strategies	3.67	0.65	4	High
Negotiation for meaning strategies	3.58	0.72	5	Average
Message reduction and alteration strategies	3.57	0.66	6	Average
Attempt to think in English	3.16	0.88	7	Average
Message abandonment strategies	2.71	0.87	8	Average

Table 2 Summary of grades obtained by respondents in PCS course

Grades	No. of respondents	Level
A	12	High
A−	44	High
B+	23	Average
B	21	Average

Table 3 Ranking of oral communication strategies employed by students of high and average grades

Strategies	Grade high	Mean score	Grade average	Mean score
Social affective strategies	1	3.96	1	3.80
Nonverbal strategies	2	3.90	2	3.76
Fluency-oriented strategies	3	3.80	4	3.67
Accuracy-oriented strategies	4	3.66	5	3.63
Negotiation for meaning strategies	5	3.65	6	3.54
Message reduction and alteration strategies	6	3.56	3	3.68
Attempt to think in English	7	2.95	7	3.34
Message abandonment strategies	8	2.62	8	2.83
Fluency-oriented strategies	3	3.80	4	3.67

they employed to cope with speaking problems when delivering their oral presentations. The main difference is on the degree of usage of the strategies. However, an important finding in this analysis is that among the top three strategies highly employed by the average students (scored B+ and B in PCS) were the message reduction and alteration strategies, which, for the high-ability students (scored A and A− in PCS), these particular strategies were ranked as number six. The summary of the ranking is shown in Table 3.

The following section will further discuss the results.

7 Discussion

The communication strategies highly employed by the respondents in this study, which were social affective strategies, nonverbal strategies, fluency-oriented strategies, and accuracy-oriented strategies, are categorized as achievement or compensatory strategies (Nakatani 2006). These strategies are used by speakers to expand their linguistic resources to achieve communicative goals. According to Corder (1983), in the ultimate concept of oral communication, communication strategies are important in terms of the relationship between the means and the ends

of communication. It is usually expected that the speakers have enough linguistic knowledge to express the message; however, second language learners might have limited linguistic resources that may not permit them to express successfully. The respondents in this study decided to use the achievement strategies due to their linguistic ability. Based on the grades they obtained for this course, it can be concluded that they are not low-ability students; thus, they are willing to take the risk in coping with problems they face while delivering oral presentations in class in order to achieve their goals, which is to deliver their presentations effectively.

In PCS class, the students were taught to deliver their presentations extemporaneously, which means the students should plan, prepare, and rehearse for the presentation before delivering the actual presentation. As such, by preparing for the presentation, the delivery should seem spontaneous and effortless. This could be one of the factors that explain their high and average grades, as well as the risk-taking strategies they employed as they are well aware of the contents of the presentations beforehand.

The results of this study confirm the findings of other studies (Chen 2009; Chuanchaisit and Prapphal 2009; Metcalfe and Noom-Ura 2013; Nakatani 2006) that high- and average-ability students frequently employed the risk-taking strategies. However, this study reveals that there is a difference between the high- and average-ability students where the average students highly used the reduction or avoidance strategies, which is message reduction and alteration. It is also worth noting that for both groups of respondents, the message abandonment strategies are the least utilized by them. This may be due to the nature of the task which is oral presentation. The evaluations for this course, professional communication skills, are four oral presentations; thus, the students must be well prepared to present as they are being evaluated. With this in mind, the students could not afford to abandon their goal for the oral presentation because in doing so, they would not be able to present their message clearly and thus would result in a lower grade.

8 Impact of Study on Teaching and Learning

Even though the respondents in this study are proficient in the language, there is still room for improvement in their communication competence as well as oral presentation ability. It is often the case that learners are unaware of the fact that they need to have strategies in their oral communication. Hence, it is recommended that communication strategies be included in the syllabus of this professional communication skills course. Numerous studies have been done on the teaching of communication strategies in language classes. These studies found that introducing communication strategies to language learners will help them to improve their communication abilities (Lam 2010; Mirzaei and Heidari 2012; Metcalfe and Noom-Ura 2013; Nakatani 2006, 2012; Norafini and Sarjit 2014). Furthermore, Nakatani (2012) suggested that second and foreign language learners "should consciously use their interlanguage system to control their performance and to

maintain interaction. To achieve these goals, learners should have experience authentic task training through raising their awareness of managing and supervising specific strategy use" (p. 774). He further added that the success of learners in oral communication could be attributed to their conscious participation in communicative achievements through strategy training. On that note, it would be interesting to introduce oral communication strategies in the effort to further enhance the students' oral presentation abilities.

9 Conclusion

The objective of this study is to find out the oral communication strategies used by the students in professional communication skills course when coping with speaking problems while delivering their presentations in class. The findings highlighted that the strategies employed are in the category of achievement strategies which are frequently used by competent ESL or EFL learners. This indicates that the students in this particular higher institution are in the right track to become good or better and still effective presenters. Nonetheless, the average students did use the reduction or avoidance strategies which could interfere in their effort to improve their oral presentation skills. As such, it is recommended that oral communication strategies be introduced in the course syllabus in addition to delivering strategies to further assist students to become effective oral presenters. This is extremely important in the effort of this university to produce marketable graduates.

References

Canale, M. (1983). Communicative competence to communicative language pedagogy. In J. C. Rihards & R. Schmidt (Eds.), *English for cross-cultural communication*. New York: Longman.

Canale, M., & Swaine, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1–47.

Chen, H. W. (2009). *Oral communication strategies used by English major college students in Taiwan.* Unpublished master's thesis, Chaoyang University of Technology, Taichung, Taiwan. Retrieved Sept 30, 2014, from http://www1.lib.cyut.edu.tw/eThesys/index.htm

Chuanchaisit, S., & Prapphal, K. (2009). A study of english communication strategies of Thai university students. MANUSYA. *Journal of Humanities, 17*, 100–126.

Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.

Corder, S. P. (1983). Strategies of communication. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication*. New York: Longman.

Darling, A. L., & Dannels, D. (January, 2003). Practicing engineers talk abut the importance of talk: A report on the role of oral communication in the workplace. *Communication Education*, *52*(1), 1–16.

Dornyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly, 29*, 55–85.

Dornyei, Z., & Scott, M. L. (1997). Communication strategies in second language: Definitions and taxonomies. *Language Learning, 47*(1), 173–210.
Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
Halizah, O., Mohamad, A. E., & Melor, Md Y. (2012). Learners' use of communication strategies in an online discussion via facebook. *Social and Behavioral Sciences, 64*, 535–544.
Lam, W. Y. K. (2010). Implementing communication strategy instruction in the ESL oral classroom: What do low-proficiency learners tell us? *TESL Canada Journal*, *27*(2), 11–30 (Spring).
Landell, K. (1997). *Management by menu*. London: Wiley.
Lazaraton, A. (2002). A qualitative approach to the validation of oral language testing. In *Studies in language testing 14*. Cambridge: Cambridge University Press.
Long, M. H. (1993). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics, 4*, 126–141.
Metcalfe, J., & Noom-Ura, S. (2013). Communication strategy use of high and low roficiency learners of English at a Thai university. *LEARN Journal: Laguge Education and Acquisition Research Network, 6*(1), 68–89.
Mirzaei, A., & Heidari, N. (2012). Exploring the use of oral communication strategies by (Non) fluent L2 speakers. *The Journal of Asia TEFL*, *9*(3), 131–156 (Autumn).
Nakatani, Y. (2006). Developing an oral communication inventory. *The Modern Language Journal, 90*, 151–168.
Nakatani, Y. (2012). Exploring the implementation of the CEFR in Asian contexts: Focus on communication strategies. *Procedia-Social and Behavioral Sciences, 46*, 771–775.
Norafini, Z., & Sarjit, K. (2014). Oral English communication difficulties and coping strategies of Diploma of Hotel Management students at UiTM. *3L: The Southeast Asian Journal of English Language Studies, 20*(3), 93–112.
Peng, J. E. (2007). Willingness to communicate in the Chinese EFL classroom. A cultural perspective. In J. Liu (Ed.), *English language teaching in China: New approaches, perspectives and standards* (pp. 250–269). London: Continuum.
Raed, L. U., Nur Illianis, A., & Mohamad Jafre, Z. A. (2013). Study of the communication strategies used by Malaysian ESL students at tertiary level. *International Journal of English Language Education, 1*(1), 35–52.
Riemer, M. J. (2002). English and communication skills for the global engineers. *Global Journal of Engineering Education, 6*(1), 91–100.
Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language, 10*(3), 209–232.
Sibet, M. P. (2005). Leaping out of the unemployment line. Retrieved on July 23, 2012. http://www.calm.unimas.my/insite6/article_a.html
Tan, K. H., Nor Fariza, M. N., & Mohd Nayef, J. (2012). Communication strategies among EFL students—an examination of frequency of use and type of strategies used. *GEMA Online Journal of Language Studies, Special Edition, 12*(3), 831–848.
Ting, S. H., & Phan, G. Y. (2008). *Adjusting communication strategies to language proficiency Australian TESOL Journal, 23*(1), 28–36.
Varadi, T. (1983). Strategies of target language learner communication: Message adjustment. *IRAL, 18*(1), 59–71.
Yule, G., & Tarone, E. (1991). The other side of the page: Integrating the study of communication strategies and negotiated input in SLA. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), *Foreign/Second Language Pedagogy Research* (pp. 162–171). Clevedon: Multilingual Matters.