Designing an Instrument for Providing Better Student Feedback on Teaching Effectiveness

[1] mytee22@yahoo.com
University of Malaya,

[2] Inti College Subang Jaya, Malaysia

Tee Meng Yew [1], Fauziah K.P. Dawood Kannaki a/p S. Narayansany,
M Kamala a/p Palaniappa Manickam, Leong Siok Jen, Kuan Chin Hoay [2]

ABSTRACT

This article describes the process of designing a SET (student evaluation of teaching) instrument that is mapped against good practices in undergraduate education. Research has shown that teachers can significantly impact student learning. But how do teachers (and academic administrators) know when their teaching is effective? How do they go about acquiring feedback on their teaching? SET instruments have been widely utilized by universities and colleges as a form of summative and formative evaluation. Researchers— after reviewing fifty years of literature— say that there is little reason to doubt that SET can provide valid and useful information for both faculty members and administrators. The evidence suggests that students can make valid and reliable judgments about classroom teaching performance if asked the right questions. But there is no known instrument universally accepted by colleges and universities. In other words, each institution of learning has to develop its own contextually appropriate instrument, so that it is compatible with institutional needs and is aligned with good practices in teaching and learning. The description in this article is based on the experience and data collected by an action research project team tasked with developing a universally accepted instrument for measuring instructional effectiveness.

Keywords: *student evaluation of teaching; evaluating instructional effectiveness; student feedback*

An earlier version of this paper was presented at a conference and several discussion forums.

INTRODUCTION

The evaluation of teacher effectiveness has become essential in higher education and a widely used approach is through student feedback or ratings (Tagomori & Bishop, 1994; Murray, 2005; Pounder, 2007), often referred to in this domain of research as student evaluation of teaching or SET. Historically, the use of student evaluation was supported by three different stakeholders for distinct reasons: first, students who wanted to have a say in the quality of teaching; second, administrators who were concerned with accountability and tracking of faculty development; and third, faculty members who wanted their salary, promotion and tenure evaluations to depend on something other than the number of publications alone (Clark, 1993, cited in Koh & Tan, 1997; Murray, 2005). Whoever the stakeholder, SET instruments are fundamentally intended as a proxy for direct measurement of student learning. Measuring student learning, as educators have discovered, is not as simple as it sounds. Limited by the technical difficulties, SET instruments are a systematic attempt to do the next best thing by assessing teacher or course characteristics believed to contribute to student learning as observed by students (Murray, 2005).

SET can also be used for providing feedback to teachers for improvement, guiding teacher training and development, assessing teacher performance for administrative reasons, or even giving assurance of

effective classroom instruction to students (Bosetti, 1994, cited in Koh & Tan, 1997; Murray, 2005). Murray (2005) also pointed out that student evaluation of teaching can lead to improved teaching, particularly when student feedback is accompanied by expert consultation. The data from SET instruments can also help colleges and universities to improve areas such as student achievement and planning allocation of available funds for educational materials and development programs (Ngware & Ndirangu, 2005).

However, the quality of SET data can be severely compromised if the instrument is not designed properly in the first place. Tagomori and Bishop (1994), for example, have found severe design flaws in the numerous SET instruments that they analyzed. As the significance and the stakes are often high in relation to usage of SET data, the exercise of designing robust SET instruments has become all the more important.

Problem Statement

Much of the research done on student evaluation of teaching has focused on the issue of reliability and validity (Cashin, 1995; McKeachie & Kaplan, 1996; Murray, 2005; Tagomori & Bishop, 1994), but the literature is limited when it relates to the process of designing a contextually appropriate instrument compatible with institutional needs and aligned with research-based good practices in teaching and learning. This is what a group of academics in a private college in Malaysia found when faced with the challenge of revamping the institution's highly influential SET form. The description in this article is based on the experience and data collected by the core project team.

Literature Review

Research has shown that teachers can significantly impact student learning (Sanders & Rivers, 1996). But how do teachers (and academic administrators) know when their teaching is effective? How do they go about acquiring feedback on their teaching? SET instruments have been widely utilized by universities and colleges as a form of summative and formative feedback. McKeachie and Kaplan (1996), after reviewing fifty years of research literature, say that there is little reason to doubt that SET can provide valid and useful information for both faculty members and administrators. This has been found to be particularly so when students are asked the right questions (Tagomori & Bishop, 1994). In general, student rating tends to be statistically reliable, valid and relatively free from bias compared to any other data used including colleagues' ratings and administrators' ratings (Cashin, 1995). It has also been argued that students can provide meaningful feedback because students spend more instructional time with the teacher than anyone else and the students have a deep and unique understanding of how the course's or teacher's characteristics affect them (Tagomori & Bishop, 1994).

Perhaps more importantly, meaningful feedback from students provides valuable opportunities for improvement. For example, an analysis by Murray (2005) of data collected via three types of research— faculty opinion surveys, field experiments, and longitudinal comparisons—suggests that student feedback on teaching can help teachers make improvements. These improvements can become more significant if the feedback is accompanied with consultation by a teaching and learning expert.

However, the use of student evaluations of teaching is not without limitations. It fails to take into account extraneous influences that affect ratings but which are beyond the educator's control such as class size, student academic motivation, course/disciplinary difficulty, gender bias, rank of teacher, age of teacher and student expectation with regard to grade (Koh & Tan, 1997; Pounder, 2007; Worthington, 2002, cited in Rowley, 2003). Some faculty members have argued that SET can lead to grade inflation and lowering of academic standards (Murray, 2005). And yet others have argued and demonstrated that students are not qualified to judge specific aspects commonly evaluated in SET instruments such as the appropriateness of the instructor's objectives, the relevance of assignments or readings, the degree to which the subject content is balanced and up-to-date, and the degree to which grading standards are unduly lax or severe (Cashin, 1996). In spite of its limitations, as mentioned earlier, SET can still provide relatively useful information whether for summative or formative purpose.

As for the process of designing SET instruments, Tagomori and Bishop (1994) provided some key guidelines. They suggested that the process must involve these three elements: first, describing or defining

the characteristics of effective teaching; second, phrasing the questions by mapping against those characteristics; and third, selecting the most appropriate responses and scaling to each question. They also warned that the design and development process is long and requires careful planning and formulation. In addition, Rowley (2003) proposed five factors that require attention when designing a SET instrument. They are: balancing between questions on teaching, support, resources and learning; designing questions that do not assume a specific mode of delivery; promoting independent learning; reflecting student learning experience that could provide comments on their own learning experience; as well as protecting anonymity of the students to enable their honest feedback. However, none of these authors provide a processual or change process guideline in designing a SET instrument.

METHOD

The project team relied on action research methods (McNiff & Whitehead, 2002) to guide the SET instrument design process. The action research process was ideal for this project as there was a very specific focus to solve a real-world issue. It also allowed for a reflective as well as an active process of progressive problem solving led by individuals working with others in a community of practice to improve the instrument used to attain student ratings on teaching. The objective and challenges for the community of practice are detailed in the case report.

Different formats of community of practice discussion meetings and focus groups were held to collect data, which were then used as a platform for action, primarily working toward developing iterations of a SET instrument. The working draft SET instrument was then shared with all academic staff to elicit feedback during several sessions of focus groups and over email. After incorporating the feedback, the final draft was submitted to the institution's senior management group for review.

Case Report

The project duration was about eight months, beginning with developing a better understanding through a literature review related to SET instrument development and best practices in teaching and learning, and eventually culminating in completing the task of designing a new SET instrument to be proposed to and reviewed by the institution's senior management team. The members of the community of practice met regularly, usually every fortnight with actionable tasks being completed in between meetings. Participation in the community was usually above and beyond the member's usual work responsibilities.

Background and Project Objectives

More information has been produced in the last 30 years than in the previous 5,000 years combined (Lyman & Varian, 2004; Wurman, 1989). As the world's information-base continue to grow, both nations and organizations today are seeking those individuals—Peter Drucker calls them knowledge workers—who can make sense and make use of existing knowledge, and use it as springboard to pursue and create new knowledge. Unfortunately, the classrooms of today are all too often focused on knowledge transmission with not enough time spent on creating learning conditions that promote deep learning (Herrington & Oliver, 2000; Marton & Saljo, 1976; Resnick, 1987; Whitehead, 1929). Education systems around the world are responding to this need for change. Singapore, for example, has overhauled its education system so that its students can be better prepared to be knowledge workers. Malaysia is also embarking on this change, as planned and documented in its *National Higher Education Action Plan 2007-2010* (Ministry of Higher Education, 2007a) and *National Higher Education Strategic Plan* (Ministry of Higher Education, 2007b).

One of the seven thrusts of Malaysia's National Higher Education Strategic Plan (Ministry of Higher Education, 2007b) is to improve teaching and learning through greater active and problem-based learning pedagogies. Similarly, internal and external quality audit entities are more geared toward advancing more diverse pedagogies in the classroom. These forces for change reinforced as well as sped up the institution's transition toward implementing a teaching and learning strategy that has a greater emphasis on active learning orientations. With this came the immediate need to create a new SET instrument.

At this point, the institution was using an existing SET instrument that had several issues. First, the questions in the instrument lacked balance as much of the focus was on the teacher's actions at the expense of aspects directly related to the learning experience and support of learning. Second, the questions had a strong leaning toward the lecture mode as the main mode of delivery. Third, it provided limited opportunities to better understand the student's reflection of their own learning experience.

The newly designed SET instrument hence needed a balance between questions on teaching, the learning experience, and learning support to provide more meaningful data on the teaching and student learning experience. In practical terms, the new instrument should not focus on lecture-based pedagogies at the expense of active learning pedagogies. Numerous academic staff had felt somewhat limited by the framing of the existing SET instrument. Much of this thinking was consistent with recommendations made by Tagomori and Bishop (1994) and Rowley (2003). In addition, there was also an implicit and explicit recognition by the members of this community of practice that not all the teaching staff would be immediately ready to be rated based on active learning practices as many (including students) were still used to the lecture mode. In other words, the instrument should positively challenge the learning culture at the institution, functioning as an impetus as well as a bridging tool to advance toward more active and problem-based learning orientations.

The Design Process

1. Setting up a cross-disciplinary task-oriented voluntary team or community of practice. As the purpose became clearer to the college teaching staff, the momentum began to build for voluntary participation in a more task-oriented community of practice. The initial community that focused on discussion had a fluid membership of about 15 academic staff. When the needs of the community became more task oriented (i.e., let's begin designing the instrument), an appeal was made to ensure a steady commitment by participating members, and ideally for the community to be represented by each of the five faculties at the college. On both accounts, the appeals were fulfilled as the number of the community reduced and stabilized to about eight people. Among this group was also a member from the institution's senior management team who agreed to communicate the findings and progress to the college principal. Another member was the campus teaching and learning director. The diverse composition of this community was critical to facilitate the change process that would come with revamping a high-stakes SET instrument.

2. Reviewing SET research literature, samples and question banks. As members of the community were mostly not experts on SET, a survey of the field and in-depth discussions were essential. A summary of the research literature is presented in the literature review earlier in this article. Numerous samples and questions banks were also acquired for review and critiqued. One question bank that was particularly helpful was downloaded from Syracuse University's teaching and learning support center (Center for Support of Teaching and Learning, 2007). It is worth noting that the site also contains numerous other helpful SET-related resources. It was while reviewing this body of knowledge that it became quite clear that the community needed to describe or define the characteristics of effective teaching and learning.

3. Describing characteristics of effective teaching and learning in order to "ask the right questions." Tagomori and Bishop (1994) suggested this as a key step as it provides the foundation and framework on which the SET instrument can be built. A number of studies were reviewed—all of them extremely informative (Bransford, Brown, & Cocking, 2000; Bain, 2004; Kuh, Kinzie, Schuh, Whitt, & Associates, 2005; Pascarella & Terenzini, 2005)—but Chickering and Gamson's (1987) *Seven principles for good practice in undergraduate education* seemed to fit the institutional needs quite nicely. These principles encouraged student and instructor contact, cooperation and collaboration among students, and greater involvement in active learning approaches. Educators were also asked to give prompt and meaningful feedback, emphasize time on task, communicate high expectations and respect diverse talents and ways of learning of their students. For the purpose of guiding the design of the SET instrument, the seven principles were concise and straightforward, and in many ways provided a simple unifying platform to achieve for both teachers and students. Most important, the principles were in alignment with the institution's emerging teaching and learning strategy.

4. Developing SET questions. Questions were mapped or matched with the "seven principles" and were logically grouped according to the following dimensions: student reflection, course evaluation, and teaching and learning evaluation. This part—including using the different dimensions to strike some balance between questions on teaching, support, resources, and student learning—is also consistent with recommendations made by Tagomori and Bishop (1994) and Rowley (2003). The student self-reflection dimension was to encourage students to think about their own responsibility before giving feedback on teaching. It would also enable correlation analyses to be performed against the teaching and learning ratings (e.g., what kind of feedback did students in the course who behaved conscientiously give in relation to teaching or learning ratings?). As for the course evaluation dimension, the primary aim was to provide an indicator of the overall health of the course in terms of its syllabus and the learning experience. With respect to teaching and learning evaluation items, questions were formed with three considerations in mind: to align each item with Chickering and Gamson's good practices, to encourage active learning and not to discourage other approaches of teaching and learning aside from the lecture method. At a later point, hygiene factors were added. Please refer to *Table 1* for a summary of the sample design that is indicative of the priorities set out for this particular institution.

5. Inviting academic staff to review and critique the working draft. All academic staff were invited to give feedback, either via focus groups or email. In total, some 40 percent of the staff provided input into the development of the final draft – with a large portion doing so during open focus group discussions facilitated by members of the community of practice. Four focus groups were scheduled, and with attendance averaging in the teens, many discussions were quite vibrant. Written feedback was also requested from the participants to provide a completely anonymous platform for candid and constructive critiques. Based on the cumulative feedback, adjustments were made before the final SET form was presented to the senior management group. It is important to note that this exercise along with the involvement of a grassroots-driven community of practice helped create a greater sense of ownership and empowerment – both important in any change management process.

Table 1 Distribution of SET items based on Chickering and Gamson's (1987) Seven principles for good practice in undergraduate education

Dimension	Sample questions	Number of question related to this dimension*
Student reflection on responsibility**	I participated and asked questions during class activities. I came to class having completed reading assignments, homework, or other activities assigned to me.	6
Course evaluation**	In this class, I had to remember and recall information such as theories, facts, formulas, etc. [Memorizing] In this class, I learnt to break down different ideas, theories, methods, experiences, cases, etc. to smaller parts for further evaluation. [Analyzing] In this class, I learnt to make choices after evaluating different ideas, theories, methods, experiences, cases, etc. [Making judgments]	5
Teaching and learning evaluation:		
Principle 1: Engagement	The lecturer was willing to assist us in this subject. The lecturer was available for consultation at scheduled times.	5
Principle 2: Cooperation	The lecturer organized class/group discussions which contributed to my learning.	2

Dimension	Sample questions	Number of question related to this dimension*
Principle 3: Active learning	The lecturer encouraged us to ask questions during class interactions. The lecturer helped us to learn to identify problems and explore different explanations/solutions. The lecturer encouraged us to relate the theories to practical situations.	7
Principle 4: Prompt feedback	The lecturer advised us on how to allocate time in completing assignments/projects, class activities, etc.	2
Principle 5: Time management	The lecturer advised us on how to allocate time in completing assignments/projects, class activities, etc.	2
Principle 6: High expectations	The lecturer discussed with us the challenging aspects of the coursework/topics of this subject The lecturer encouraged us to constantly improve the quality of our work.	3
Principle 7: Diverse talents and ways of learning	The lecturer used different activities to encourage learning.	2
Hygiene factors	The lecturer's explanations were clear. The lecturer was available for consultation at scheduled times.	5

* Does not add up to the final total of 28 items as some questions overlap

** These dimensions were also influenced by the works of Kuh, Kinzie, Schuh, Whitt, & Associates's *Student success in college: Creating conditions that matter* (2005)

Lessons Learned and Limitations

The lessons learned and realizations of limitations were directly influenced by the intense participation of the community of practice members along with focus group sessions involving many of the institution's academic staff. Involving individuals with different professional responsibilities and expertise in this process created a productive environment to recognize realistic conditions for the effective design of a new SET instrument. In many ways, the following is a summary of their input.

SET instruments, no matter how well designed, have inherent limitations. It is dependent on the quality of the students' response (e.g., Were they honest? Did they understand the question? How did they complete the 28 items in less than 2 minutes?). In addition, as discussed in the literature review, extraneous influences such as gender- bias and syllabus requirements can significantly affect the way students provide feedback. In this regard, institutions of learning should never rely on students ratings alone to evaluate teaching (Murray, 2005). For a more comprehensive discussion about this, please refer to Cashin (1995, 1996).

SET instruments are only tools. SET instruments are only as effective and useful as educators, administrators and students allow them to be. For example, research has shown that teaching performance can be improved when student feedback is accompanied by expert consultation (Cashin, 1995; Murray, 2005). Yet, many institutions do not provide such support, using SET data only for administrative purposes. In another example, if a dean is concerned that the SET form is not relevant enough to the faculty's discipline, it should be suggested that the respective faculty consider developing an instrument to complement the campus-wide SET form. In fact, Purvanova (2002) would recommend such an action as different disciplines may have different evaluation needs.

There must be a commitment to allocating adequate resources and support to develop a SET instrument. Perhaps this is stating the obvious, but there must be explicit recognition of this. For example,

there may be a real temptation to create a SET instrument by borrowing, adopting or adapting from other institutions. But if the SET instrument is critical to the college's administrative system, the development process must be contextualized – taking into consideration key drivers such as the college mission and the professional development opportunities that should come following an evaluation. In other words, creating a SET instrument will necessitate significant time and resources (Tagomori & Bishop, 1994).

This project did not seek students' feedback. Should it have, considering the increasing focus on greater student engagement, student learning and learner-centered pedagogies? In retrospect, involving students would have added great value to the development process.

CONCLUSION

Research indicates that effective teaching is a complex, multidimensional process and this should be reflected in the design of SET instruments. As such, institutions and individuals tasked with designing a SET instrument should expect it to be a deliberate and long process that needs to be carefully formulated. The design process involves five key things. Firstly, an empowered and dedicated cross-disciplinary team. Secondly, a clear purpose and understanding of where an institution is and where it wants to be, particularly with regard to its teaching and learning culture. The SET initiative must be aligned with this mission. Thirdly, an in-depth review of the SET and teaching and learning literature. The first three items set up the fourth aspect recommended by Tagomori and Bishop (1994) and Rowley (2003): describe and define the characteristics of effective teaching; phrase and map the SET questions applying to those characteristics; and try to strike a balance between questions on teaching, support, resources and learning. And finally, involve the necessary stakeholders including teachers, administrators and students. Recognize that it is a change process (Fullan, 2001), and that intense feelings will likely surface particularly in institutions that place significant importance on its SET data. Getting necessary stakeholders involved early will create a greater sense of ownership and empowerment and should bode well for the next phases: the validation and the implementation process.

Note: An earlier version of this paper was presented at a conference and several discussion forums.

REFERENCES

Bain, K. (2004). *What the best college teachers do.* Boston, MA: Harvard University Press.

Bransford, J., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=9853&page=R3

Cashin, W. (1995). *Student ratings of teaching: The research revisited.* Manhattan, Kansas: IDEA Paper no.32, Kansas State University Center for Faculty Evaluation and Development .

Cashin, W. (1996). *Developing an Effective Faculty Evaluation System.* Manhattan, Kansas: IDEA Paper no.33, Kansas State University Center for Faculty Evaluation & Development.

Center for Support of Teaching and Learning. (2007). *Student ratings of teaching effectiveness: Using the CSTL item bank to create your own form.* Retrieved from Syracuse University: Center for Support of Teaching and Learning: http://cstl.syr.edu/cstl2/home/Teaching%20Support/Teaching%20at%20SU/Student%20Ratings/12A310.

Chickering, A., & Gamson, Zelda F. (1987). Seven principles for good practice in undergraduate education. *American Association for Higher Education, 39*(7)*,* 3-7.

Fullan, M. (2001). *Leading in a culture of change.* San Francisco, CA: Jossey- Bass.

Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environment. *Educational Technology Research and Development, 48*(3)*,* 23-48.

Koh, H. C., & Tan T. M. (1997). Empirical investigation of the factors affecting SET results. *International Journal of Educational Management, 11*(4)*,* 170-178.

Kuh, G., Kinzie, J., Schuh, J., Whitt, E., & Associates. (2005). *Student success in college: Creating conditions that matter.* San Francisco, CA: Jossey-Bass.

Lyman, P., & Varian, H. (2004). *How much information.* Retrieved from http://www.sims.berkeley.edu/how-much-info-2003

Marton, F., & Saljo, R. (1976). On qualitative differences in Learning I: Outcome and process. *British Journal of Educational Psychology, 46,* 4-11.

Ministry of Higher Education Malaysia. (2007a). *National Higher Education Action Plan 2007-2010.* Retrieved from http://planipolis.iiep.unesco.org/upload/Malaysia/Malaysia%20Higher%20education%20action%202007-2010.pdf

Ministry of Higher Education Malaysia. (2007b). *Pelan strategik pengajian tinggi Negara: Melangkaui tahun 2020.* [*National higher education strategic plan: Laying the foundation beyond 2020*] Putrajaya: Ministry of Higher Education Malaysia. Retrieved from http://engine.um.edu.my/doc/FKUM/PelanStrategikPengajianTinggi2020.pdf

McKeachie, W., & Kaplan, M. (1996). Persistent problems in evaluating college teaching. *AAHE Bulletin, 48*(6). Retrieved from http://cedar.olemiss.edu/depts/vc_academic_affairs/problems.html

McNiff, J., & Whitehead, J. (2002). *Action research: Principles and practice*. London, UK: Routledge.

Murray, H. (2005). *Student evaluation of teaching: Has it made a difference?* Paper presented at Annual Meeting of the Society for Teaching and Learning in Higher Education. Charlottetown, Prince Edward Island.

Ngware, M. W., & Ndirangu, M. (2005). An improvement in instructional quality: Can evaluation of teaching effectiveness make a difference? *Quality Assurance in Education, 13*(2)*,* 183-201.

Pascarella, E., & Terenzini, P. (2005). *How college affects students (Vol. 2): A third decade of research.* San Francisco, CA: Jossey-Bass.

Pounder, J. S. (2007). Is student evaluation of teaching worthwhile: An analytical framework for answering the question. *Quality Assurance in Education, 15,* 178-191.

Purvanova, R. K. (2002)*. Empirical examination of factors that influence student ratings of instructors*. Paper presented at the Annual Conference of the American Psychological Association, Chicago.

Resnick, L. (1987). Learning in school and out. *Educational Researcher, 16*(9)*,* 13-20.

Rowley, J. (2003). Designing student feedback questionnaires. *Quality Assurance in Education, 11*(3)*,* 142-149.

Sanders, W., & Rivers, June C. (1996). *Cumulative and residual effects of teachers on future student academic achievement.* Knoxville, TN: University of Tennessee. Retrieved from http://www.cgp.upenn.edu/pdf/Sanders_Rivers-TVASS_teacher%20effects.pdf

Tagomori, H., & Bishop, L. (1994). *Content analysis of evaluation instruments used for student evaluation of classroom teaching performance in higher education.* Paper presented at the Annual Meeting of the American Educational Research Association. New Orleans, LA.

Whitehead, A. (1929). *The aims of education and other essays.* New York, NY: Macmillan.

Wurman, R. (1989). *Information anxiety* (1st ed). New York, NY: Doubleday.