

RESPONSE RATE AND RATINGS FOR STUDENT EVALUATION OF TEACHING: DOES ONLINE ADMINISTRATION MATTER?

Debbie Chor Shan Liu
Educational Development Centre
The Hong Kong Polytechnic University
HKSAR, CHINA
debbie.liu@polyu.edu.hk

Christine Armatas
Educational Development Centre
The Hong Kong Polytechnic University
HKSAR, CHINA
christine.armatas@polyu.edu.hk

ABSTRACT

A comparative study was conducted at the time of switching from a traditional paper-based system for student evaluation of teaching to an online one to examine if there were any significant differences in response rates and ratings between the two modes of administration, and whether lower response rates were associated with lower ratings. 'Early' versus 'late' respondents' ratings were also looked at to see if students with negative feedback were likely to be the first to respond. Results of the study suggest that, consistent with much prior research, the mean response rate for online administration was significantly lower by about 21%. However, it had little effect on the mean ratings. In addition, no significant differences were found in ratings obtained via paper-based and online administrations, nor that given by 'early' and 'late' respondents. The findings provide a starting point for dispelling some of the worries about reliability issues associated with online administration, and sheds light on ways that may better the system at the university under study.

Keywords: Online Student Evaluation; Response Rates; Student Ratings, Student Satisfaction with Teaching.

INTRODUCTION

A study comparing response rates for, and student ratings on, surveys evaluating teaching is reported in this paper. Data for comparison was collected as part of a pilot implementation of a change from in-class and paper-based to online and out-of-class administration of an existing teaching evaluation instrument at a Hong Kong university. Consistent with experiences elsewhere, response rates to the survey dropped dramatically when online administration was introduced. The drop in response rates prompted voluble and persistent complaints from teaching staff about the threat to the integrity of the survey results, prompting investigation of the effect of administration mode on both response rates and student ratings for the same survey.

Our experience is not new. The switch to online administration of student evaluations in lieu of the traditional paper-based system has added a new dimension to the ongoing debate on the reliability and validity of student evaluation of teaching in higher education. Many universities worldwide have implemented an online system to benefit from significant time and cost savings in administration and processing of feedback, as well as enhanced data integrity. Other advantages of online administration include not using class time to collect feedback and eliciting more feedback in terms of quantity and quality, as students can have more time to reflect on their responses when completing the evaluations in their own time at their own pace and place (Hmieleski & Champagne, 2000; Hardy, 2003; Johnson, 2003; Ardan, Ardan, Coppage, & Crouch, 2007; Dommeyer, Baum, Hanna, & Chapman, 2004; Sorenson & Reiner, 2003; Stowell,

Addison, & Smith, 2012). Despite these benefits, many, including staff at our university, express grave concerns over the reliability of online student evaluations because of the generally lower response rates consistently documented in the research literature (e.g. Capa-Aydin, 2016; Goodman, Anson, & Belcheir, 2014; Nulty, 2008). There is also a perception that online evaluations tend to induce more negative feedback as students are thought to be more susceptible to negative influences outside the classroom (Risquez, Vaughan, & Murphy, 2015; Stowell et al., 2012; Hardy, 2003). Moreover, as the online medium allows every student an equal opportunity to provide feedback (an advantage over the traditional in-class paper-based evaluations often cited in the literature), many teachers question whether students who do not attend classes are able to make accurate judgements on teaching quality (Risquez et al., 2015). Concerns then arise about potential sampling bias and whether the mode of administration will alter students' behaviours and attitudes in responding to the same surveys administered in a different way, and whether any change found in the results can be attributed to the student experience or the mode of survey administration.

With evaluation results being increasingly tied to the making of important personnel decisions (Collings & Ballantyne, 2004; Ory, 2000; Avery, Bryant, Mathios, Kang, & Bell, 2006), online evaluations are seen by some to have further undermined the reliability and validity of student evaluation of teaching. This perceived threat has resulted in a number of comparison studies on the traditional paper-based evaluation versus the online evaluation, investigating differences between the two administration modes in relation to response rates, student responses, psychometric properties, class characteristics (e.g. discipline, class size, etc.) and student demographics (e.g. gender, grade-point average). Traditionally, student evaluations have been administered in class to a captive audience at the end of the semester using paper forms, with response rates averaging out to 70–80% (Goodman et al., 2014). Without a captive audience, online student evaluations tend to yield lower response rates than in-class paper-based evaluations by 20–30% (Goodman et al., 2014; Nulty, 2008) with the average response rates generally falling between 30% and 60% (e.g. 33% in Nulty, 2008; 41% in Anderson, Brown, & Spaeth, 2006; 44% in NSSE, 2003; 50% in Johnson, 2003, and in Kulik, 2009; 55% in IDEA Centre, 2011) – although response rates vary significantly across institutions and classes (Dommeyer et al., 2004; Anderson et al., 2006).

While comparisons of the response rates between the paper and online administration modes have offered relatively consistent findings, those of quantitative ratings have shown mixed results. For example, in Capa-Aydin's (2016) review of 15 studies comparing the response rates and mean ratings of in-class and online evaluations, some reported higher mean ratings in favour of in-class evaluation at item or overall level; some showed slight differences in favour of online evaluation; and some found no differences. On the whole, only two of the 15 studies yielded significant results in favour of the traditional in-class paper-based evaluation in regard to overall ratings. In terms of qualitative responses, a number of the studies have reported lengthier, more frequent and more informative feedback being provided in online evaluations (e.g. Hmieleski & Champagne, 2002; Dommeyer et al., 2004; Morrison, 2011; Johnson, 2003; Layne, DeCristoforo, & McGinty, 1999; Donovan, Mader, & Shinsky, 2006). For example, Collings and Ballantyne (2004) examined three datasets, each from a different survey and student type, one of which was a student survey of teaching. They found that for all three datasets, online surveys clearly yielded comments 25–50% lengthier, although a marginally greater proportion of respondents provided comments in paper-

based surveys. Hmieleski and Champagne (2000) reported that students in an online evaluation wrote an average of four times as many comments as those in a paper-based survey. In their study, Layne et al. (1999) found that 26% more respondents provided written comments online. Similarly, Donovan et al. (2006) found that 27% more online respondents provided comments than paper-based survey respondents did, and the online comments were 54% lengthier than those submitted via paper-based survey. Furthermore, the amount of formative feedback made by online respondents was found to be 6% more. Ardalan et al. (2007), in comparing the quantity and quality of written responses, also found that online comments were lengthier and more meaningful than those from paper-based administration. On the other hand, in terms of the favourableness of written comments, Morrison (2011) found no evidence that online respondents tend to provide more positive or negative feedback than paper-based survey respondents.

Overall, the research on the impact of switching from paper-based to online administration of surveys evaluating teaching shows that response rates are substantially lower for online, but that the ratings do not really change. However, students provide more detailed feedback for qualitative responses for online surveys of this nature. Taken together, the research findings suggest that, provided adequate response rates are maintained, there are benefits of implementing online surveys, including efficiencies in survey administration, data management and reporting, as well as richer written feedback. Despite this, at the university where the current study was conducted, the proposal to move to online administration of the existing survey of teaching was met with considerable resistance. In addition to concerns about response rates dropping and impacting on the representativeness of results, staff had other concerns. These included the appropriateness of allowing all students to complete the SFQ online regardless of their history of class attendance and that online administration would encourage students with negative views to express them more readily. Given the concerns of staff, analysis of data from the pilot implementation of an online version of the existing survey at the university was undertaken with a specific focus on differences in response rates and student ratings between different administration types.

BACKGROUND TO THE PRESENT STUDY

Student Evaluation of Teaching and Learning at the University

At the university where this study was conducted, student feedback on teaching and learning has been collected regularly for developmental and judgemental purposes for the past 20 years via an evaluation instrument known as the Student Feedback Questionnaire (SFQ). In 2011/12, the university decided to explore the feasibility of replacing the traditional in-class, paper-based evaluation system with a custom-designed online system built in-house, known as the eSFQ system. Pilot implementation of the system was carried out in the subsequent two academic years, gradually involving all subjects offered at the university. The SFQ exercise normally takes place in the last two weeks of each semester. Consistent with practices elsewhere around the globe, the paper-based SFQ was administered in class by a departmental administrative staff or a student representative, in the absence of the teaching staff member concerned. Students were usually given 10 to 15 minutes to complete the SFQ forms and the completed forms were then collected and sent in a sealed envelope to the central unit responsible for scanning the forms and data processing. With the paper-based SFQ, staff members could choose a time they deemed appropriate to conduct the survey, for example, in the last class (by which time students' learning experience is thought to be most complete) or in a class where high attendance is expected (e.g. a class reviewing for the

final exam). With the new eSFQ system, students are invited to complete the SFQ online via email and short message system (SMS) notifications, with a link to the system login page embedded in the message. In order to access and complete the eSFQ forms, students are required to log in to the system using their student ID and password. During the survey period, non-respondents are reminded of any incomplete eSFQ periodically until all available forms for that student are completed or the survey period ends. Staff members can check the real time response rates via the eSFQ system and they are encouraged to ask their students to respond to the survey via in-class explanation or additional emails should they see fit. After the overall examination results are finalised, the SFQ results are made available to staff members in the form of a report, containing the means, standard deviations and percentage distributions for each of the ratings of the SFQ items. For the online administration, the written comments are de-identified, compiled into a list, and appended to the report as well; otherwise, the paper forms are returned to the staff concerned for their perusal of the written feedback. In addition, the mean ratings on the subject are made available to students as well.

Instrument

The SFQ consists of two sections: one about the subject and one about the teaching of the staff member. Items in the SFQ forms differ slightly across schools/faculties and subject types but the forms, nonetheless, contain a set of standard items used across all schools and faculties to elicit students' feedback on the subject and the teaching of the staff member concerned. Both the paper-based system and eSFQ system use the same set of standard items, except that the eSFQ allows subject leaders and subject teachers to add a maximum of five additional close-ended and/or open-ended questions of their choice in the corresponding section. There are six close-ended items and four open-ended items in the standard item set. The close-ended items are rated on a five-point Likert scale, ranging from 'strongly disagree (1)' to 'strongly agree (5)'. For the current study, ratings for the six close-ended items in the standard item set were used to compare quantitative responses to the same survey with online or paper-based administration.

Table 1. Standard items (close-ended items)

Section I About the subject

Item 1 I have a clear understanding of what I am expected to learn from this subject.

Item 2 The teaching and learning activities (e.g. lectures, discussions, case studies, projects, etc.) have helped me to achieve the subject learning outcomes.

Item 3 The assessments require me to demonstrate my knowledge, skills and understanding of the subject.

Item 4 I understand the criteria according to which I will be graded.

Section II About the staff member

Item 5 The teaching of the staff member has provided me with a valuable learning experience.

Item 6 Overall, I think that the staff member is an effective teacher.

Research Questions

Utilising the data collected for the pilot implementation of the online administration of the existing SFQ at a university in Hong Kong, the present study compared the response rates and mean ratings for surveys completed online with those from the traditional in-class, paper-based system to examine the following questions:

- (1) How different are the mean and distribution of eSFQ response rates from those of paper-based SFQ?
- (2) Are there any significant differences in the SFQ ratings given by students who respond early to the eSFQ as compared to those of the late respondents?
- (3) Do classes with lower response rates tend to receive lower SFQ scores?
- (4) Are there any significant differences in the reliability between eSFQ and paper-based SFQ in terms of their internal consistency?
- (5) Are there any significant differences in the means and distributions of the SFQ scores (i.e., student ratings) collected via eSFQ as compared to those collected via paper-based SFQ?

METHODOLOGY

Sample

This study was conducted across two academic years. In the 2012/13 academic year, a total of 443 eSFQ surveys were conducted for 105 subjects. A total of 13994 questionnaires were sent out to 6914 students. In Semester 1, 2013/14, a total of 660 eSFQ surveys were conducted on 127 subjects, involving 21190 questionnaires and 9639 students.

Data Sources

The data used in the study primarily came from the pilot conducted in four consecutive semesters (i.e., Semesters 1 and 2 and Summer Term in 2012/13 and Semester 1 in 2013/14). Data obtained from paper-based administration in earlier years were also used for comparison. Since it is university policy that all SFQ for General University Requirements (GUR) subjects be conducted online, all data from the GUR subjects were collected via the eSFQ for the pilot. For comparison purposes, the most recent data obtained from the paper-based SFQ for the same subjects taught by the same teachers in the previous year were also used. In addition to collecting SFQ data online for GUR subjects, an invitation to take part in the pilot of the eSFQ system with non-GUR subjects was sent to all academic staff. Interested participants were followed up with an email explaining the purpose of the pilot and subject selection criteria, and highlighting some key dates for the pilot. A short meeting was also scheduled for each participating staff member to brief them again on the purpose of the study, work out the logistics associated with eSFQ for their subject, and to answer any questions they might have. Confirmation of the agreed arrangements were then sent to all participating staff members via email.

For the non-GUR subjects, where possible, data were collected via the traditional in-class paper-based system and the eSFQ system under a split-half setting, in which classes of the same subject were randomly assigned to complete the SFQ using either mode of administration. For those assigned to use the online administration, the eSFQ was conducted mainly out of class except for three classes where the split-half setting was conducted within the same class, i.e., half of the class was given the paper forms to complete, and the other half was asked to do it online (also in class).

The response rates to the eSFQ of these three classes ranged from 66.67% to 100%. For subjects where staff members agreed that only eSFQ be conducted, the most recent paper-based SFQ data for the same subjects taught by the same teachers were used for comparison purposes (inter-semester comparisons). To ensure uniformity of the in-class administration practices, all in-class paper-based SFQ and in-class eSFQ for the participating subjects were conducted by the same researcher. The purpose of the study was explained and instructions were given to the students using a set of standardised presentation slides.

RESULTS

Differences In Mean Response Rates Between The Paper-based SFQ And eSFQ

The response rate means and distributions of scores for eSFQ versus paper-based SFQ are shown in Table 2. Results show that the mean response rate for the paper-based SFQ was significantly higher than that of the eSFQ, with a mean difference of 20.62% ($t=34.503$, $p<.001$). Significant differences were also found in the distribution of the response rates between the two modes of administration ($\chi^2=1061.724$, $p<.001$). Over 72% of the paper-based SFQ surveys received a response rate of 70% or above, which was achieved by only about 27% of the eSFQ surveys. On the other hand, about 6% of the eSFQ surveys had a response rate of below 30%, as compared to 1.62% for the paper-based SFQ ones. These analyses show that, on average response rates are higher for paper-based administration and that fewer classes had response rates of 70% or higher with online administration.

Table 2. Differences in the means and distributions of the response rates of eSFQ versus paper-based SFQ (All subjects, 2012/13 and all participating subjects in the eSFQ pilot, Semester 1, 2013/14)

	Paper	Online	Sig.
No. of SFQ surveys	6992	1103	
Mean response rate	78.97%	58.35%	$t=34.503$, $p<.001$
Standard Deviation	18.44	18.44	
Response Rate	Below 30%	1.62%	$\chi^2=1061.724$, $p<.001$
	30-39.99%	2.17%	
	40-49.99%	4.41%	
	50-59.99%	7.77%	
	60-69.99%	11.38%	
	70-79.99%	15.39%	
	80-89.99%	21.85%	
	90-100%	35.41%	4.44%

Early Versus Late Respondents

An independent t-test was conducted to test the commonly expressed view that students with strong negative views about their learning experience will have a much higher inclination than

others to respond to eSFQ. Differences in the mean SFQ scores were compared for the six standard items between the first 25% of respondents versus the last 25% respondents taking part in the eSFQ pilot for non-GUR subjects in Semester 1 of 2013/14 (see Table 3). Results showed that the mean SFQ scores given by the first 25% of respondents were consistently higher than those given by the last 25% respondents, although none of the differences were statistically significant at the .05 level except for Item 6 of Section II ($t=2.991$, $.01 < p < .05$). For all of the items tested, a larger proportion of the early respondents gave a high rating (scores 4 and 5) than the late respondents did, but the difference is statistically significant only for Section II Item 6 ($\chi^2=14.132$, $.01 < p < .05$). This finding is not consistent with the view that students with strong negative views tend to complete the eSFQ.

Table 3. Comparison of the means and distributions of eSFQ ratings of the first and last 25% of respondents of selected non-GUR subjects, Semester 1, 2013/14

		Ratings					Valid N	Mean	SD	Sig.
		1	2	3	4	5				
Item 1	First 25%	0.9%	2.3%	13.9%	57.9%	25.0%	216	4.04	.75	$t=1.629$, $p > .05$; $\chi^2=3.595$, $p > .05$
	Last 25%	1.4%	3.6%	19.5%	53.4%	22.2%	212	3.91	.82	
Item 2	First 25%	0.9%	3.2%	18.1%	46.8%	31.0%	216	4.03	.84	$t=1.841$, $p > .05$; $\chi^2=9.089$, $p > .05$
	Last 25%	0.9%	3.2%	20.1%	57.1%	18.7%	219	3.89	.77	
Item 3	First 25%	0.5%	0.9%	11.1%	56.5%	31.0%	216	4.17	.69	$t=1.929$, $p > .05$; $\chi^2=4.242$, $p > .05$
	Last 25%	0.9%	2.7%	14.5%	56.1%	25.8%	212	4.03	.77	
Item 4	First 25%	0.5%	3.8%	15.5%	54.5%	25.8%	213	4.01	.78	$t=1.125$, $p > .05$; $\chi^2=2.007$, $p > .05$
	Last 25%	1.4%	3.6%	18.6%	53.6%	22.7%	202	3.93	.82	
Item 5	First 25%	0.8%	3.6%	15.1%	47.2%	33.3%	252	4.08	.83	$t=1.782$, $p > .05$; $\chi^2=5.292$, $p > .05$
	Last 25%	0.8%	3.2%	21.2%	49.2%	25.6%	250	3.96	.82	
Item 6	First 25%	0.8%	3.2%	11.9%	45.5%	38.7%	253	4.18	.82	$t=2.991$, $.01 < p < .05$; $\chi^2=14.132$, $.01 < p < .05$
	Last 25%	1.2%	2.4%	20.2%	51.2%	25.0%	248	3.96	.81	

Relationship Between Response Rates and Mean SFQ Scores

A very weak, positive (albeit statistically significant) correlation was observed between response rates and mean SFQ scores for online administration for five of the six items, with the value of the correlation coefficients (r) ranging from .088 (Section I Item 2) to .122 (Section II Item 7). Based on the R-squared values for these correlations, the observed variation in response rates can, at best,

account for less than 1.5% of the variation in mean SFQ scores, suggesting that a lower response rate has little meaningful effect on mean SFQ scores. The correlations between the response rates to eSFQ and the mean SFQ scores for the six standard items in Sections I and II are shown in Table 4.

Table 4. Correlations between response rates and mean eSFQ ratings (All eSFQ surveys conducted in 2012/13 and 2013/14)

Variables	Adjusted Response Rate (N=1103)	Item 1 Mean (N=1051)	Item 2 Mean (N=1051)	Item 3 Mean (N=1050)	Item 4 Mean (N=1051)	Item 5 Mean (N=1051)	Item 6 Mean (N=1049)
Adjusted Response Rate	1.00	.054	.088*	.106*	.108**	.096*	.122**
Mean	58.27%	4.00	4.00	4.03	3.93	4.07	4.15
SD	18.39	.35	.36	.35	.37	.42	.41

* $p < .01$. ** $p \leq .001$.

Internal Consistency Of The eSFQ

Table 5 shows the results of a series of independent t-tests on the differences in the mean values of the Cronbach's alpha (a measure of internal consistency) for Section I Q1-4 and Section II Q6-7 between the paper-based and eSFQ. No statistically significant differences were found in the reliability of the eSFQ and paper-based SFQ with regards to their internal consistency for Section I Q1-4. On the other hand, the mean alpha value for Section II Q6-7 tended to be higher for eSFQ than for the paper-based SFQ (0.92 versus 0.87), and the difference was statistically significant at .05 level ($t = -2.725$, $.05 < p < .01$). The results suggest that the reliability of the eSFQ in terms of internal consistency is comparable to that of the paper-based SFQ.

Table 5. Comparison of the internal consistencies of Section I Items 1-4 and Section II Items 6-7 between paper-based and eSFQ

Type of Administration	Section I Items 1-4		Section II Items 6-7	
	Paper-based	Online	Paper-based	Online
No. of SFQ Surveys	34	32	45	42
Mean Alpha Value	0.84	0.85	0.87	0.92
Std. Deviation	0.11	0.12	0.10	0.07
Std. Error Mean	0.018	0.021	0.015	0.011
Sig.	$t = -.394$, $p > .05$		$t = -2.725$, $.05 < p < .01$	

Comparability of Mean SFQ Scores Between Paper-based SFQ and eSFQ

Differences in the mean SFQ scores of the same subject taught by the same teacher collected via eSFQ and paper-based SFQs were examined using a 2 x 2 analysis of variance with the factors of type of SFQ administration (paper or online) and time of administration (same semester or adjacent

semesters). The dependent variable used was the mean of the two items about the teacher which is the summary score used by the university for evaluation of teaching quality. No significant main or interaction effects were found, indicating that there were no statistically significant differences between the eSFQ and paper-based SFQ in the mean SFQ scores for these two items for split-half administrations (i.e., intra-semester comparisons). Nor were there statistically significant differences between the mean of these two items for the same subject taught by the same teacher between adjacent semesters (i.e., inter-semester comparisons). Means and standard deviations for conditions compared in this analysis are detailed in Table 6.

Table 6. Descriptive statistics for comparison of the mean SFQ score for paper and e-administrations

Type of comparison	Type of SFQ administration	Mean	Standard Deviation	N
Inter-semester	Paper	4.05	.49	24
	Online	4.12	.46	24
	Total	4.08	.50	48
Intra-semester	Paper	4.06	.27	28
	Online	3.97	.40	28
	Total	4.01	.35	56
Total	Paper	4.05	.38	52
	Online	4.04	.43	52
	Total	4.05	.40	104

DISCUSSION

Four key questions were addressed through the analysis of data collected for the pilot implementation of the online administration of the existing SFQ at a Hong Kong university. These were whether response rates and scores differ depending on whether the SFQ is administered online or paper-based and whether lower response rates are associated with lower scores. In addition, analyses were conducted to compare ratings of students at the start and end of the survey period.

Consistent with prior research, the response rates to the eSFQ, which averaged out to 58.35%, were significantly lower than those of paper-based SFQ administered in class, with a mean difference of 20.62%. However, we found that lower response rates had little effect on the mean SFQ ratings, as suggested by the very weak positive correlation (albeit statistically significant at .05 level) found between the two variables. In addition, students who complete SFQs for their subjects 'early' do not give lower rating compared to students who respond towards the end of the survey period. We also found no evidence that the mean ratings for eSFQ are lower than that of paper-based SFQ. In fact, no statistically significant differences were found in the mean SFQ scores for the two items on the staff member. In sum, the study found no systematic bias in the mean SFQ ratings for online and out-of-class administration, despite the significantly lower response rates. The empirical evidence from our study is inconsistent with some of the commonly cited concerns of staff about the effect on rating scores and potential sampling bias in online administration, and lends support to the implementation of the eSFQ system as a viable alternative

to the paper-based SFQ system at the university in question. However, the widespread concern about the lowered reliability of survey results associated with the lower response rates remains a salient issue that needs to be addressed. The research literature has provided an array of strategies to improve response rates (e.g. Nair, Adams, & Mertova, 2008; Bennett & Nair, 2010; Goodman et al., 2014; Crews & Curtis, 2011). Some commonly used methods with more immediate results include sending repeated reminders to non-respondents, offering incentives to students, and encouraging staff members to urge their students to respond to the surveys (Nulty, 2008). Other approaches (with less immediate results) include avoiding over surveying students – as too many surveys lead to survey fatigue and disinterest (Leckey & Neill, 2001; Adams & Umbach, 2012); and ‘closing the loop’ – as students not seeing or believing that their feedback is being acted upon diminishes their willingness to participate in subsequent surveys (Powney & Hall, 1998).

As recommended (Nulty, 2008; Goodman et al., 2014), the university where this study was conducted has adopted a wide range of approaches recommended in the research literature to raise awareness and promote student participation in the SFQ exercise. However, despite the measures taken to promote student participation, the mean response rates declined over the course of the study. Response rates therefore need to be monitored to ensure they do not decrease to a level where the results are meaningless.

One possible approach to improve the response rates is to allow staff members to have an option to administer the eSFQ either in-class or out-of-class. As suggested in the research literature, the generally higher response rates obtained from the paper-based evaluation are probably due to administration to a captive audience (Nulty, 2008), so conducting the online surveys in class may yield response rates similar to those of the paper-based surveys. Although administering the eSFQ in class negates the benefits of not taking up valuable class time with surveys and allowing students more time to respond to the surveys at their own pace and place, the higher response rates it tends to yield may address the concern that substantial effort is required to attain response rates similar to those of paper-based SFQ. It may also alleviate the sense of loss of autonomy in administration in terms of timing and schedule, as staff members can conduct the survey at a time they deem appropriate, instead of rigidly following the dates the University prescribes. This suggested option, therefore, seems to offer the next best alternative for maintaining a balance between individual and institutional needs and concerns.

Developing a mobile version of the eSFQ system for students’ easier access to the survey could be another way to encourage student participation. With increasing accessibility to mobile technology and smart devices, a mobile version of the eSFQ system would offer an even greater degree of mobility, flexibility, and convenience for students to provide feedback that goes beyond the physical limitation of having to complete the survey using a desktop computer at a fixed location. However, boosting the response rates may only partially assuage the apprehension of academics about online evaluations as it is suggested that part of the reliability concern may lie in the use of survey results by personnel committees and administrators to inform important personnel decisions as well (McKeachie, 1997). In examining academics’ anxieties about online evaluations, Rienties (2014) argued that academics’ resistance to online evaluations does not necessarily stem from the surveying method but rather from the ‘dual nature’ of student evaluation – that is, the survey results are being used for both improvement and judgemental purposes, with increasing emphasis placed on the latter (Collings & Ballantyne, 2004; Ory, 2000; Avery et al.,

2006). Moreover, as Nulty (2008) notes, while the response rates obtained in course evaluation surveys are, in many cases, inadequate regardless of the mode of administration, more often than not, faculty and administrators (mis)use the survey results as a sole basis to appraise teaching effectiveness, while neglecting other factors such as sample size that will also affect the representativeness of the results, and misinterpreting the data when making inferences to inform important personnel decisions (Cohen, 1990; Theall, 2002).

Part of the remedy may lie in strengthening the teaching evaluation process. Apart from employing multiple strategies to encourage student participation to boost the response rates, academics, faculty and administrators should take into consideration other contextual factors that may affect survey results when interpreting the feedback obtained, and use multiple sources of teaching evidence in conjunction with the survey results to inform a more comprehensive view of a teacher's teaching effectiveness (Nulty, 2008) – especially when the evaluation is geared towards judgemental purposes regarding tenure, promotion and (re)appointment decisions. Furthermore, understanding different stakeholders' attitudes and perceptions about teaching evaluation may help to understand the complexities involved in the interplay of practicality and sentiments from various organisational issues (Cohen, 1990) and could shed light on a possible avenue for academics and administrators to find a middle ground.

There were some limitations to the study that need to be acknowledged, one of which was that the teachers whose subjects were included in this study were self-selected. Furthermore, not all students in selected subjects completed the teaching evaluation and some of the sample sizes for comparison were not large. However, as completing teaching evaluations for subjects is not compulsory at our university, in this regard, this would be true regardless of mode of administration and so reflects actual practice. While acknowledging these limitations to the study, the findings are useful for several reasons. First, the study examined response rates and student ratings together, comparing paper and online administration of the same survey. Second, the study was conducted at a university located in Asia – most of the previous work in this area has come from western universities, so findings will be of interest to other higher education institutions in the region which share similar characteristics.

In summary, the results from this study show that while response rates dropped substantially when the SFQ was administered online, this did not affect the ratings. Also, it does not seem that students wanting to express a negative view of the teacher or the subject rush to do so when left to complete the SFQ in their own time. These findings are encouraging, but convincing staff that the lower response rate does not negatively impact ratings and that eSFQ is 'a good thing' may be another matter. However, analysis of SFQ data that demonstrates this provides a good place to start.

REFERENCES

- Adams, M. J., & Umbach, P. D. (2012). Nonresponse and Online Student Evaluations of Teaching: Understanding the Influence of Salience, Fatigue, and Academic Environments. *Research in Higher Education*, 53(5), 576-591. doi:10.1007/s11162-011-9240-5.
- Anderson, J., Brown, G., & Spaeth, S. (2006). Online student evaluations and response rates reconsidered. *Innovate*, 2(6). Retrieved from <http://www.innovateonline.info/index.php?view=article&id=301>.

- Ardalan, A., Ardalan, R., Coppage, S., & Crouch, W. (2007). A comparison of student feedback obtained through paper-based and web-based surveys of faculty teaching. *British Journal of Educational Technology*, 38(6), 1085-1101. doi:10.1111/j.1467-8535.2007.00694.x.
- Avery, R. J., Bryant, W., Mathios, A., Kang, H., & Bell, D. (2006). Electronic Course Evaluations: Does an Online Delivery System Influence Student Evaluations? *The Journal of Economic Education*, 37(1), 21-37. doi:10.3200/JECE.37.1.21-37.
- Bennett, L., & Nair, C. S. (2010). A recipe for effective participation rates for web-based surveys. *Assessment & Evaluation in Higher Education*, 35(4), 357-365. doi:10.1080/02602930802687752.
- Capa-Aydin, Y. (2016). Student evaluation of instruction: comparison between in-class and online methods. *Assessment & Evaluation in Higher Education*, 41(1), 112-116. doi:10.1080/02602938.2014.987106.
- Cohen, P. A. (1990). Bringing research into practice. *New Directions for Teaching and Learning*, 1990(43), 123-132.
- Collings, D., & Ballantyne, C. (2004). Online Student Survey Comments: A Qualitative Improvement? Paper presented at the 2004 Evaluation forum, Melbourne, Australia. Retrieved from http://our.murdoch.edu.au/Educational-Development/_document/Publications/Eval_forum_paper.pdf.
- Crews, T. B., & Curtis, D. F. (2011). Online course evaluations: Faculty perspective and strategies for improved response rates. *Assessment & Evaluation in Higher Education*, 36(7), 865-878. doi:10.1080/02602938.2010.493970.
- Dommeyer, C. J., Baum, P., Hanna, R. W., & Chapman, K. S. (2004). Gathering faculty teaching evaluations by in-class and online surveys: their effects on. *Assessment & Evaluation in Higher Education*, 29(5), 611-323. doi:10.1080/02602930410001689171.
- Donovan, J., Mader, C. E., & Shinsky, J. (2006). Constructive student feedback: Online vs. traditional course evaluations. *Journal of Interactive Online Learning*, 5(3), 283-296.
- Goodman, J., Anson, R., & Belcheir, M. (2015). The effect of incentives and other instructor-driven strategies to increase online student evaluation response rates. *Assessment & Evaluation in Higher Education*, 40(7), 958-970. doi:10.1080/02602938.2014.960364.
- Hardy, N. (2003). Online ratings: Fact and fiction. *New Directions for Teaching & Learning*, 2003(96), 31-38.
- Hmieleski, K., & Champagne, M. V. (2000). Plugging in to Course Evaluation. *The Technology Source*. Retrieved from http://technologysource.org/article/plugging_in_to_course_evaluation/.
- IDEA Centre. (2011). Research notes no. 4: Paper versus online survey delivery. Retrieved from http://ideaedu.org/wp-content/uploads/2014/11/IDEA_Research_Notes_4-Online_v_Paper_Delivery.pdf.
- Johnson, T. D. (2003). Online student ratings: Will students respond? *New Directions for Teaching & Learning*, 2003(96), 49-59.
- Kulik, J. A. (2009). Response Rates in Online Teaching Evaluation Systems. Retrieved from https://www.wku.edu/senate/archives/archives_2015/e-4-1-response-rates-research.pdf
- Layne, B. H., DeCristoforo, J. R., & McGinty, D. (1999). Electronic versus traditional student ratings of instruction. *Research in Higher Education*, 40(2), 331-232. doi:10.1023/A:1018738731032.
- Leckey, J., & Neill, N. (2001). Quantifying quality: The importance of student feedback. *Quality in Higher Education*, 7(1), 19-32. doi:10.1080/13538320120045058.

- McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist*, 52(11), 1218-1225.
- Morrison, R. (2011). A comparison of online versus traditional student end-of-course critiques in resident courses. *Assessment & Evaluation in Higher Education*, 36(6), 627-641. doi:10.1080/02602931003632399
- NSSE (National Survey of Student Engagement). (2003). The College Student Report: 2003 Overview. Retrieved from http://nsse.indiana.edu/pdf/2003_inst_report/nsse_2003_overview.pdf.
- Nair, C. S., Adams, P., & Mertova, P. (2008). Student engagement: The key to improving survey response rates. *Quality in Higher Education*, 14(3), 225-232. doi:10.1080/13538320802507505.
- Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: what can be done? *Assessment & Evaluation in Higher Education*, 33(3), 301-314. doi:10.1080/02602930701293231.
- Ory, J. C. (2000). Teaching evaluation: Past, present, and future. *New Directions for Teaching and Learning*, 2000(83), 13-18.
- Powney, J., & Hall, S. (1998). Closing the loop: The impact of student feedback on students' subsequent learning. Edinburgh: Scottish Council for Research in Education. Retrieved from https://www.researchgate.net/publication/234775216_Closing_the_Loop_The_Impact_of_Student_Feedback_on_Students'_Subsequent_Learning_Research_Report_Series.
- Rienties, B. (2014). Understanding academics' resistance towards (online) student evaluation. *Assessment & Evaluation in Higher Education*, 39(8), 987-1001. doi:10.1080/02602938.2014.880777.
- Risquez, A., Vaughan, E., & Murphy, M. (2015). Online student evaluations of teaching: what are we sacrificing for the affordances of technology? *Assessment & Evaluation in Higher Education*, 40(1), 102-134. doi:10.1080/02602938.2014.890695.
- Sorenson, D. L., & Reiner, C. (2003). Charting the uncharted seas of online student ratings of instruction. *New Directions for Teaching and Learning*, 2003(96), 1-24.
- Stowell, J. R., Addison, W. E., & Smith, J. L. (2012). Comparison of online and classroom-based student evaluations of instruction. *Assessment & Evaluation in Higher Education*, 37(4), 465-473. doi:10.1080/02602938.2010.545869.
- Theall, M. (2002). Student ratings: Myths vs. research evidence. (L. D. Sorenson, Ed.) *Focus on Faculty*, 10(3), 2. Retrieved from <https://studentratings.byu.edu/info/faculty/myths.asp>.