

University Students Refuse To Give Useful Feedback Under Compulsory Evaluation Systems

Huang Zexi*

Yingcai Experimental School, University of Electronic Science and Technology of China

Abstract

Previous research in students evaluation of teaching (SET) has revealed tight relationships between participants perception, motivation and quality of their feedback. This study takes compulsoriness of SET systems into consideration and aims to investigate its impact on SET through the factors mentioned before. We used Internet-based questionnaires and conducted correlation and regression analyses, which showed a significant correlation between compulsoriness and motivation and a tendency to significance between compulsoriness and feedback. These indicated that under a compulsory SET system, students motivation is negatively affected and as a consequence of their casual involvement, the feedback they yield is possibly useless. These findings provide evidence that reliability and validity of SET can be ruined where compulsory systems are adopted.

1 Introduction

Students' evaluation of teaching system has enjoyed its prosperity ever since it was invented, and nowadays it becomes a routine activity in colleges and universities around the globe because of its formative and summative functions in the field of education (Spooren et al., 2013). Therefore, validity and reliability of SET systems have become a major concern among universities, which are studied in detail in aspects such as instruments (Marsh, 1982, 1987;

Spooren et al., 2014), analytical methods (Pozo-Munoz et al., 2000; Radmacher & Martin, 2001; Barth, 2008; Onwuegbuzie et al., 2009) and result interpretation (Pratt et al., 1999; Felton* et al., 2004; Clayson & Sheffet, 2006; Feldman, 2007; Otani et al., 2012).

However, even with appropriate instruments and effective methods in analyzing and interpreting, satisfactory results cannot be guaranteed when participants of SET systems, i.e. university students are biased (Feldman, 1977). Kember* et al. (2004) have revealed that students submit their rating according to their own conceptions, which indicates students' evaluation are basically subjective. While factors like teachers' leniency and after-class workload are not generators of bias (Marsh & Roche, 2000), it has been pointed out that students' motivation for taking part in evaluation procedure and perception of the system are remarkably essential since "the usefulness of SET data will be severely undermined unless students willingly provide quality input" (Chen & Hoshower, 2003) and perception of students could be "a relatively important source of rating bias" (Worthington, 2002). All these studies led to a conclusion that students' willingness to evaluate the teaching effectiveness is *sine qua non* if the SET system is expected to function well.

Therefore, here, we took students willingness into account and investigated the influence from an important characteristic of the system itself, that is, compulsoriness: we tried to ascertain the difference of students' behaviors in SET between different levels of compulsoriness of the SET systems they participated in, as well as reasons for the discrepancy. Hence, participants in this study were required to fill in a questionnaire which consisted of five dimensions: compulsoriness of the system in their university, their motivation for participation in evaluation of

*Corresponding author. Tel.: +86 134 3881 7133
E-mail address: Eitima@163.com (Huang Zexi)

teaching effectiveness, their perception of usefulness of the system, their perception of compulsoriness in SET and quality of their feedback respectively.

We conducted correlation and regression analysis in the research in order to examine to what extent compulsoriness posed an impact on other dimensions. Considering students under compulsory systems are obliged to present their feedback instead of out of their own willingness, we predicted that their motivation for taking part would be lower than their counterparts who made their own decision on whether to join in or not and their feedback more invalid. However, we had no prediction regarding the relation between compulsoriness and students' perception because adopting compulsoriness itself could be considered a method to improve participation rate and thus enhance effectiveness of the system.

2 Method

2.1 Participants

A total of 145 Chinese undergraduates (eighty-eight males) were recruited from different universities in both Chinese mainland and Hong Kong. Internet-based informed consent was given to participants, as well as a formal announcement that any information concerning personal privacy wouldn't be exposed to the authorities of their universities. In addition, all participants reported to have at least once taken part in the SET systems of their universities in the past three months and possess basic knowledge of filling in the questionnaire on the Internet.

2.2 Instrument

We designed the instrument of the study, which was composed of five dimensions, compulsoriness of SET in participants' universities (4 items), participants' motivation for SET (5 items), participants' perception of SET systems (5 items), participants' perception of compulsoriness (4 items) and participants' feedback (6 items) in SET. All the items were scored on a 5-point Likert scale ranging from strongly disagree (1) to strongly agree (5). Post-hoc reliability analysis was conducted and all but the first dimension were significant reliable ($\alpha > .8$). Items in the first dimension were designed to find whether the universi-

ties adopted a given policy to ensure students' participation in SET. Since these policies could vary dramatically among different universities, the statistical reliability analysis for the first dimension was not conducted.

2.3 Procedure

We administered the questionnaires to participants based on an online survey website Sojump (www.sojump.com). Participants read instructions in the very beginning before filling in the form. There was no time limit for the filling procedure. When each participant completed and submitted the questionnaire, his or her feedback was recorded automatically on the database of Sojump and was retrieved later.

2.4 Data analysis

First, according to their total score in the dimension of compulsoriness, participants were divided into two groups (See Table 1), compulsory group (total score > 12) and non-compulsory group (total score ≤ 12). Following independent t-tests with respect to other dimensions were conducted between the two groups to find whether the respective levels of the two groups were significantly different. Second, for the purpose of investigating the potential correlations between all dimensions, Pearson product-moment indices were calculated. Last, dimensions significantly correlated with feedback ($p < .05$) were put into multiple regression to see whether they held predictive power. It's noteworthy that negative statements in all dimensions had been reversely scored before any statistical analyses were conducted.

3 Results

3.1 Mean comparison results

Groups were assumed to have no significant difference in demographic or personality characteristics. Thus, to examine whether or not subjects in different groups showed significant distinction in other four dimensions, independent t-tests were conducted. Mean comparison results (see Figure 1, Table 2) showed that for motivation (mean difference = -2.484 , $p = .048$), S-perception (mean difference

Table 1. Descriptive statistics of motivation score in two groups

| Groups | Mean | N | Std. Deviation | % of Total N | Range |
|----------------|-------|-----|----------------|--------------|-------|
| Compulsory | 16.54 | 123 | 2.28 | 84.8% | 13-20 |
| Non-compulsory | 9.73 | 22 | 2.57 | 15.2% | 4-12 |
| Total | 15.50 | 145 | 3.37 | 100.0% | 4-20 |

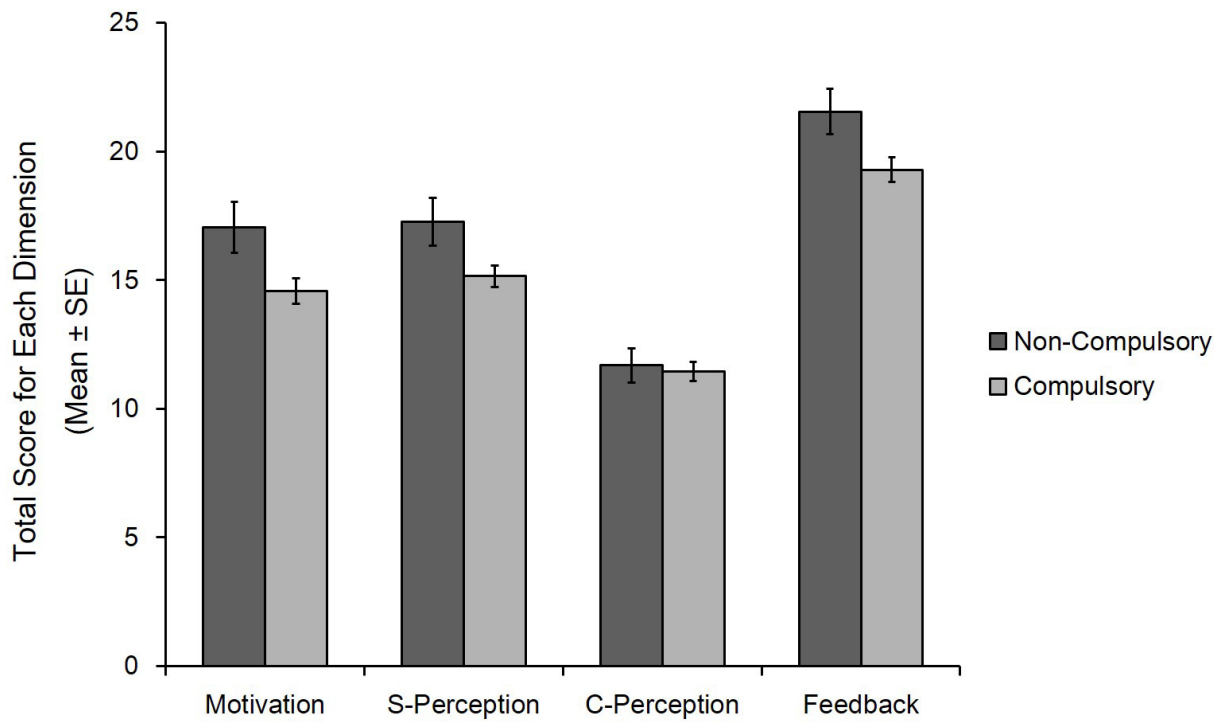


Figure 1. Total score of motivation, S-perception (perception of SET systems), C-perception (perception of compulsoriness) and feedback. Error bars depict SEs.

= -2.126, $p = .045$) and feedback (mean difference = -2.253, $p = .056$), differences between groups were significant. However, participants reported similar attitudes towards compulsoriness (mean difference = -.227, $p = .805$).

3.2 Correlational findings

For the purpose of obtaining the complete insight of the potential relationship among SET compulsoriness, students' motivation, students' perception of SET, students' perception of compulsoriness and students' feedback in SET, Pearson correlations between each two variables were computed (See Table 3). Students' high-quality feedback tended to be impeded by low motivation ($p < .01$, $r = .674$), negative perception of SET systems ($p < .01$, $r = .684$) and compulsory SET policies ($p < .1$, $r = -.147$). However, though compulsoriness adversely affected students' motivation ($p < .01$, $r = -.229$), no significant correlation was observed between compulsoriness and S-perception ($p > .1$, $r = .069$).

3.3 Regression findings

Since we found motivation and perception of SET systems were significantly and directly correlated with feedback, stepwise regression analysis was undertaken to determine what proportion of feedback variance could be explained by these variables. Two one-predictor regressions were conducted respectively (see Figure 2). Perception of SET alone accounted for a total of 46.4% of variance ($\beta = .684$, $p = .000$), and motivation accounted for 45.4% ($\beta = .674$, $p = .000$). When combining the two variables together, a total of 59.4% of variance was explained ($\beta_1 = .449$, $\beta_2 = .426$, $p = .000$), which showed the high predictive power of motivation and S-perception.

4 Discussion

This study investigated how compulsoriness of SET systems affects students' motivation for SET, students' perception of SET, and students' feedback on SET. We designed an Internet-based questionnaire, through which participants reported their feelings about the evaluation process and their attention level when assessing their teachers. As

hypothesized before, students obliged to evaluate the effectiveness of teaching were less motivated and presented their feedback more casually than their counterparts who had the choice to leave no comments without being punished. It is noteworthy that regardless of the compulsoriness level of their universities, all participants had at least once taken part in SET systems in the past three months.

This finding highlights the impact of compulsoriness on university students who make up the majority of SET participants by examining their motivation and perception levels. Mean comparisons between groups showed that generally students whose universities apply SET as compulsory turn to have modulated motivation for assessment tasks, negative perception of the systems and also invalid submissions of feedback. This might be explained by participants' disappointment with compulsoriness regardless of their groups.

We observed that students' motivation and S-perception were significantly correlated with students' feedback, and so was compulsoriness with motivation. A tendency to significance was also found between compulsoriness and feedback, further identifying negative impact on SET of compulsoriness. However, no such relation was found between compulsoriness and S-perception. Thus, we contend that compulsoriness undermines the authenticity of students feedback mainly by reducing their motivation to a lower level instead of jeopardising their faith in summative and formative functions of SET. In addition, multiple regression analyses demonstrated that the quality of feedback is primarily determined by motivation and S-perception, indicating for decision makers of universities to improve quality of SET feedback, creating motivators or changing students' opinions may be efficient ways. Although it seems that abolishing compulsory systems might be the optimal approach to improving validity of SET, issues such response rates should also be taken into consideration. Thus, for administrations who insist on maintaining compulsory policies, we suggest that they take measures in aspects of promoting students' perception of SET as alternatives, such as "showing students that their feedback has been used successfully", "listing prominently the uses of the teaching evaluation on the evaluation instrument" (Chen & Hoshower, 2003), etc., since compulsoriness influences students' motivation enormously while makes little difference to perception.

It is noteworthy that dimensions of the instrument used

Table 2. T test for equalities of means between compulsory group and non-compulsory group.

| | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
|---------------------------|-----------------|-----------------|-----------------------|---|-------|
| | | | | Lower | Upper |
| Motivation | .048 | -2.484 | 1.243 | -4.942 | -.027 |
| S-Perception ^a | .045 | -2.126 | 1.054 | -4.209 | -.043 |
| C-Perception ^b | .805 | -.227 | .917 | -2.040 | 1.587 |
| Feedback | .056 | -2.253 | 1.171 | -4.567 | .062 |

^a Perception of SET systems

^b Perception of compulsoriness

Table 3. Pearson correlations among compulsoriness, motivation, perception of SET and feedback.

| Variables | Compulsoriness | Motivation | S-Perception | Feedback |
|----------------|----------------|------------|--------------|----------|
| Compulsoriness | — | | | |
| Motivation | -.229** | — | | |
| S-Perception | -.069 | .552** | — | |
| Feedback | -.147* | .674** | .684** | — |

* $p < .1$

** $p < .01$

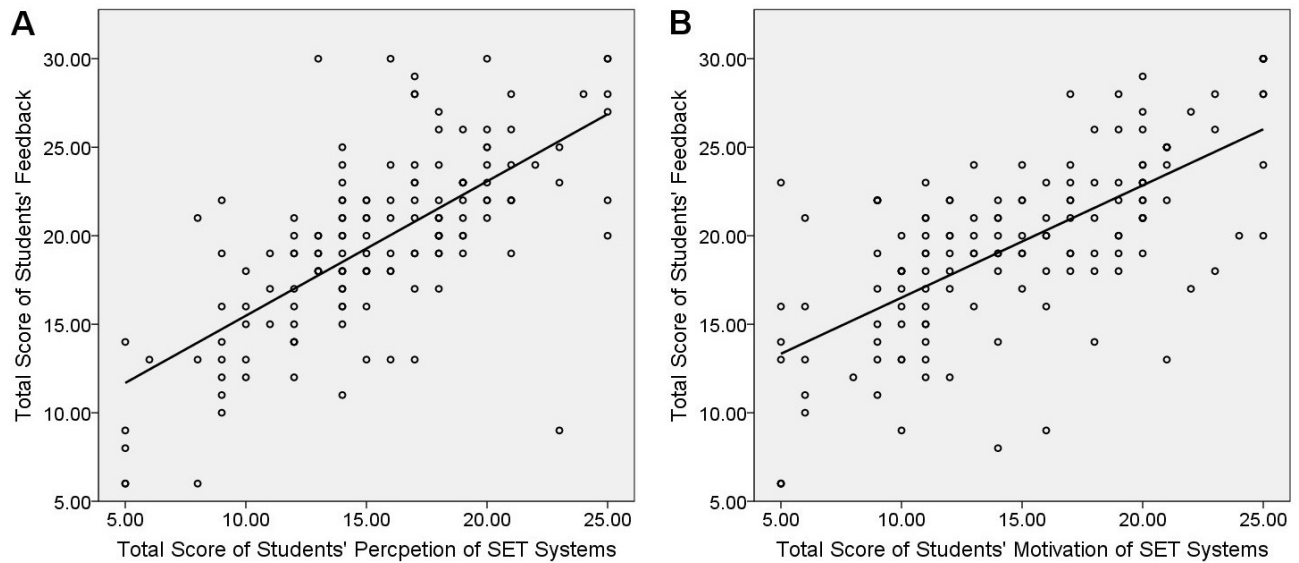


Figure 2. One-predictor regression models. A, Regression between feedback and perception of SET. B, Regression between feedback and perception of SET.

in this study were quite general to gain a quick insight of the mechanism, therefore it may be not enough to draw the whole picture under the phenomenon. In addition, students' background was not strictly controlled, which may affect the results. Future research with a dedicated instrument and a detailed standard for participants is needed.

5 Conclusion

To summarize, our study demonstrates a tangible modulation of motivation and feedback quality from students under a compulsory SET system, which renders a threat to validity and reliability of SET. It is also remarkable that students' perception of SET, a *sine qua non* for high-quality feedback, is basically unchanged regardless of compulsory policies applied in their universities. This provides a solution for authorities who hope to collect useful SET information while maintaining compulsory policies as they take measures to promote students' perception as alternatives.

Acknowledgements

We thank Liu Hao and Zhang Jiahua for their assistance of recruiting participants and their advice on questionnaire design. In addition, an earlier draft of this paper was prepared during the first semester in 2015. Without implicating them for any remaining errors and oversights appearing in the paper, we want to express our gratitude to Ju Hao, Lei Chucai, Chen Wenjun, Wang Haoming for useful comments, suggestions and critiques. And we are also deeply grateful to Gao Zhao and Gao Shan, under whose guidance and supervision this research was conducted.

References

- Barth, M. M. (2008). Deciphering student evaluations of teaching: A factor analysis approach. *Journal of Education for Business*, 84(1), 40–46.
- Chen, Y., & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment & evaluation in higher education*, 28(1), 71–88.
- Clayson, D. E., & Sheffet, M. J. (2006). Personality and the student evaluation of teaching. *Journal of Marketing Education*, 28(2), 149–160.
- Feldman, K. A. (1977). Consistency and variability among college students in rating their teachers and courses: A review and analysis. *Research in Higher Education*, 6(3), 223–274.
- Feldman, K. A. (2007). Identifying exemplary teachers and teaching: Evidence from student ratings¹. In *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 93–143). Springer.
- Felton*, J., Mitchell, J., & Stinson, M. (2004). Web-based student evaluations of professors: The relations between perceived quality, easiness and sexiness. *Assessment & Evaluation in Higher Education*, 29(1), 91–108.
- Kember*, D., Jenkins, W., & Chi Ng, K. (2004). Adult students' perceptions of good teaching as a function of their conceptions of learning part 2. implications for the evaluation of teaching. *Studies in Continuing Education*, 26(1), 81–97.
- Marsh, H. W. (1982). Seeq: A reliable, valid, and useful instrument for collecting students' evaluations of university teaching. *British journal of educational psychology*, 52(1), 77–95.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International journal of educational research*, 11(3), 253–388.
- Marsh, H. W., & Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluations

- of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, 92(1), 202.
- Onwuegbuzie, A. J., Daniel, L. G., & Collins, K. M. (2009). A meta-validation model for assessing the score-validity of student teaching evaluations. *Quality & Quantity*, 43(2), 197–209.
- Otani, K., Kim, B. J., & Cho, J.-I. (2012). Student evaluation of teaching (set) in higher education: How to use set more effectively and efficiently in public affairs education. *Journal of Public Affairs Education*, 531–544.
- Pozo-Munoz, C., Reboloso-Pacheco, E., & Fernandez-Ramirez, B. (2000). The 'ideal teacher'. implications for student evaluation of teacher effectiveness. *Assessment & Evaluation in Higher Education*, 25(3), 253–263.
- Pratt, D., Kelly, M., & Wong, W. (1999). Chinese conceptions of 'effective teaching' in hong kong: Towards culturally sensitive evaluation of teaching. *International Journal of Lifelong Education*, 18(4), 241–58.
- Radmacher, S. A., & Martin, D. J. (2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. *The Journal of psychology*, 135(3), 259–268.
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83(4), 598–642.
- Spooren, P., Mortelmans, D., & Christiaens, W. (2014). Assessing the validity and reliability of a quick scan for student's evaluation of teaching. results from confirmatory factor analysis and g theory. *Studies in Educational Evaluation*, 43, 88–94.
- Worthington, A. C. (2002). The impact of student perceptions and characteristics on teaching evaluations: a case study in finance education. *Assessment & Evaluation in Higher Education*, 27(1), 49–64.