



Received: 27 March 2016  
Accepted: 08 June 2016  
Published: 30 June 2016

\*Corresponding author: Ahmed Al Kuwaiti, Deanship of Quality and Academic Accreditation, University of Dammam, P.O. Box 40065, Al-Khobar 31952, Saudi Arabia  
E-mails: [akuwaiti@uod.edu.sa](mailto:akuwaiti@uod.edu.sa), [qaa@uod.edu.sa](mailto:qaa@uod.edu.sa)

Reviewing editor:  
Jody S. Piro, WCSU, USA

Additional information is available at the end of the article

## EDUCATIONAL ASSESSMENT & EVALUATION | RESEARCH ARTICLE

# Understanding the effect of response rate and class size interaction on students evaluation of teaching in a higher education

Ahmed Al Kuwaiti<sup>1\*</sup>, Mahmoud AlQuraan<sup>2</sup> and Arun Vijay Subbarayalu<sup>1</sup>

**Abstract:** *Objective:* This study aims to investigate the interaction between response rate and class size and its effects on students' evaluation of instructors and the courses offered at a higher education Institution in Saudi Arabia. *Study Design:* A retrospective study design was chosen. *Methods:* One thousand four hundred and forty four different courses belonging to all the Colleges ( $N = 21$ ) located across seven different campuses of University of Dammam (UOD) were considered in this study. All the course evaluation surveys (CES) ( $N = 168,574$ ) conducted during the academic year 2013–2014 were analyzed to investigate response rate and class size interaction effect on students evaluation of courses and instructors. *Results:* It is observed that when the class size is at the medium level, the ratings of instructors and courses increase as the response rate increases. On the contrary; when the class size is small, a high response rate is required for the evaluation of instructors and at least medium response rate is required for evaluation of courses. The study suggests that the interaction between response rate and class size is an important factor that needs to be taken into account while interpreting the students' evaluation of instructors and courses. *Originality:* This research study examined the effect of interaction



Ahmed Al Kuwaiti

### ABOUT THE AUTHORS

Dr Ahmed Al Kuwaiti is presently working as an Associate Professor and General Supervisor of the Deanship of Quality and Academic Accreditation, University of Dammam, Kingdom of Saudi Arabia. His research interests lie predominantly within the domain of medical education and tends to be focused on quality of higher education, accreditation and academic research.

Dr Mahmoud Al-Quraan working as Associate Professor in the Department of Counseling and Educational Psychology, College of Education, Yarmouk University, Irbid-Jordan. His primary research interest focuses on educational assessment, educational pedagogy, and Psychometrics.

Dr Arun Vijay Subbarayalu is working as Assistant Professor at Deanship of Quality and Academic Accreditation, University of Dammam and his research interest is to improve the quality of higher education using six sigma methods. He is an academician with more than a decade of experience in teaching, research and has developed many innovative teaching methods.

### PUBLIC INTEREST STATEMENT

University of Dammam (UOD) is currently performing several evaluations by students as a measure to improve the quality of the institution as well as its programs. While carrying out these evaluations, several factors influence the outcome of students' evaluations and most critical one is the influence of class size and response rate of students. This research paper addresses the interaction of class size and response rate and it concluded that class size has a significant effect on the students' evaluation of instructors and courses in Higher Education. This study will help the higher education institutions to understand the optimal response rate required for different class sizes while planning students' evaluation surveys of courses and instructors.

between the response rate (RR) and class size (CS) on students' evaluation of instructors and courses in higher education. This study will help the policy planners of higher education institutions to understand the response rate required for different class sizes while planning students' evaluation surveys for courses and instructors.

**Subjects:** Education; Quality Management; Statistics & Probability

**Keywords:** higher education; student evaluation of teaching; response rate; class size; Saudi Arabia

## 1. Introduction

Students' Evaluation of Teaching Effectiveness is one of the most widely accepted indicators of measuring the Quality of Higher Education worldwide, and it has gained "marvelous" attention in the field of psychology, quality control, and quality assurance in the last few decades (Ginns, Prosser, & Barrie, 2007; Vijay, 2014). The process of improving the quality of Higher Education is a dynamic one, and universities ought to continuously improve their teaching based on the students' perceptions (Konting, Kamaruddin, & Man, 2009). Utilizing students perception in improving the quality of Higher Education is a commonly utilized practice in almost every university across the globe (Zabaleta, 2007). Higher Education Institutions (HEIs) in the Kingdom of Saudi Arabia (KSA) are also becoming increasingly aware of the importance of quality in the knowledge delivery due to increasing numbers of students entering the educational institutions (Al-Kuwaiti & Subbarayalu, 2015). Since students are the individuals who are most exposed to and the most affected by the teacher's teaching, their perceptions and point of views about both instructors and courses are of paramount importance. Research also indicates that students are the most qualified sources of data and information about teaching and learning settings (Archibong & Nja, 2011). Thus, the quality of university courses should certainly be evaluated by its recipients (Nikolaidis & Dimitriadis, 2014) and these evaluations need to include all the elements of the teaching and learning processes (Wachtel, 1998; Chen & Hoshower, 2003; Clayson, 2009; Berk, 2005). This practice of evaluating the courses through students feedback has been included as one of the key mechanisms in the internal quality-assurance processes as a way of demonstrating institution's performance in accounting and auditing practices (Blackmore, 2009; Johnson, 2000). In order to facilitate that, several instruments are available to measure the level of student satisfaction of the courses (Coffey & Gibbs, 2001; Ramsden, 1991).

Traditionally, these surveys comprised a series of closed-ended questions about courses and teaching effectiveness with at least one question pertaining to overall teaching effectiveness. The surveys were typically anonymous and mostly conducted near the end of the semester using either a paper-based or electronic format (Kogan, Schoenfeld-Tacher, & Hellyer, 2010). Once the data were collected, reports were generated across the instructors, departments, and colleges, which are considered as evidence of teaching effectiveness (Sproule, 2000). An important part of any such evaluation is the communication of results in such a way that allows fair and meaningful interpretations and comparisons of results to make judgment on quality of teaching, career advancement, and funding of teaching (Burden, 2008; Kuzmanovic, Savic, Andric Gusavac, Makajic-Nikolic, & Panic, 2013; Neumann, 2000). These evaluations have contributed to quality in the educational process, especially if the proper reliability coefficients are used to assess the psychometric properties of the instruments (Al-Kuwaiti, 2014; Burden, 2008; Morley, 2014).

There are several reasons that make HEI's use students' evaluations and assessments (Spooren, Brockx, & Mortelmans, 2013), namely (i) quick feedback assuming that instructors make changes based on students' evaluation; (ii) students evaluation is used for critical decisions such as promotion; and (iii) for accreditation and governmental agencies that require such evaluations. Besides this there are several benefits that the students' evaluations can have for institutions, some of the benefits being: (i) instructors value the input and make improvements in their teaching; (ii) instructors are rewarded for having excellent ratings; (iii) instructors with very low ratings are encouraged to seek help; (iv) students perceive and use evaluations as a way to suggest improvements in

teaching; (v) students have more information on which to make their course selections; and (vi) evaluations motivate instructors to improve teaching (Archibong & Nja, 2011; Neumann, 2000; Ory & Ryan, 2001). University administrators also utilize students' evaluations to make administrative decisions; despite the fact that a number of administrators expressed concern about the validity and value of student evaluations of teaching (Beran, Violato, & Kline, 2007; Kogan & Shea, 2007).

A previous study conducted by Addison, Best, and Warrington (2006) identified three key factors influencing students' rating during the course evaluation survey (CES) (i.e. Course, instructors, and students) and examined the relationship between the students' perceptions related to course difficulty and expected grades. The results revealed that perceived difficulty is associated with grade expectations and the ratings that students give on formal evaluations. Students who have high academic achievement evaluated instructors higher than those who have lower academic achievement. Further, it is also observed that regardless of the academic achievement, higher evaluations were given by students who found the course easier than expected compared to those who found it harder than initially anticipated.

Several studies have been conducted to study the effect of class size on the outcome of students' evaluation of courses in higher education. Gleason (2012) demonstrated that medium classes (i.e. 30–55 students) had little to no benefit over large classes (i.e. 110–130 students) in student learning and student achievement, whereas large classes have small-to-medium positive effect over medium classes in the area of student satisfaction. The only area in which the small classes had a small positive effect was in the area of student engagement, while Pezzella, Paladino, Zoller, and Mandery (2014) compared the performance of students in large-sized classes to students in small classes to assess the efficacy of student learning in large classes. The results of that study indicated that large classes are as efficacious as small classes and that student performance is quite nuanced. On exploring the existing literature, no studies have examined the effect of Interaction between the response rate (RR) and class size (CS) on students' evaluation of Instructors. Thus, the primary objective of this study is to find out the effect of interaction between response rate and class size on students' evaluation of courses and instructors.

## 2. Methodology

### 2.1. The instrument

In this study, CES used to collect the data, which contains 14 five-point Likert Scale items, divided into two subscales (instructor and course related items). The questionnaire has been developed based on the guidelines of the National Commission for Academic Accreditation and Assessment (NCAAA) in Saudi Arabia for the purpose of accreditation. CES was developed by a panel of experts in the related areas, and several studies investigated its psychometric properties and usefulness (Al Rubaish, Wosornu, & Dwivedi, 2011, 2012; Al-Kuwaiti & Maruthamuthu, 2014). Besides that, Corrected-Item-To-Total Correlation and Cronbach's Alpha were calculated from a random sample ( $n = 50$ ) selected from the current data. The results show that Cronbach's Alpha equals 0.963 and the Corrected-Item-To-Total Correlation ranges from 0.568 to 0.888 which adds evidences for the reliability and the validity of CES.

### 2.2. Data collection

University of Dammam (UOD) through the Deanship of Quality and Academic Accreditation (DQAA) has developed a special application called "UDQUEST" to collect data electronically, which is related to several attributes of the courses offered at different programs in the university. Dommeyer, Baum, Hanna, and Chapman (2004) found that online evaluations do not produce significantly different mean evaluation scores than traditional in-class evaluations, even when different incentives are offered to students to complete online evaluations. Contrary to this, Capa-Aydin (2014) indicated that in-class evaluations have a significantly higher response rate than online evaluation. In addition, Rasch's analysis showed that mean ratings in in-class evaluation were significantly higher than those in online evaluations. Several evaluations have been conducted at UOD for the purpose of accreditation, as stipulated by NCAAA. One among them is the CES, which is used to evaluate courses

and the related instructors at the end of every semester to all the students who registered in every course offered in that semester.

A total of 1,443 different undergraduate courses belonging to all the colleges ( $N = 21$ ) located across seven different campuses of UOD were considered in this study. Accordingly, all the CES surveys ( $N = 168,574$ ) conducted during the academic year 2013–2014 were included in the analysis.

### 2.3. Statistical analysis

A Two-Way Analysis of Variance (Full model) was used to investigate whether there is any significant difference in students' evaluations with respect to different class sizes and response rates. Students' evaluations of both instructors and courses are considered as dependent variables and the independent variables were class size and response rate. "CS" is defined as the number of students registered in the course or the number of students who targeted to fill the survey whereas "RR" is defined as the proportion of students who responded to the survey with respect to the number of students who were supposed to fill the survey. If the main effect (i.e. neither "RR" nor "CS") and the interaction effect (both "RR" and "CS") are found significant, then the Tukey Post Hoc test was used for pairwise comparison. In this study, the response rate and class size were categorized into three levels; 1 - low [class size less than 60 and response rate less than 73%]; 2 - medium [class size (61–200) and response rate (74–91) and]; 3 - high [class size over 200 and response rate over 92%] (Nulty, 2008). The above classification was designed assuming 3% sampling error and 95% confidence level conditions (Nulty, 2008).

### 3. Results

Table 1 shows Descriptive statistics of students' evaluation of instructors according to class size and response rate. In order to explore the interaction effect between response rate and class size on students' evaluation of instructors and courses, the descriptive statistics of two independent variables (response rate and class size) with respect to the dependent variable i.e. mean score of two sub-scale (i.e. effectiveness of the instructors and courses) were calculated using the statistical package for the social science (SPSS), version 19.

**Table 1. Descriptive statistics of students' evaluation of instructors according to class size and response rate at UOD**

Response rate	Class size	Mean	Std. deviation	N
Low response rate	Small class size	3.88	1.07	12,937
	Medium class size	3.88	1.07	30,880
	Large class size	3.89	1.08	21,710
	Total	3.88	1.07	65,527
Medium response rate	Small class size	3.92	1.03	16,094
	Medium class size	3.90	1.01	26,916
	Large class size	3.97	1.03	28,428
	Total	3.93	1.02	71,438
High response rate	Small class size	3.96	0.98	14,311
	Medium class size	4.00	0.94	8,483
	Large class size	3.96	0.98	8,815
	Total	3.97	0.97	31,609
Total	Small class size	3.92	1.03	43,342
	Medium class size	3.91	1.03	66,279
	Large class size	3.94	1.04	58,953
	Total	3.92	1.03	168,574

Table 1 shows that the mean ratings of students' evaluation of instructors which ranges from 3.88 to 4.00. The lowest mean is [3.88] when the response rate is low across all level of the class size. Likewise, highest mean is [4.00] when the response rate is high and the class size is medium. A Two-Way ANOVA was used to test the differences between these means and the results are presented in Table 2.

From Table 2, it is shown that both the main effects (response rate; class size) as well as the interaction effect are statistically significant and hence the data are further subjected to pairwise comparisons.

From Table 3, it is observed that the highest mean differences in students rating is found between high and low response rate, taken into account that all the mean difference are statistically significant.

**Table 2. Two-Way ANOVA summary table where the dependent variable is students' evaluation of instructors and the independent variables are response rate and class size**

Source	Sum of squares	df	Mean square	F	Sig.	Partial eta squared
Response rate	167.584	2	83.792	78.437	.000	.001
Class size	12.037	2	6.019	5.634	.004	.000
Response rate × Class size	41.116	4	10.279	9.622	.000	.000
Error	180,073.198	168,565	1.068			
Total	2,773,328.395	168,574				

**Table 3. Pairwise comparisons showing the mean difference in the rating of students' evaluation of instructors between different response rates**

(I) Response rate	(J) Response rate	Mean difference (I-J)	Std. error	Sig.
Low response rate	Medium response rate	-.047	.006	.000
	High response rate	-.089	.007	.000
Medium response rate	Low response rate	.047	.006	.000
	High response rate	-.042	.007	.000
High response rate	Low response rate	.089	.007	.000
	Medium response rate	.042	.007	.000

**Table 4. Pairwise comparisons showing the mean difference in the rating of students' evaluation of instructors between different class sizes**

(I) Class size	(J) Class size	Mean difference (I-J)	Std. error	Sig.
Small class size	Medium class size	-.009	.007	.192
	Large class size	-.023	.007	.001
Medium class size	Small class size	.009	.007	.192
	Large class size	-.014	.007	.037
Large class size	Small class size	.023	.007	.001
	Medium class size	.014	.007	.037

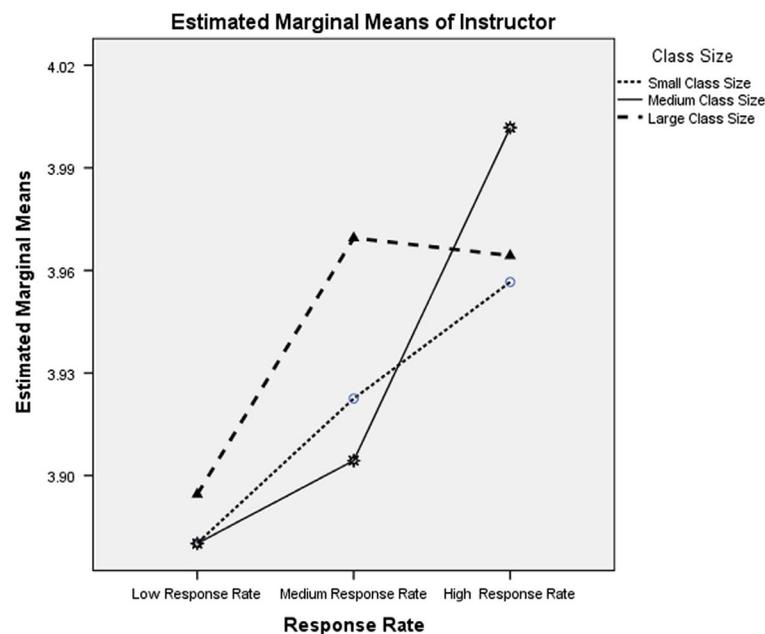
From the above Table 4, it is observed that the mean students rating differ significantly between large class size and smaller or medium class sizes. Such result is in conformance with the finding of previous studies done by Koh and Tan (1997) and Badri, Abdulla, Kamali, and Dodeen (2006). The investigators explored further to study the interaction effect of response rate and class size on the students rating of instructors.

From Figure 1, it is inferred that when the class size is medium or small, rating of students evaluation of instructors increases as long as the response rate increases. On the other hand; when the class size is large, rating of students evaluation of instructors increases when the response rate is moving from low response rate to medium response rate, and then it gets stable when it is moving to high response rate. Across all levels of response rate, the mean rating of students' evaluation of instructors is found to be less for the small class size than that for the large class size. This suggests that when the class size is small, the response rate needs to be at least at the medium level to have stable students' evaluation of instructors, and the medium response rate is sufficient, when the class size is large.

Table 5 shows descriptive statistics of students' evaluation of courses according to response rate and class size. It ranges from 3.66 to 3.88. The lowest mean ratings is observed (i.e. 3.66) when the response rate is low and the class size is small, while the highest rating (i.e. 3.88) is found when the response rate is high and the class size is medium. Further, a Two-Way ANOVA was used to test whether observed mean differences are statistically significant and the results are shown in Table 6.

From Table 6, it is inferred that both the main effects (i.e. CS and RR) and the interaction effect (i.e. CS and RR) are statistically significant. Further, the data is subjected to pairwise comparison and the results are shown in Tables 7 and 8.

**Figure 1. The interaction between response rate and class size on rating of students' evaluation of instructors.**



**Table 5. Descriptive statistics of students' evaluation of courses according to class size and response rate at UOD**

Response rate	Class size	Mean	Std. deviation	N
Low response rate	Small class size	3.66	1.09	12,937
	Medium class size	3.69	1.12	30,880
	Large class size	3.71	1.11	21,710
	Total	3.69	1.11	65,527
Medium response rate	Small class size	3.73	1.07	16,094
	Medium class size	3.76	1.05	26,916
	Large class size	3.79	1.06	28,428
	Total	3.77	1.06	71,438
High response rate	Small class size	3.73	1.04	14,311
	Medium class size	3.88	0.98	8,483
	Large class size	3.81	1.01	8,815
	Total	3.79	1.02	31,609
Total	Small class size	3.71	1.07	43,342
	Medium class size	3.74	1.08	66,279
	Large class size	3.76	1.07	58,953
	Total	3.74	1.07	168,574

**Table 6. Two-Way ANOVA summary table where the dependent variable is students' evaluation of courses and the independent variables are response rate and class size**

Source	Sum of squares	df	Mean square	F	Sig.	Partial eta squared
Response rate	311.279	2	155.640	135.456	.000	.002
Class size	118.806	2	59.403	51.700	.000	.001
Response rate × Class size	72.856	4	18.214	15.852	.000	.000
Error	193,681.231	168,565	1.149			
Total	2,556,210.120	168,574				

**Table 7. Pairwise comparisons showing the mean difference in the students' evaluation of courses between different response rates**

(I) Response rate	(J) Response rate	Mean difference (I-J)	Std. error	Sig.
Low response rate	Medium response rate	-0.071	.006	.000
	High response rate	-0.118	.008	.000
Medium response rate	Low response rate	0.071	.006	.000
	High response rate	-0.047	.007	.000
High response rate	Low response rate	0.118	.008	.000
	Medium response rate	0.047	.007	.000

Table 7 shows the mean differences in students rating between three different response rates. It is observed that all the mean difference are statistically significant at 0.05 levels. Specifically, highest mean difference is observed between the high and low response rate (i.e. 0.118).

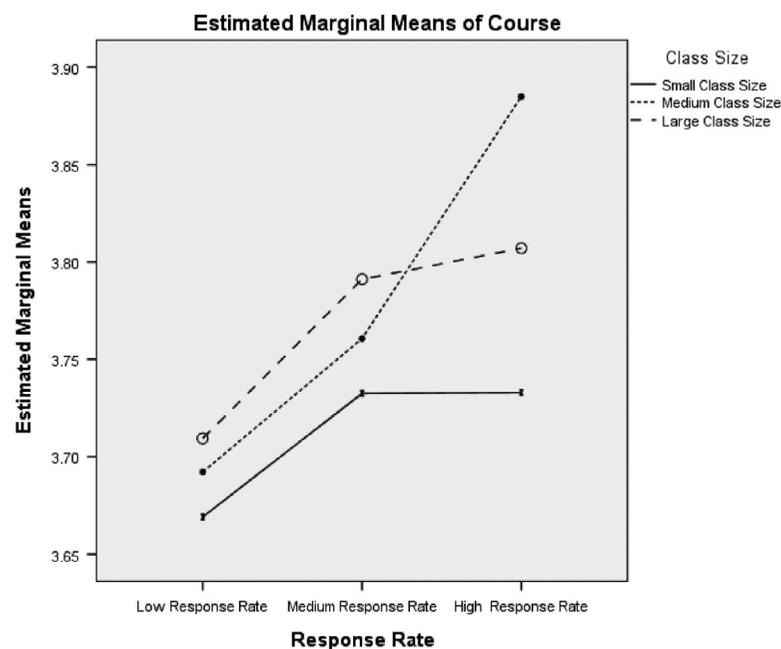
From Table 8, it is inferred that all the observed mean difference in the students evaluation of courses between three different class sizes are statistically significant and the highest mean difference is noted between medium and small class sizes (i.e. 0.068).

Irrespective of the response rate (RR), the mean rating of the students' evaluation of courses are found to be less if the class size is small, however, it improves correspondingly in a steady pace with an increase in the size of the class. Also, rating of students' evaluation of courses increases when the response rate is moving from "low" to "medium" and then it gets relatively stable when moving to "high" response rate. This suggests that when the class size is small, the response rate needs to be at least at the medium level to have stable students' evaluation of courses. Moreover, we observe that when the class size is medium, rating of students' evaluation of courses increases as the response rate increases (Figure 2).

**Table 8. Pairwise comparisons showing the mean difference in the students' evaluation of courses between different class sizes**

(I) Class size	(J) Class size	Mean difference (I-J)	Std. error	Sig.
Small class size	Medium class size	-.068	.007	.000
	Large class size	-.058	.007	.000
Medium class size	Small class size	.068	.007	.000
	Large class size	.010	.007	.149
Large class size	Small class size	.058	.007	.000
	Medium class size	-.010	.007	.149

**Figure 2. The interaction between response rate and class size on rating of students' evaluation of courses.**



#### 4. Discussion

This study explored the effect of interaction between response rate and class size on ratings of students' evaluations of instructors and courses. The frequent use of students' evaluation of courses is largely due to the easiness of collecting the data, presenting, and interpreting the results (Penny, 2003). The interpretation of this evaluation is more complicated than it looks, and it entails a risk of inappropriate use by both teachers and administrators for both formative and summative purposes (Franklin, 2001). This study indicated that if the students' evaluation of courses and instructors are conducted using a small class size and low response rate, it might lead to misinterpretation of findings. Previous study also suggested that inadequate results of students' evaluation of teaching could not be used for formative and faculty decision purposes (Conle, 1999; Franklin, 2001; Galbraith, Merrill, & Kline, 2012).

Precisely, an unstable students' evaluation of courses and instructors might occur when the response rate is low and the class size is small. So it is recommended that at least medium response rate is required if the class size is small in order to get stable ratings of students' evaluation of both instructors and courses. Further, we observe that when the class size is medium, the ratings of instructors and courses increase as the response rate increases. From the statistical point of view, poor response rate among the students in these evaluation surveys would make it difficult to arrive at generalisations and utmost caution need to be taken while interpreting the results of these evaluations (Al-Kuwaiti & Subbarayalu, 2015). Therefore, the HEIs need to be aware of these interpretations while using the students' evaluation surveys as a measure to take quality improvement and accountability purposes (Nulty, 2008).

Moreover, the ways in which administrators engage in students' evaluation of teaching effectiveness could be considered one of the greatest threats to the purposes of these evaluations (Penny, 2003). Many users are not sufficiently trained to handle these data, and they may even be unaware of their own ignorance about the collection and interpretation of these evaluations (Spooren et al., 2013). Therefore, misuse of these evaluations might have consequences for both the improvement of teaching and the career development (Boysen, Kelly, Raesly, & Casner, 2013). At the same time, a response rate should be considered large enough for the survey data to provide adequate evidence for accountability and the improvement purposes in order to maximize the benefits from these evaluations (Nulty, 2008). Also, it is very important to consider the types of students participated in the evaluation to ascertain whether the perceptions of participants differ from those of non-participants, particularly when the response rate is low and class size is small.

Besides class size and response rate, there are other factors that might interfere with the students evaluation of teaching. These factors are grouped under three categories viz. Students, faculty, and course factors. The students' factors include gender (Campbell & Bozeman, 2007); cultural background of the students (Capa-Aydin, 2014); Domain-specific vocational interests (Chen & Watkins, 2010) and; psychosocial dynamics such as instructors' attractiveness (Freng & Webber, 2009); Faculty factors include gender (Smith, 2009) and teachers characteristics (Chen & Watkins, 2010; Clayson, 2009; Clayson & Sheffet, 2006). The Course factors include grades or expected grades (Al-Kuwaiti & Maruthamuthu, 2014; Campbell & Bozeman, 2007; Dommeyer et al., 2004); course level (Galbraith et al., 2012); and course difficulty (Ginns et al., 2007). The present study also contributes to this discussion by

**Table 9. Suggested response rate for different class sizes while conducting the students' evaluation of instructors and courses**

Class size	Response rate		
	Low	Medium	High
Small		☒	☑
Medium			☑☒
Large		☑☒	

☑: Recommended response rate and class size required to evaluate instructors.

☒: Recommended response rate and class size required to evaluate courses.

investigating the interaction between response rate and class size (total number of students on the course) and its effects on students' evaluation of instructors and the courses offered at HEIs in Saudi Arabia. This study adds value to the literature by providing an appropriate response rate required for different class sizes while conducting students' evaluation of courses and instructors (Table 9).

## 5. Conclusions

We conclude that the interaction between class size and response rate has a significant effect on the students' evaluation of instructors and courses in higher education. Specifically, if the class size is small, a high response rate is required for evaluation of instructors and at least medium response rate is required for evaluation of courses. The results of the study will help the policy planners of HEIs to understand the response rate required for different class sizes while planning students' evaluation surveys of courses and instructors.

### Funding

The authors received no direct funding for this research.

### Author details

Ahmed Al Kuwaiti<sup>1</sup>

E-mails: [akuwaiti@uod.edu.sa](mailto:akuwaiti@uod.edu.sa), [qaa@uod.edu.sa](mailto:qaa@uod.edu.sa)

Mahmoud AlQuraan<sup>2</sup>

E-mail: [mahmoud.q@yu.edu.jo](mailto:mahmoud.q@yu.edu.jo)

Arun Vijay Subbarayalu<sup>1</sup>

E-mail: [ausubbarayalu@uod.edu.sa](mailto:ausubbarayalu@uod.edu.sa)

<sup>1</sup> Deanship of Quality and Academic Accreditation, University of Dammam, P.O. Box 40065, Al-Khobar 31952, Saudi Arabi.

<sup>2</sup> Department of Counseling and Educational Psychology, College of Education, Yarmouk University, P.O. Box 566, Irbid 21163, Jordan.

### Citation information

Cite this article as: Understanding the effect of response rate and class size interaction on students evaluation of teaching in a higher education, Ahmed Al Kuwaiti, Mahmoud AlQuraan & Arun Vijay Subbarayalu, *Cogent Education* (2016), 3: 1204082.

### References

- Addison, W. E., Best, J., & Warrington, J. D. (2006). Students' perceptions of course difficulty and their ratings of the instructor. *College Student Journal*, 40, 409–416.
- Al-Kuwaiti, A. (2014). Students evaluating teaching effectiveness process in Saudi Arabian medical colleges: A comparative study of students' and faculty members perception. *Saudi Journal of Medicine and Medical Sciences*, 2, 165–172.
- Al-Kuwaiti, A., & Maruthamuthu, T. (2014). Factors influencing student's overall satisfaction in course evaluation surveys: An exploratory study. *International Journal of Education and Research*, 2, 661–674.
- Al-Kuwaiti, A., & Subbarayalu, A. V. (2015). Appraisal of students experience survey (SES) as a measure to manage the quality of Higher Education in the Kingdom of Saudi Arabia: An institutional study using six sigma model. *Educational Studies*, 41. doi:10.1080/03055698.2015.1043977
- Al Rubaish, A., Wosornu, L., & Dwivedi, S. N. (2011). Using deductions from assessment studies towards furtherance of the academic program: An empirical appraisal of institutional student course evaluation. *iBusiness*, 3, 220–228.
- Al Rubaish, A., Wosornu, L., & Dwivedi, S. N. (2012). Appraisal of using global student rating items in quality management of Higher Education in Saudi Arabian University. *iBusiness*, 4(1), 1–9.
- Archibong, I. A., & Nja, M. E. (2011). Towards improved teaching effectiveness in Nigerian public universities: Instrument design and validation. *Higher Education Studies*, 1, 78–91.
- Badri, M. A., Abdulla, M., Kamali, M. A., & Dodeen, H. (2006). Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the UAE. *International Journal of Educational Management*, 20, 43–59.
- Beran, T., Violato, C., & Kline, D. (2007). What's the "use" of student ratings of instruction for administrators? One university's experience. *Canadian Journal of Higher Education*, 37, 27–43.
- Berk, R. A. (2005). Survey of 12 strategies to measure teaching effectiveness. *International Journal of Teaching and Learning in Higher Education*, 17, 48–62.
- Blackmore, J. (2009). Academic pedagogies, quality logics and performative universities: Evaluating teaching and what students want. *Studies in Higher Education*, 34, 857–872. <http://dx.doi.org/10.1080/03075070902898664>
- Boysen, G. A., Kelly, T. J., Raesly, H. N., & Casner, R. W. (2013). The (mis) interpretation of teaching evaluations by college faculty and administrators. *Assessment & Evaluation in Higher Education*, 1–16.
- Burden, P. (2008). Does the use of end of semester evaluation forms represent teachers' views of teaching in a tertiary education context in Japan? *Teaching and Teacher Education*, 24, 1463–1475. <http://dx.doi.org/10.1016/j.tate.2007.11.012>
- Campbell, J. P., & Bozeman, W. C. (2007). The value of student ratings: Perceptions of students, teachers, and administrators. *Community College Journal of Research and Practice*, 32, 13–24. <http://dx.doi.org/10.1080/10668920600864137>
- Capa-Aydin, Y. (2014). Student evaluation of instruction: Comparison between in-class and online methods. *Assessment & Evaluation in Higher Education*, 1–15.
- Chen, Y., & Hoshower, L. B. (2003). Students evaluation of teaching effectiveness: An assessment of students' perceptions and motivation. *Assessment & Evaluation in Higher Education*, 28, 71–88.
- Chen, G. H., & Watkins, D. (2010). Stability and correlates of student evaluations of teaching at a Chinese university. *Assessment & Evaluation in Higher Education*, 35, 675–685.
- Clayson, D. E. (2009). Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature. *Journal of Marketing Education*, 31, 16–30.
- Clayson, D. E., & Sheffet, M. J. (2006). Personality and the student evaluation of teaching. *Journal of Marketing Education*, 28, 149–160. <http://dx.doi.org/10.1177/0273475306288402>
- Coffey, M., & Gibbs, G. (2001). The evaluation of the Student Evaluation of Educational Quality Questionnaire (SEEQ) in UK Higher Education. *Assessment & Evaluation in Higher Education*, 26, 89–93.

- Conle, C. (1999). Moments of interpretation in the perception and evaluation of teaching. *Teaching and Teacher Education*, 15, 801–814.  
[http://dx.doi.org/10.1016/S0742-051X\(99\)00026-8](http://dx.doi.org/10.1016/S0742-051X(99)00026-8)
- 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 response rates and evaluations. *Assessment & Evaluation in Higher Education*, 29, 611–623.
- Franklin, J. (2001). Interpreting the numbers: Using a narrative to help others read student evaluations of your teaching accurately. *New Directions for Teaching and Learning*, 87, 85–100.
- Freng, S., & Webber, D. (2009). Turning up the heat on online teaching evaluations: Does “hotness” matter? *Teaching of Psychology*, 36, 189–193.  
<http://dx.doi.org/10.1080/00986280902959739>
- Galbraith, C. S., Merrill, G., & Kline, D. (2012). Are student evaluations of teaching effectiveness valid for measuring student learning outcomes in business related classes? A neural network and bayesian analyses. *Research in Higher Education*, 53, 353–374.  
<http://dx.doi.org/10.1007/s11162-011-9229-0>
- Ginns, P., Prosser, M., & Barrie, S. (2007). Students’ perceptions of teaching quality in Higher Education: The perspective of currently enrolled students. *Studies in Higher Education*, 32, 603–615.  
<http://dx.doi.org/10.1080/03075070701573773>
- Gleason, J. (2012). Using technology-assisted instruction and assessment to reduce the effect of class size on student outcomes in undergraduate mathematics courses. *College Teaching*, 60, 87–94.  
<http://dx.doi.org/10.1080/87567555.2011.637249>
- Johnson, R. (2000). The authority of the student evaluation questionnaire. *Teaching in Higher Education*, 5, 419–434.  
<http://dx.doi.org/10.1080/713699176>
- Kogan, J. R., & Shea, J. A. (2007). Course evaluation in medical education. *Teaching and Teacher Education*, 23, 251–264.  
<http://dx.doi.org/10.1016/j.tate.2006.12.020>
- Kogan, L. R., Schoenfeld-Tacher, R., & Hellyer, P. W. (2010). Student evaluations of teaching: Perceptions of faculty based on gender, position, and rank. *Teaching in Higher Education*, 15, 623–636.  
<http://dx.doi.org/10.1080/13562517.2010.491911>
- Koh, H. C., & Tan, T. M. (1997). Empirical investigation of the factors affecting SET results. *International Journal of Educational Management*, 11, 170–178.
- Konting, M. M., Kamaruddin, N., & Man, N. A. (2009). Quality assurance in Higher Education institutions: Exit survey among University Putra Malaysia graduating students. *International Educational Studies*, 2, 25–31.
- Kuzmanovic, M., Savic, G., Andric Gusavac, B., Makajic-Nikolic, D., & Panic, B. (2013). A Conjoint-based approach to student evaluations of teaching performance. *Expert Systems with Applications*, 40, 4083–4089.  
<http://dx.doi.org/10.1016/j.eswa.2013.01.039>
- Morley, D. (2014). Assessing the reliability of student evaluations of teaching: Choosing the right coefficient. *Assessment & Evaluation in Higher Education*, 39, 127–139.
- Neumann, R. (2000). Communicating student evaluation of teaching results: Rating Interpretation Guides (RIGs). *Assessment and Evaluation in Higher Education*, 25, 121–134.  
<http://dx.doi.org/10.1080/02602930050031289>
- Nikolaidis, Y., & Dimitriadis, S. G. (2014). On the student evaluation of university courses and faculty members’ teaching performance. *European Journal of Operational Research*, 238, 199–207.  
<http://dx.doi.org/10.1016/j.ejor.2014.03.018>
- Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: What can be done? *Assessment & Evaluation in Higher Education*, 33, 301–314.
- Ory, J. C., & Ryan, K. (2001). How do student ratings measure up to a new validity framework? *New Directions for Institutional Research*, 109, 27–44.  
[http://dx.doi.org/10.1002/\(ISSN\)1536-075X](http://dx.doi.org/10.1002/(ISSN)1536-075X)
- Penny, A. R. (2003). Changing the agenda for research into students’ views about university teaching: Four shortcomings of SRT research. *Teaching in Higher Education*, 8, 399–411.  
<http://dx.doi.org/10.1080/13562510309396>
- Pezzella, F. S., Paladino, A., Zoller, C., & Mandery, E. (2014). The efficacy of student learning in large-sized criminal justice preparatory classes. *Journal of Criminal Justice Education*, 25, 106–130. <http://dx.doi.org/10.1080/10511253.2013.875214>
- Ramsden, P. (1991). A performance indicator of teaching quality in Higher Education: The course experience questionnaire. *Studies in Higher Education*, 16, 129–150.  
<http://dx.doi.org/10.1080/03075079112331382944>
- Smith, B. P. (2009). Student ratings of teaching effectiveness for faculty groups based on groups based on race and gender. *Education*, 129, 615–624.
- 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, 598–642.  
<http://dx.doi.org/10.3102/0034654313496870>
- Sproule, R. (2000). Student evaluation of teaching: A methodological critique of conventional practices. *Education Policy Analysis Archives*, 8, 125–142.  
<http://dx.doi.org/10.14507/epaa.v8n50.2000>
- Vijay, S. A. (2014). Appraisal of student rating as a measure to manage the quality of Higher Education in India: An institutional study using six sigma model approach. *International Journal of Quality Research*, 7, 493–508.
- Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education*, 23, 191–212.
- Zabaleta, F. (2007). The use and misuse of student evaluations of teaching. *Teaching in Higher Education*, 12, 55–76.  
<http://dx.doi.org/10.1080/13562510601102131>



© 2016 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

