

Decision-making by Heads of Academic Department using Student Evaluation of Instruction (SEI)

Mohammed Saleh Alkathiri

Imam Abdulrahman bin Faisal University, Dammam, Saudi Arabia
<https://orcid.org/0000-0002-4079-7347>

Abstract. This study investigates the ways in which heads of academic department use student evaluation of instruction (SEI) to make decisions about individual faculty members and/or whole academic departments. The study utilized a convenience sample of 57 heads of department, who completed an online questionnaire with two main constructs, which were assessed at the interval level of measurement. The results of the study revealed significant differences between heads of department who tend to trust SEI results compared to those who tend not to trust SEI results. The findings suggest there is a significant association between how heads of department perceive SEI and how they use it to make decisions about individual faculty members and their academic departments. In addition, analysis of the respondents as per two groups, according to their attitudes of trust or distrust toward SEI, showed that disparities within these groups were greater with respect to issues or decisions that affect individuals as opposed to whole departments. Therefore, the study concludes that decisions should not be made based solely on the results of SEI; rather, multiple sources of evaluation should be utilized to make proper decisions. The author strongly recommends that academic leaders should use SEI across multiple years or courses in order to obtain more reliable information. Future research may include qualitative studies on the topic and discipline-specific studies within certain academic departments or college clusters.

Keywords: faculty members; heads of academic department; higher education; student evaluation of instruction; teaching and learning

1. Introduction

The use of student evaluation of instruction (SEI) is, and has long been, one of the most common assessment practices in higher education. In the United States, SEI is the predominant form of faculty evaluation, and approximately 88% of all liberal arts colleges use SEI for summative decisions (Seldin, 1999). In 1991, the U.S. Department of Education reported that 97% of 40,582 heads of department who participated in a survey used SEI to assess teaching (Cashin, 2003). Today, many universities still use SEI to determine whether to grant faculty tenure,

promote them, raise their pay based on merit, and offer them opportunities for professional development (Kelly, Ponton & Rovai, 2007). In fact, at institutions where the emphasis is on teaching, SEI is an influential measure used in promotion decisions (Emery, Kramer & Tian, 2003). Furthermore, universities use SEI for various other purposes, such as establishing the credibility of the education they offer, planning strategy, and improving curricula (Massy & French, 2001; Scott & Hawke, 2003). Nonetheless, the use of SEI in higher education has been controversial. Whilst supporters of SEI view it as a valid and reliable tool that can be used to facilitate decisions, those who oppose SEI claim that it is biased because of many factors that influence its results.

SEI has been researched more often than other topics in higher education for decades (Theall & Franklin, 2001). Whereas many studies have supported the use of SEI as a valid indicator of quality instruction, several studies have questioned its validity and reliability for faculty evaluation. This paper provides a literature review of support for and opposition to SEI. Moreover, it identifies a gap in the literature around the question of how heads of department use SEI results to make decisions about faculty members and academic departments in the context of Saudi higher education. It then presents original research devised to fill that gap. The results inform the reader of the present practical reality of SEI and show how practitioners and policy makers use SEI results in their decisions. Finally, this paper provides conclusions and recommendations for academic leaders on the use of SEI in higher education. These will enable well-informed evaluations of individual faculty members and overall academic departments, and thus ultimately facilitate better decision-making in future.

2. Literature Review

2.1 History of student evaluation of instruction

The primary purpose of using SEI is to improve the quality of teaching and learning (Delvaux et al., 2013). Since ancient times, students have had a voice on how they are taught. For example, at the time of Socrates and in the medieval period (Tucker, 2015), students expressed their opinions about their teachers. However, the use of SEI, as it is currently known and applied, to evaluate faculty teaching started in the early 1900s (Algozzine et al., 2004). Wachtel (1998) provided a brief review of SEI, indicating that “the first teacher rating scale was published in 1915” (p. 191). In the 1920s, several universities in the United States started student evaluation procedures (Wachtel, 1998). Today, SEI is used in universities worldwide. Taking into consideration the fact that faculty members today perform multiple responsibilities (Alkathiri, 2018), the purpose of SEI has been expanded by universities, as well as by quality assurance bodies. It is now used, for example, to allocate performance funding, to gather evidence to decide which faculty members to promote, and to select the winners of teaching awards (Arthur, 2009; Hendry & Dean, 2002; Massy & French, 2001; Scott & Hawke, 2003; Shah & Nair, 2012; Tucker, 2014).

According to Centra (1993), research on SEI went through four distinct periods. First, the period from 1927 to 1960 was known by the pioneer work of Remmers and his colleagues at Purdue University. Second, in the 1960s, the use of SEI was

voluntary in almost all universities. Third, the 1970s was the “golden age of research on student evaluations” (Centra, 1993, p. 49). During this period, new research on SEI evolved, including studies that showed evidence of the validity of SEI and advocated the use of SEI for formative and summative evaluations. Lastly, the fourth period started in the early 1980s, when research on SEI continued to expand, providing further illustration of research findings, including studies of the meta-analysis type.

According to Theall and Franklin (2001), more studies have been conducted on SEI than any other topic in higher education. Many studies have been done in the United States, Australia, and Europe. In contrast, Saudi Arabia, and the Arab world more generally, has been relatively understudied. Research on SEI has covered various subtopics such as the validity, reliability, and usefulness of SEI; the dimensions of effective teaching to be evaluated; the bias in student and instructor responses; and the identification of teaching excellence (Tucker, 2014). Furthermore, many reviews of the literature and meta-analysis studies on SEI have been conducted (Alderman, Towers & Bannah, 2012; Perry & Smart, 2007; Richardson, 2005). Nonetheless, research seldom includes discussions of the use of SEI by heads of department to make decisions that may affect individual faculty members or academic departments. The current paper is intended to contribute to the field of higher education by rectifying that omission, presenting and analyzing new evidence from Saudi Arabia where the study took place.

2.2 Support for student evaluation of instruction

Many researchers have claimed that SEI is an important indicator of where quality is improving and where it needs to be improved in teaching and learning, and in student satisfaction (Alkathiri, 2020; Zineldin, Akdag & Vasicheva, 2011). In addition, for decades, scholars have suggested that SEI can be considered a valid indicator of effective instruction. For example, as indicated by Liu (2012), SEI can predict ratings gathered from other sources, such as former students and colleagues. In addition, McKeachie (1997) claimed that SEI is a source of evaluation of teaching effectiveness that is more valid than any other. According to Liu (2012), early studies on SEI acknowledged its importance in teaching and learning because of the way in which SEI can actually reflect the quality of teaching based on student perceptions. Furthermore, literature (see Liu, 2012) has urged the use of SEI since students are able to furnish information on (1) learning goals, (2) student-instructor rapport, (3) teaching methods, (4) student-instructor communication, and (5) consumer data.

Research comparing SEI in distance education and face-to-face courses has concluded that there is little difference between the two modes of teaching in terms of the ratings of whole courses, and the quality of their instruction (Kelly et al., 2007; McGhee & Lowell, 2003; Waschull, 2001). Supporters of the use of SEI claim that if students are trained in using SEI, “evaluative judgements [given] on a regular basis have strong positive impact on the improvement of [faculty’s] instructional skills” (Spooren, Mortelmans & Denekens, 2007, p. 667). With regard to factors that might be expected to affect SEI, Aleamoni (1999) indicated that there is no relationship between SEI and class size, gender of student, time of day when

a course is offered, level of course, or rank of instructor. In addition, many reviews of SEI have concluded that gender roles have no effect on it, or where such effects exist, they are not significant (Liu, 2012; Radmacher & Martin, 2001).

2.3 Concerns about student evaluation of instruction

Although many studies support the use of SEI in higher education, others express opposition to it. For example, some studies have revealed a gender bias against female faculty members in SEI, with students evaluating male and female faculty members based on different dimensions (Basow, 1995; Chamberlin & Hickey, 2001; Liu, 2012). Basow (1995) analyzed 2,000 SEIs collected from undergraduate students over four years and found that male faculty members scored much better than female faculty members on most questions most of the time, aside from one year when the women scored better on two criteria (i.e., sensitivity and student comfort). Another issue of bias is that female students consistently score their female professors higher than their male peers do (Centra & Gaubatz, 2000). Overall, it is evident that the gender of faculty member and students has an effect on SEI.

Another concern about SEI is the impact of teaching mode, in that, sometimes, faculty members receive disproportionately lower ratings in face-to-face courses compared to faculty members who teach online classes (Carle, 2009). Students studying online show more diverse opinions than students in face-to-face courses when scoring the delivery mode for effectiveness (Kelly et al., 2007; Liu, 2012; McGhee & Lowell, 2003; Rovai et al., 2006). Furthermore, class size has been found to have an influence on SEI. For example, many studies examining the impact of class size on SEI have found that higher SEI scores correlate with smaller classes (Badri et al., 2006; Liaw & Goh, 2003; Liu, 2012).

Another problem with SEI is the impact of evaluation instruments on results. According to Landrum and Braitman (2008), SEI scores decrease significantly when the number of points on an evaluation scale are changed from 10 to 5. On the 10-point scale, students would use a range of values that is larger than on the 5-point scale. Other studies have examined subtle factors that impact SEI. For example, elective courses score better than compulsory ones (Marsh & Roche, 1997); SEI at the end of a semester can be significantly predicted by students' first impressions of the instructor (Buchert et al., 2008); undergraduate students give lower ratings than graduate students do (Marsh, 2007; Whitworth, Price & Randall, 2002); the faculty member's rank and experience influence SEI (Rovai et al., 2006); and faculty-member characteristics such as enthusiasm and humor can positively impact SEI (Obenchain, Abernathy & Wiest, 2001).

According to Centra and Gaubatz (2000), SEI can be biased because characteristics of students and instructors that are irrelevant to teaching may potentially affect ratings. For example, a class right at the start of the day might receive a worse score than the same class at a later, less awkward time. Concerns around the reliability of SEI need to be taken seriously, considering the influences of extraneous factors on SEI. When making decisions regarding faculty members, academic leaders should take care when using SEI results from undergraduate

courses or small classes. SEI of a small sample (i.e., in small classes) might not be accurate nor reflect the actual quality of faculty teaching. In order to make accurate decisions, the types of courses and the reasons that students took the courses should be considered. Another suggestion for increasing the reliability of SEI and reducing distortions is to ask students to rate the extent to which they have attained their educational objectives (McKeachie, 1997; Zhao & Gallant, 2012).

3. Statement of Purpose and Research Question

Conducting research using student-satisfaction data is a common practice in higher education (Alkathiri, 2020). Moreover, SEI is a crucial aid to decision-making there. SEI is used in colleges and universities for various purposes. These include providing formative feedback to faculty for instructional improvement; measuring teaching effectiveness in order to make administrative decisions on career advancement; helping students choose classes and instructors; and for research on teaching (Zhao & Gallant, 2012). Furthermore, according to Algozzine et al. (2004), a major reason for universities to use SEI is to make decisions on salary. That said, when making decisions based on SEI, institutional administrators need to be aware of the various findings of the ongoing research concerning its validity and reliability. The present study was devised to investigate the views of heads of department concerning the use of SEI to facilitate decisions. The author investigated the attitudes of heads of department and the ways in which these affect their use of SEI. The primary research question was: Does the difference in views of heads of department, at a public university in the Eastern Province of Saudi Arabia, have a significant effect on their use of SEI to make decisions about individual faculty members and about academic departments? The author investigated the use of SEI results by heads of department to make decisions about: (1) individual faculty members (including on promotion and awarding of tenure, effectiveness of teaching, professional development needs, and contribution to student learning experience); and (2) academic departments (including on effectiveness of teaching, professional development needs, and contribution to student cohort learning experience).

4. Methodology

The present study focused on heads of department at a public university in the Eastern Province of Saudi Arabia. Its purpose was to determine how significantly the differences between the heads of department in their views on SEI affect their use of it to make decisions about individual faculty members and academic departments. This study's main hypothesis was that the tendency of heads of department to trust or not to trust the results of SEI makes no difference on their decision-making about individual faculty members or the academic department. The criterion variable was their overall tendency to trust the use of SEI to make decisions. The author utilized a quantitative research design and a convenience sample of heads of department, who were asked to take an online questionnaire to provide data. Convenience sampling involves the sample being drawn from the population that is available to the researcher (Taherdoost, 2016). Analyzing and evaluating the data on two constructs, the author developed an evaluation scale to assess the respondents' responses regarding their experiences using SEI

to make decisions about individual faculty members (i.e., construct A) and about their academic department (i.e., construct B).

4.1 Respondents

An online survey was distributed to 112 heads of department from 20 colleges in four clusters (i.e., health, engineering, sciences and management, and arts and education). Fifty-seven heads of department completed the survey, which equates to a 50.89% response rate. Just under six out of every ten respondents were women. Table 1 displays the counts and percentages of respondents by category.

Table 1: Respondent demographic data

Categories of participating heads of department		Overall sample count (n = 57)	% M
Sex	Male	24	42.1
	Female	33	57.9
College cluster	Health	8	14.0
	Engineering	19	33.3
	Sciences and management	15	26.3
	Arts and education	15	26.3

4.2 Procedure

The researcher sent out an email to 112 heads of department from all clusters, asking them to take an online questionnaire. Respondents completed the questionnaire voluntarily. No compensation was offered for completing the questionnaire. Respondents were shown a consent form prior to taking the questionnaire. Respondents' completion and submission of the questionnaire were used to indicate consent.

4.3 Instrument

The author developed an online questionnaire of eight question items (see Appendix 1). The questionnaire asked respondents to indicate their attitude toward each item using a 6-point Likert-type scale with three levels of agreement and three of disagreement. The first item, "Overall, I tend to trust the SEI results to make decisions in my job," was intended to identify the level of trust in general terms. The seven other question items were subscales of two constructs, A and B (see Table 2). The two main constructs were assessed at the interval level of measurement.

Table 2: Percentages of some form of agreement, mean scores, and standard deviations of scores for questions and constructs (strongly disagree = 1, strongly agree = 6)

Item no.	Question	Some form of agreement (%)	M	SD
1	Overall, I tend to trust the SEI results to make decisions in my job as a head of department	40.3	3.2	1.0
2	As a head of department, I would use the SEI results to make decisions about the individual faculty members	38.2	3.2	0.72
2.1	Individual faculty members' promotion/awarding of tenure	3.5	1.9	0.69
2.2	Individual faculty members' effectiveness of teaching	24.6	2.9	1.2
2.3	Individual faculty members' professional development needs	64.9	3.8	0.93
2.4	Individual student learning experience	59.7	4.0	0.93
3 (B)	As a head of department, I would use the SEI results to make decisions about the academic department	95.9	5.0	0.66
3.1	Academic department's effectiveness of teaching	87.7	4.6	0.96
3.2	Academic department's professional development needs	100	5.2	0.57
3.3	Student cohort learning experience	100	5.2	0.63

Construct A, concerning the use of SEI results to make decisions about individual faculty members, is measured by questions 2.1, 2.2, 2.3, and 2.4. Construct B, concerning the use of SEI results to make decisions about an academic department, is measured by questions 3.1, 3.2, and 3.3. Using the online questionnaire, two kinds of scores were calculated for each construct: subscale scores and an overall score. The subscale score for each question item from 2.1 onwards is the mean value calculated from all of its responses. The higher the subscale score for an item, the more likely the participating heads of department would be to use SEI to make a decision about the matter in question. For instance, the high M value on question 3.2 implies that the participating academic leaders are more likely to use SEI to make decisions about an academic department's professional development needs. Finally, an average of the subscale scores for each construct gives its overall tendency score.

4.4 Analysis

The researcher used an independent t-test to assess if there was an association between overall trust in results of SEI by heads of department (i.e., "Overall, I tend to trust the SEI results to make decisions in my job as a head of department") and their willingness to use SEI to make decisions about individual faculty members

and overall academic departments. As mentioned above, to obtain the overall tendency score for each construct (A or B), the relevant subscale scores were averaged. Table 2 displays percentages of some form of agreement, mean scores, and standard deviations of scores for each of the items and constructs. The Cronbach alpha coefficient of internal consistency was computed to report the reliability and correlations for each of the constructs, as presented in Table 3.

Table 3: Correlation of constructs and measures of internal consistency

Construct	Question numbers	Subscale construct	A	α
A	2.1, 2.2, 2.3, 2.4	Decisions about individual faculty members		.76
B	3.1, 3.2, 3.3	Decisions about overall academic department	.78*	.87

* $p < .01$.

5. Results

It is remarkable that all of the responses to question items 3.2 and 3.3 indicated some form of agreement. Both concern construct B: using SEI results to make decisions about the academic department. The least agreement was with question item 2.1: using SEI to make decisions about an individual faculty member's promotion and awarding of tenure.

Overall, the percentage of some form of agreement could be considered low for the questions within construct A (concerning individual faculty members), in contrast to the percentage of some form of agreement for the questions within construct B (concerning the academic department). The individual items within the constructs were averaged. Table 3 shows that the results of the survey have high reliability, as well as significant correlations between the constructs.

Based on the respondents' answers to the first question, two groups of heads of department were identified: one group with a tendency to trust the use of SEI to make decisions (answering 4, 5, or 6), and another group with a tendency not to (answering 1, 2, or 3). The t-test results showed statistically significant differences between construct means for the two groups. Therefore, the study's main hypothesis is rejected because the tendency of the respondents to trust or not to trust the results of SEI makes a significant difference on their decision-making about the individual faculty members or the academic department.

For construct A, respondents in the "trusting" first group had a mean of 3.85, whereas respondents in the "distrusting" second group had a mean of 2.68. The difference was statistically significant ($t(55) = 9.749, p > .01$). For construct B, respondents in the first group had a mean score of 5.58, whereas the second group's mean score was 4.59. The difference was statistically significant ($t(55) = 8.315, p > .01$). Furthermore, Cohen's effect size was computed, and alpha level was reported in Table 3.

6. Discussion

According to Muammar and Alkathiri (2021), “higher education institutions are faced with a constantly evolving set of aims: to meet the needs of students while responding to societal demands and stakeholders’ expectations in a context of continually changing expectations, roles, and outcomes” (p. 1). Despite the challenges that heads of department face in their positions, they have a key role in the success of academic departments as well as the achievement of major higher education objectives (Freeman Jr., Karkouti & Ward, 2020; Reznik & Sazykina, 2017). The present study should make educators in higher education aware of how much heads of department vary in terms of their tendency to trust the results of SEI for making informed decisions.

Although the position of head of department is fundamental in higher education institutions to implement forward changes, the position is unattractive to many faculty members (Freeman Jr. et al., 2020). The expectations of heads of department can be ambiguous, especially that they receive limited training prior to assuming the position (Freeman Jr. et al., 2020). According to Freeman Jr. et al. (2020), “[c]hairs must balance the dual responsibilities of managing faculty and student affairs who they support and evaluate as they implement the mandates from higher administration. Similarly, they shuttle between their managerial roles and faculty roles while balancing work-life demands” (p. 895). Figure 1 shows the difference among respondents regarding using student evaluation of instruction to make decisions about individual faculty members and the academic department on various purposes.

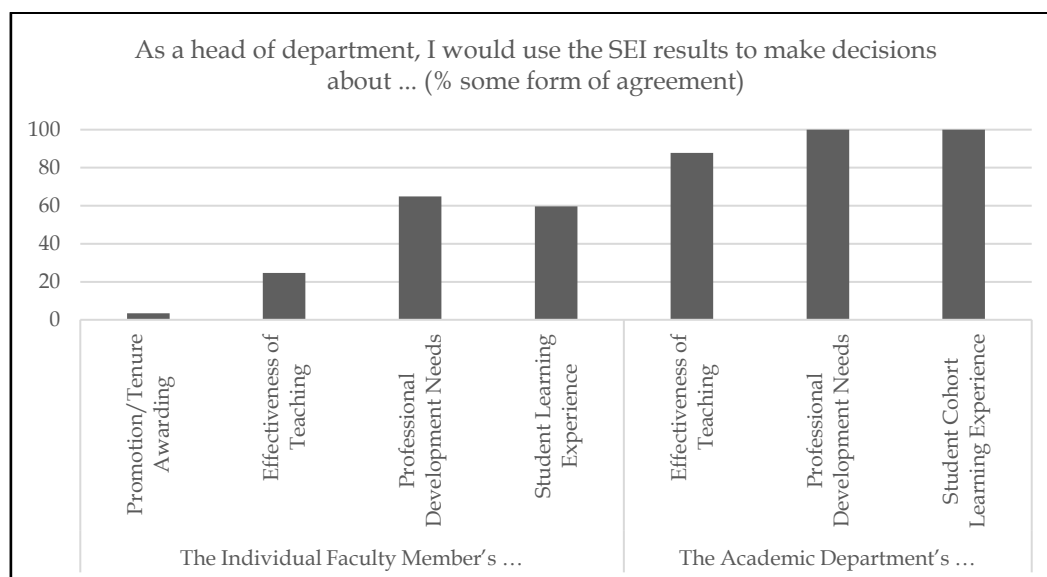


Figure 1: Form of agreement (%) by respondents regarding using student evaluation of instruction for various purposes

As seen in Figure 1, respondents tended to use SEI results to make decisions about individual faculty members’ professional development needs and about their contribution to the individual student learning experience. However, respondents were less likely to use SEI results to make decisions regarding the effectiveness of individual faculty members’ teaching or about promoting them or awarding them

tenure. According to Smith (2005), “heads of department are overloaded with work, ... large departments are difficult to manage and ... collegiality is the ‘preferred’ model of decision-making” (p. 463). On the other hand, respondents were more likely to use SEI results to make decisions about their whole academic departments on matters such as effectiveness of teaching, professional development needs, and the student cohort learning experience. Learning about the significant differences amongst heads of department concerning the use of SEI to make decisions will help academic leaders to address the issue. It is hoped that this will result in specific measures that facilitate a better understanding of these different views on SEI, and promote well-informed decisions in higher education.

Many studies have reported that, in general, SEI is a valid indicator of the quality of instruction (Marsh & Roche, 2000; Theall & Franklin, 2001). Many faculty members in higher education have exhibited reasonably positive attitudes toward the validity of SEI and its usefulness for improving instruction (Nasser & Fresko, 2002). However, others have expressed concerns about SEI and its uses (Nasser & Fresko, 2002) because of various factors that may cause biases (Badri et al., 2006; Kelly et al., 2007).

This study set out to investigate the views of heads of academic department concerning SEI, and the effect of their views on the ways in which they tend to use SEI when making decisions about individual faculty members and their overall academic department. The sample of heads of department who tended to trust SEI results showed more agreement with the use of SEI to make decisions, in comparison to those who tended not to trust SEI. The comparisons were statistically significant with respect to two constructs: using SEI results to make decisions about (A) individual faculty members and (B) overall academic department. Therefore, it can be concluded that whether or not heads of department trust SEI results has an effect on their use of such information when making decisions about faculty members under their supervision as well as about the academic department that they chair.

Although there were significant differences between the group who trusted SEI and the one who did not, the mean scores for each group on the use of SEI results to make decisions were lower where those decisions concerned individual faculty members as opposed to whole academic departments. Furthermore, the disparity between the two groups was biggest when it came to decisions that affect individuals. Therefore, it is clear that decisions cannot be made based solely on the results of SEI; rather, multiple sources of evaluation should be utilized. Moreover, when evaluating the effectiveness of a faculty member, the multi-dimensional nature of SEI should be considered. Further research may include qualitative studies on the topic to further our understanding of the use of SEI by heads of department. Additionally, discipline-specific studies within certain academic departments or college clusters are recommended.

7. Conclusion

Fifty-seven heads of department completed an online questionnaire with two main constructs: using SEI results to make decisions about (A) individual faculty

members and (B) overall academic department. The study aimed at investigating the ways in which heads of academic department use SEI to make decisions about individual faculty members and their academic department. The results revealed statistically significant differences between those heads of department who tended to trust the results of SEI and those who tended not to trust the results of SEI. The study concludes that there is a significant association between how heads of department perceive SEI and how they use it to make decisions about their academic department and individual faculty members.

In addition, the disparities within the groups of respondents, according to their attitudes of trust or distrust toward SEI, were greater with respect to issues or decisions that affect individuals as opposed to overall department. Therefore, decisions should not be made based solely on the results of SEI; rather, multiple sources of evaluation should be utilized to make proper decisions. Based on the findings of the current study, the author strongly suggests that academic leaders should use SEI across multiple years or courses in order to obtain more reliable information.

8. References

- Alderman, L., Towers, S., & Bannah, S. (2012). Student feedback systems in higher education: A focused literature review and environmental scan. *Quality in Higher Education*, 18(3), 261-280. <https://doi.org/10.1080/13538322.2012.730714>
- Aleamoni, L. M. (1999). Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*, 13, 152-66. <https://doi.org/10.1023/A:1008168421283>
- Algozzine, B., Gretes, J., Flowers, C., Howley, L., Beattie, J., Spooner, F., Mohanty, G., & Bray, M. (2004). Student evaluation of college teaching: A practice in search of principles. *College Teaching*, 52(4), 134-41. <https://doi.org/10.3200/CTCH.52.4.134-141>
- Alkathiri, M. S. (2018). Using art-based techniques in faculty training programmes. In T. Chemi & X. Du (Eds.), *Palgrave studies in business, arts and humanities: Arts-based methods and organizational learning* (pp. 265-290). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-63808-9_12
- Alkathiri, M. S. (2020). Teaching and learning experiences of college students in Saudi higher education in the 21st century. *The International Journal of Learning in Higher Education*, 27(1), 15-30. <https://doi.org/10.18848/2327-7955/CGP/v27i01/15-30>
- Arthur, L. (2009). From performativity to professionalism: Lecturers' responses to student feedback. *Teaching in Higher Education*, 14(4), 441-454. <https://doi.org/10.1080/13562510903050228>
- Badri, M., Abdulla, M., Kamali, M., & 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. <https://doi.org/10.1108/09513540610639585>
- Basow, S. (1995). Student evaluations of college professors: When gender matters. *Journal of Educational Psychology*, 87, 656-665. <https://doi.org/10.1037/0022-0663.87.4.656>
- Buchert, S., Laws, E. L., Apperson, J. M., & Bregman, N. J. (2008). First impressions and professor reputation: Influence on student evaluations of instruction. *Social Psychology of Education*, 11(4), 397-408. <https://doi.org/10.1007/s11218-008-9055-1>

- Carle, A. C. (2009). Evaluating college students' evaluations of a professor's teaching effectiveness across time and instruction mode (online vs. face-to-face) using a multilevel growth modeling approach. *Computers & Education*, 53(2), 429-435. <https://doi.org/10.1016/j.compedu.2009.03.001>
- Cashin, W. E. (2003). Evaluating college and university teaching: Reflections of a practitioner. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 531-593). Kluwer Academic. https://doi.org/10.1007/978-94-010-0137-3_10
- Centra, J. A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. Jossey-Bass. <https://eric.ed.gov/?id=ED363233>
- Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *Journal of Higher Education*, 70(1), 17-33. <https://doi.org/10.1080/00221546.2000.11780814>
- Chamberlin, M. S., & Hickey, J. S. (2001). Student evaluations of faculty performance: The role of gender expectations in differential evaluations. *Educational Research Quarterly*, 25(2), 3-14. <https://search.proquest.com/docview/216185804?pq-origsite=gscholar&fromopenview=true>
- Delvaux, E., Vanhoof, J., Tuytens, M., Vekeman, E., Devos, G., & Van Petegem, P. (2013). How may teacher evaluation have an impact on professional development? A multilevel analysis. *Teaching and Teacher Education Journal*, 36, 1-11. <https://doi.org/10.1016/j.tate.2013.06.011>
- Emery, C. R., Kramer, T. R., & Tian, R. G. (2003). Return to academic standards: A critique of students' evaluations of teaching effectiveness. *Quality Assurance in Education: An International Perspective*, 11(1), 37-47. <https://doi.org/10.1108/09684880310462074>
- Freeman Jr., S., Karkouti, I. M., & Ward, K. (2020). Thriving in the midst of liminality: Perspectives from department chairs in the USA. *Higher Education*, 80, 895-911. <https://doi.org/10.1007/s10734-020-00521-6>
- Hendry, G. D., & Dean, S. J. (2002). Accountability, evaluation of teaching and expertise in higher education. *The International Journal for Academic Development*, 26(4), 327-414. <https://doi.org/10.1080/13601440210156493>
- Kelly, H. F., Ponton, M. K., & Rovai, A. P. (2007). A comparison of student evaluations of teaching between online and face-to-face courses. *Internet and Higher Education*, 10, 89-101. <https://doi.org/10.1016/j.iheduc.2007.02.001>
- Landrum, R. E., & Braitman, K. A. (2008). The effect of decreasing response options on students' evaluation of instruction. *College Teaching*, 56, 215-217. <https://doi.org/10.3200/CTCH.56.4.215-218>
- Liaw, S., & Goh, K. (2003). Evidence and control of biases in student evaluations of teaching. *The International Journal of Educational Management*, 17(1), 37-43. <https://doi.org/10.1108/09513540310456383>
- Liu, O. L. (2012). Student evaluation of instruction: In the new paradigm of distance education. *Research in Higher Education*, 53, 471-486. <https://doi.org/10.1007/s11162-011-9236-1>
- Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education* (pp. 319-383). Springer. https://doi.org/10.1007/1-4020-5742-3_9
- Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, 52(11), 1187-1197. <https://doi.org/10.1037/0003-066X.52.11.1187>
- Marsh, H. W., & Roche, L. A. (2000). Effects of grading lenience and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent

- bystanders? *Journal of Educational Psychology*, 92(1), 202-228. <https://doi.org/10.1037/0022-0663.92.1.202>
- Massy, W. F., & French, N. J. (2001). Teaching and learning quality process review: What the programme has achieved in Hong Kong. *Quality in Higher Education*, 1, 34-45. <https://doi.org/10.1080/13538320120045067>
- McGhee, D. E., & Lowell, N. (2003). Psychometric properties of student ratings of instruction in online and on-campus courses. *New Directions for Teaching & Learning*, 96, 39-48. <https://doi.org/10.1002/tl.121>
- McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist*, 52, 1218-1225. <https://doi.org/10.1037/0003-066X.52.11.1218>
- Muammar, O. M., & Alkathiri, M. S. (2021). What really matters to faculty members attending professional development programs in higher education. *International Journal for Academic Development*. <https://doi.org/10.1080/1360144X.2021.1897987>
- Nasser, F., & Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment & Evaluation in Higher Education*, 27(2), 187-198. <https://doi.org/10.1080/02602930220128751>
- Obenchain, K. M., Abernathy, T. V., & Wiest, L. R. (2001). The reliability of students' ratings of faculty teaching effectiveness. *College Teaching*, 49(3), 100-104. <https://doi.org/10.1080/87567550109595859>
- Perry, R. P., & Smart, J. C. (Eds.). (2007). *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Springer. <https://doi.org/10.1007/1-4020-5742-3>
- 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. <https://doi.org/10.1080/00223980109603696>
- Reznik, S. D., & Sazykina, O. A. (2017). Head of a university department: Competence and new activity priorities. *European Journal of Contemporary Education*, 6(1), 126-137. <https://doi.org/10.13187/ejced.2017.1.126>
- Richardson, J. T. E. (2005). Instruments for obtaining student feedback: A review of the literature. *Assessment and Evaluation in Higher Education*, 30(4), 387-415. <https://doi.org/10.1080/02602930500099193>
- Rovai, A. P., Ponton, M. K., Derrick, M. G., & Davis, J. M. (2006). Student evaluation of teaching in the virtual and traditional classrooms: A comparative analysis. *Internet and Higher Education*, 9, 23-35. <https://doi.org/10.1016/j.iheduc.2005.11.002>
- Scott, G., & Hawke, I. (2003). Using an external quality audit as a lever for institutional change. *Assessment and Evaluation in Higher Education*, 28(3), 323-332. <https://doi.org/10.1080/0260293032000059667>
- Seldin, P. (1999). Current practices – good and bad – nationally. In P. Seldin & Associates (Eds.), *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions* (pp. 1-24). Anker. <https://www.wiley.com/en-us/Changing+Practices+in+Evaluating+Teaching:+A+Practical+Guide+to+Improved+Faculty+Performance+and+Promotion+Tenure+Decisions-p-9781882982288>
- Shah, M., & Nair, C. S. (2012). The changing nature of teaching and unit evaluations in Australian universities. *Quality Assurance in Education*, 20(3), 274-288. <https://doi.org/10.1108/09684881211240321>
- Smith, R. (2005). Departmental leadership and management in chartered and statutory universities. *Educational Management Administration & Leadership*, 33(4), 449-464. <https://doi.org/10.1177/1741143205056305>

- Spooren, P., Mortelmans, D., & Denekens, J. (2007). Student evaluation of teaching quality in higher education: Development of an instrument based on 10 Likert-scales. *Assessment & Evaluation in Higher Education Journal*, 32(6), 667-679. <https://doi.org/10.1080/02602930601117191>
- Taherdoost, H. (2016). Sampling methods in research methodology: How to choose a sampling technique for research. *International Journal of Advance Research in Management*, 5(2), 18-27. <https://doi.org/10.2139/ssrn.3205035>
- Theall, M., & Franklin, J. (2001). Looking for bias in all the wrong places: A search for truth or a witch hunt in student ratings of instruction? In M. Theall, P. C. Abrami, & L. A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* [Special issue of *New Directions for Institutional Research*, 2001(109), 45-56. <https://doi.org/10.1002/ir.3>] Wiley.
- Tucker, B. M. (2015). *The student voice: Using student feedback to inform quality in higher education* (Doctoral dissertation). Curtin University, Perth. https://espace.curtin.edu.au/bitstream/handle/20.500.11937/2158/234289_Tucker%20B%202015.pdf
- Tucker, B. (2014). Student evaluation surveys: Anonymous comments that offend or are unprofessional. *Higher Education*, 68(3), 347-358. <https://doi.org/10.1007/s10734-014-9716-2>
- Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education*, 23(2), 191-212. <https://doi.org/10.1080/0260293980230207>
- Waschull, S. B. (2001). The online delivery of psychology courses: Attrition, performance, and evaluation. *Computers in Teaching*, 28, 143-147. https://doi.org/10.1207/S15328023TOP2802_15
- Whitworth, J., Price, B., & Randall, C. (2002). Factors that affect college of business student opinion of teaching and learning. *Journal of Education for Business*, 77, 282-289. <https://doi.org/10.1080/08832320209599677>
- Zhao, J., & Gallant, D. J. (2012). Student evaluation of instruction in higher education: Exploring issues of validity and reliability. *Assessment and Evaluation in Higher Education*, 37, 227-235. <https://doi.org/10.1080/02602938.2010.523819>
- Zineldin, M., Akdag, H. C., & Vasicheva, V. (2011). Assessing quality in higher education: New criteria for evaluating students' satisfaction. *Quality in Higher Education Journal*, 17(2), 231-243. <https://doi.org/10.1080/13538322.2011.582796>

Appendix 1

An online questionnaire for heads of academic department regarding whether they use student evaluation of instruction to make decisions about individual faculty members and their academic departments.

I am ...

- (1) Male
- (2) Female

I am a head of department in the following college cluster ...

- (1) Health
- (2) Engineering
- (3) Sciences and Management
- (4) Arts and Education

1. Overall, I tend to trust student evaluation of instruction results to make decisions in my job ...

- (1) Strongly disagree
- (2) Disagree
- (3) Slightly disagree
- (4) Slightly agree
- (5) Agree
- (6) Strongly agree

2. As a head of department, I would use the student evaluation of instruction results to make decisions about the ...

2.1 Individual faculty members' promotion/awarding of tenure

- (1) Strongly disagree
- (2) Disagree
- (3) Slightly disagree
- (4) Slightly agree
- (5) Agree
- (6) Strongly agree

2.2 Individual faculty members' effectiveness of teaching

- (1) Strongly disagree
- (2) Disagree
- (3) Slightly disagree
- (4) Slightly agree
- (5) Agree
- (6) Strongly agree

2.3 Individual faculty members' professional development needs

- (1) Strongly disagree
- (2) Disagree
- (3) Slightly disagree
- (4) Slightly agree
- (5) Agree
- (6) Strongly agree

2.4 Individual student learning experience

- (1) Strongly disagree
- (2) Disagree
- (3) Slightly disagree
- (4) Slightly agree
- (5) Agree
- (6) Strongly agree

3. As a Head of Department, I would use the student evaluation of instruction results to make decisions about the ...

3.1 Academic department's effectiveness of teaching

- (1) Strongly disagree
- (2) Disagree
- (3) Slightly disagree
- (4) Slightly agree
- (5) Agree
- (6) Strongly agree

3.2 Academic department's professional development needs

- (1) Strongly disagree
- (2) Disagree
- (3) Slightly disagree
- (4) Slightly agree
- (5) Agree
- (6) Strongly agree

3.3 Student cohort learning experience

- (1) Strongly disagree
- (2) Disagree
- (3) Slightly disagree
- (4) Slightly agree
- (5) Agree
- (6) Strongly agree

Thank you for your input.