Research Article

Improving Teaching Quality in Higher Education through Student Evaluation Feedback in Saudi Universities

Aisha Yousef Aljendan [1*]

[1*] Assistant Professor, Department of Educational Leadership, College of Education, King Faisal University, Alahsa, Saudi Arabia
* **Corresponding Author:** aaljendan@kfu.edu.sa

ARTICLE INFO	ABSTRACT
	Student evaluations of teaching (SET) are widely used in Saudi universities as a resource for recording and improving the quality of teaching. This paper studies the effect of SET on enhancing teaching quality in Saudi universities. It reports on findings from a narrative review in which the researcher conducted a PubMed and ERIC database for published journal articles that had been peer-reviewed. The researcher combined different search phrases to investigate how SET can lead to the improvement of teaching quality. The article is divided into sections that focus on specific issues related to SET and how to address these issues. The primary emphasis is on the fundamental concept of including student feedback to improve teacher quality. Considering the substantial allocation of resources towards SET and the high probability of their ongoing utilization for assessing teaching efficacy and educational achievements, optimizing the practical insights derived from these evaluations is imperative. The study findings indicate that designing the evaluation system, administering the evaluations, and following the subsequent procedures are crucial to improving the quality of teaching and learning through SET. The study also offers pedagogical implications to enhance understanding of the challenges faced by instructors at Saudi universities and suggests recommendations for ideal practice. **Keywords:** Student Evaluations, Teaching Quality, Higher Education, Student Feedback, Improvement.

INTRODUCTION

Student evaluations of teaching (SET) are now widely adopted within universities globally (Surgenor, 2013; Kuzmanovic, Savic, Popovic & Martic, 2012; Clayson & Haley, 2011). Many believe SETs are crucial for enhancing undergraduate teaching because they provide valuable feedback (Bangert, 2006; Nargundkar & Shrikhande, 2012). Several institutions have invested in establishing SET (Clayson & Haley, 2011).

The majority of SET is comprised of Likert-type scales and is employed for assessing different teaching components of teaching (Kuzmanovic et al., 2012). Nasser and Fresko (2002) stated that the evaluations of teaching and teachers are integral aspects of higher education and can assist with enhancing the overall quality of teaching.

Since governments and students make considerable investments in education (Pepe & Wang, 2012; Landrum & Braitman, 2008), teacher evaluations are critical to make greater returns on those investments. Increasing demands are being placed on colleges and universities to demonstrate proof of good teaching (Liu, 2012). Additionally, SET are frequently essential to the procedures necessary to achieve accreditation (Pepe & Wang, 2012).

While some instructors and administrators believe that student evaluations are a valuable tool for improving teaching quality, others criticise them for being biased, unreliable, and misused. For example, research has shown that evaluations could be affected by various aspects, including the race, gender, or prior academic performance of

students, rather than solely reflecting the quality of instruction. Additionally, some institutions base decisions about recruitment, tenure, and promotions solely or primarily on evaluations, which may require them to improve fairness and accuracy.

Surgenor (2013) determined that although faculty members' attitudes regarding the possible applications of summative SET were positive, they were concerned about the extent to which they were valid and beneficial in addition to their possible ramifications. Kulik (2001) suggests that some faculties are positively inclined towards the increasing usage of SET as it satisfies the rights of students to have a voice. Such faculties make an effort to consider student feedback. Members of the teaching staff have expressed positive responses regarding the feedback they have received, with some members proposing that the feedback obtained assisted with explaining why the teaching methodology and course content needed to be adapted (Moore & Kuol, 2005). Faculties have mentioned that positive feedback from evaluations increases their confidence in their competence as a faculty member and their characteristics and amiability (Arthur, 2009).

Nevertheless, studies show that faculty members sometimes value feedback from SET (Kogan, Schoenfeld-Tacher & Hellyer, 2010; Nasser & Fresko, 2002; Spiller & Ferguson, 2011). Spiller and Ferguson (2011) argue that specific teaching staff need more convincing about such evaluations due to the lack of relevance of SET questions to their lessons. Furthermore, faculty members have reported being concerned that student evaluations are significantly emphasised based on their belief that students do not have sufficient knowledge to conduct such evaluations (Simpson & Siguaw, 2000).

In Saudi Arabian universities, once students complete a course or semester, they request that they provide feedback on their instructors by completing surveys. Instructors and institutions use these assessments to gauge their instruction's quality, identify growth areas, and make personnel decisions. The surveys often ask students to score the instructor on various factors, including the instructor's ability to communicate clearly, their availability to answer questions, the consistency of their grading, and the overall quality of their teaching.

Although the utilization of SET for assessing the effectiveness of higher education teachers has been the subject of numerous studies in the literature, only a few researchers have investigated the benefits of SET in the Saudi Arabia context. Researchers have extensively studied SET in Western countries, and most of the findings support its usage and effectiveness in universities. However, further research is needed to investigate the effectiveness of SET in Saudi universities.

Therefore, this paper discusses how SET is administered and interpreted in universities in Saudi Arabia based on data and analysis of studies conducted predominantly in Western countries and Saudi Arabia.

This paper aims to answer the following three study questions:

1. Which factors affect the extent to which students' evaluation of teaching within higher education is reliable and valid?

2. How are SET currently implemented for higher education in Saudi Arabia, and what problems have been reported?

3. How can the implementation of SET be improved?

LITERATURE REVIEW

Student evaluation feedback is crucial for continuous endeavours to improve the quality of teaching in higher education. It is an excellent tool for assessing the efficacy of teaching techniques, curriculum, and educational experience. The knowledge acquired by student evaluations offers instructors and institutions essential data that can be utilised to recognise areas of excellence, tackle shortcomings, and execute focused enhancements. Educators can create a teaching environment that is more receptive to student needs and preferences by actively integrating and acknowledging student feedback. This approach promotes ongoing enhancement and guarantees exceptional quality education (Coelho & Ribeiro,2019).

Control theory states that people continually evaluate their actions about a behavioural objective, and if there is a difference, they strive to close it (Carver & Scheier, 1982). Professional reflection cannot occur without feedback, heightening awareness of actions and consequences. Challenging one's habitual patterns of thought and action makes it possible to break free from ingrained behaviour patterns. In addition, when there is a mismatch between objectives and actual results, feedback can help alleviate the negative feelings that come from them and encourage the positive feelings that come from achieving the desired results (Deci et al., 1999). Northcraft et al. (2011) found that performers performed better on the corresponding tasks when given better feedback.

Therefore, using student evaluation feedback to enhance teaching quality in higher education is critical to educational improvement, particularly in Saudi Arabian universities. The literature emphasises the significance of feedback in various learning contexts and its multifaceted role in evaluating student achievement, developing competencies, and enhancing motivation and confidence (Singh,2019; Abaidullah et al., 2014). Furthermore, the use of student feedback as an indirect assessment tool for evaluating teaching in higher education is widely acknowledged (Abaidullah et al., 2014). The importance of understanding students' perceptions of feedback is highlighted, as it can provide insights into the usefulness of different forms of feedback, efficient/inefficient feedback characteristics, and ways in which students can be supported by teachers to capitalise on feedback (Pǎduraru, 2023).

Studies have indicated that student evaluations of teaching positively influence the professional growth of faculty members and can significantly enhance their pedagogical methods. Utilising student evaluations as a feedback mechanism enables institutions to customise their teaching approaches to address the needs and expectations of students more effectively (Saidi & Vu,2021). This enhances the calibre of teaching and cultivates a favourable and captivating learning atmosphere. Moreover, student evaluations can be crucial to assessing faculty members' performance. In addition, using student evaluations to improve teaching quality shows a dedication to ongoing development and promotes a culture of responsibility and responsiveness throughout the academic community. Student evaluations give faculty members valuable insights into their teaching efficacy   and areas that require enhancement (Mahmoud, 2007). Educators can implement essential modifications to their instructional approaches, educational content, and classroom interactions, ultimately resulting in more captivating and efficient educational encounters for students. Student evaluations can help identify faculty members who excel in their teaching abilities and   recognise their efforts (Freng & Webber ,2009). This can contribute to faculty promotion decisions and incentivise other faculty members to improve their teaching practices.

Gaining a thorough comprehension of student evaluation feedback can also aid in adjusting institutional policies and practices to conform to the changing requirements and expectations of the student population. Universities can improve their academic standards and create a better learning environment for students by acknowledging the significance of this input in improving teaching quality.

The data obtained from Student Evaluation of Teaching (SETs) can offer valuable insights, such as the necessity to adapt teaching materials or teaching methods in a specific course. SETs have also been utilised to inform crucial decisions concerning faculty members' career advancement (Berk, 2005; Uijtdehaage & Neal, 2015; Stroebe, 2016). Nevertheless, the appropriateness of this approach is debatable when higher education institutions determine faculty members, including their promotion to higher academic ranks (typically associated with tenure and salary increments) and significant administrative positions within the institution (Berk, 2005; Stroebe, 2020). Furthermore, institutions often require new faculty members to include SETs from their previous institutions in their applications for a new position, thereby impacting their professional growth within their current institution and potential future employers (Madichie, 2011).

The focus on enhancing teaching quality has become a significant interest in Saudi Arabian universities. This is particularly important in the growing importance of student feedback and evaluations in shaping the educational system. Including students' perspectives has been identified as a critical initiative in higher education policy, emphasising the importance of understanding and utilising student feedback to improve teaching and learning outcomes.   Leveraging student evaluation surveys can provide valuable insights into students' specific needs and preferences, helping educators tailor their teaching methods to align with the requirements of student needs. Moreover, by implementing the feedback obtained from these evaluations, educational institutions in Saudi Arabia can ensure that their teaching quality meets international standards while catering to their student population's unique cultural and academic requirements. Developing sustainable feedback practices is essential to address broader concerns about staff-student relationships and the nature of learning in mass higher education (Carless et al., 2011).

Additionally, using student feedback for quality assurance in academic programs is common in various countries, including the UK, the USA, and Australia (Puteh & Habil, 2011). However, it is crucial to consider the specific cultural and contextual factors that may influence the reception and implementation of student feedback in Saudi Arabian higher education landscape. The literature highlights the importance of leveraging student evaluation feedback to enhance teaching quality in higher education, specifically focusing on the unique context of Saudi Arabian universities. Understanding student perceptions of feedback, developing sustainable feedback practices, and aligning feedback mechanisms with broader national development visions are crucial considerations for advancing the quality of teaching in higher education institutions in Saudi Arabia. The literature also underscores the need for comprehensive assessment mechanisms to measure progress towards

sustainability, aligning with the broader goals of Saudi Arabia's Vision 2030 initiative (Alshuwaikhat & Mohammed, 2017). This initiative emphasises stakeholders' active involvement and empowerment at all levels, which is pertinent to effectively utilising student evaluation feedback to improve teaching quality. As such, it is imperative   for Saudi universities to actively solicit and act upon student feedback to drive positive change and elevate the standards of higher education in the region.

METHODOLOGY

The current study follows Baumeister's (2012) and Green, Johnson and Adams (2006) recommendations for writing narrative reviews, given the importance of providing a holistic understanding of the subject. Furthermore, narrative reviews assist with answering a more comprehensive range of problems because they utilise studies with different approaches. Narrative reviews also help stimulate discussion and reflection.

A Pubmed and ERIC database search was conducted for published journal articles that had been peer-reviewed by combining different search phrases, namely 'The Effectiveness of Teaching', 'Higher Education institutions quality', 'Student Evaluations of Teaching', and 'The Evaluation of Faculty'. The included studies comprised original research studies or review papers that investigated subjects about the three main research concerns identified in the study. Exclusion criteria for this review paper encompassed studies not conducted in English and studies for which whole-text articles were unobtainable or unrelated to the three primary research issues.

George Boole, an English mathematician of the 19th century, devised a logical approach from which the Boolean operators AND, OR, and NOT were derived. The implementation involved combining three distinct concepts or themes using the 'AND' operator and combining alternative terms using the 'OR' operator and 'NOT' operator following Boolean logic principles.

When the Boolean operator "AND" is utilized to connect two phrases or words in a search query, the resulting search outcomes will exclusively consist of articles that include the specified phrases and words, limiting the scope of the research. As an illustration, the researcher accessed the databases on student feedback and teaching quality. The search resulted in the retrieval of articles encompassing the specific phrase and the individual term. Consequently, as the number of terms the researcher combined using the Boolean operator AND increased, the number of results about the intended subject matter diminished.

The utilization of the Boolean operation AND in the systematic search proved to be efficacious, as the conjunction of multiple phrases and/or words using the Boolean operator AND resulted in a reduction in the number of returned results for a match to be considered valid, all terms needed to be present. Research has revealed that implementing SET can lead to a decrease in irrelevant outcomes and an increase in outcomes relevant to establishing and operating student feedback systems. Consequently, this improvement in outcome relevance enhances the overall quality of university teaching.

When "OR" is used as a Boolean operator to connect two phrases or words in a search query, the search will retrieve articles that include any of the specified phrases or words, expanding the search scope. As an illustration, the researcher conducted a study on enhancing the quality of teaching or obtaining feedback from students in higher education. The completed search resulted in articles with specific terms, although only sometimes both phrases simultaneously. Consequently, the researcher observed increased search results pertinent to desired information by using more phrases and/or synonyms in conjunction with the Boolean operator OR.

When the Boolean operator NOT is placed between two words or phrases, the resulting search will exclude articles that contain any words or terms that appear after the Boolean operator NOT. By employing this approach, the search will be more focused and eliminate any extraneous information. During this review, the researcher encountered scholarly journal articles that specifically focused on students' feedback in higher education rather than general education. To exclude search results about general education, the researcher specifically inputted student feedback related to higher education NOT general education.

The use of Boolean operators in this study facilitated the retrieval of objects that contained either one or both of the specified phrases. This strategy proved highly effective in augmenting the quantity of search results obtained.

In addition to ERIC, the researcher utilized the Saudi Digital Library (SDL) database while searching in Saudi Arabia, selecting Some related articles due to the need for more research to be published in this field in the Saudi context.

Figure 1 summarizes search techniques applied in this study using Boolean operators.
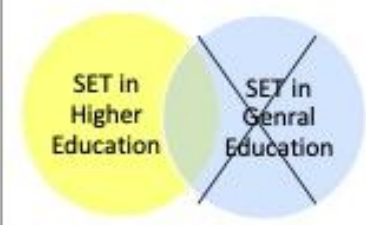
Boolean operator	Search examples	How it works?	Rustles
AND	Student feedback **AND** Teaching quality	AND **narrows** a search. It retrieves results that have both keywords.	Student feedback — Teaching quality
OR	Enhancing the quality of teaching **OR** Obtaining feedback from students in higher education	OR **expands** the search. It retrieves results that have either keyword.	Enhancing the quality of teaching in higher education — Obtaining feedback from students in higher education
NOT	SET in Higher Education **NOT** SET in General Education	NOT **narrows** the search. It excluded rustles that have unwanted keywords.	SET in Higher Education — SET in Genral Education

Figure 1. Using Boolean operators in this study

After the selection of publications, they were categorized according to specific themes and their corresponding sub-themes:

(a) Factors that affect the reliability and validity of students' evaluation of higher education teaching; (b) the current implementation of SET in Saudi higher education and the specific problems that have been reported; and (c) ways in which the implementation of SET can be improved such that a SET model with greater effectiveness can be developed.

RESULTS

Factors that affect the extent to which students' evaluation of teaching within higher education is reliable and valid:

SET has traditionally been implemented in higher education, and the literature (Madichie, 2011; Benton & Cashin, 2013) contains a significant number of studies on this particular subject. Nevertheless, many studies have also explored the various shortcomings of such evaluations and how researchers have attempted to overcome them.

Research shows that various factors influence SET scores (Madichie, 2011; Benton & Cashin, 2013; Schiekirka & Raupach, 2015; Oermann, Conklin, Rushton & Bush, 2018). When faculty and administrators attempt to interpret and use such data, their awareness of these variables must be enhanced.

The sections that follow provide an outline of these factors. Figure 2 summarizes specific elements that affect SET scores, which will be explained in more detail in the following section.
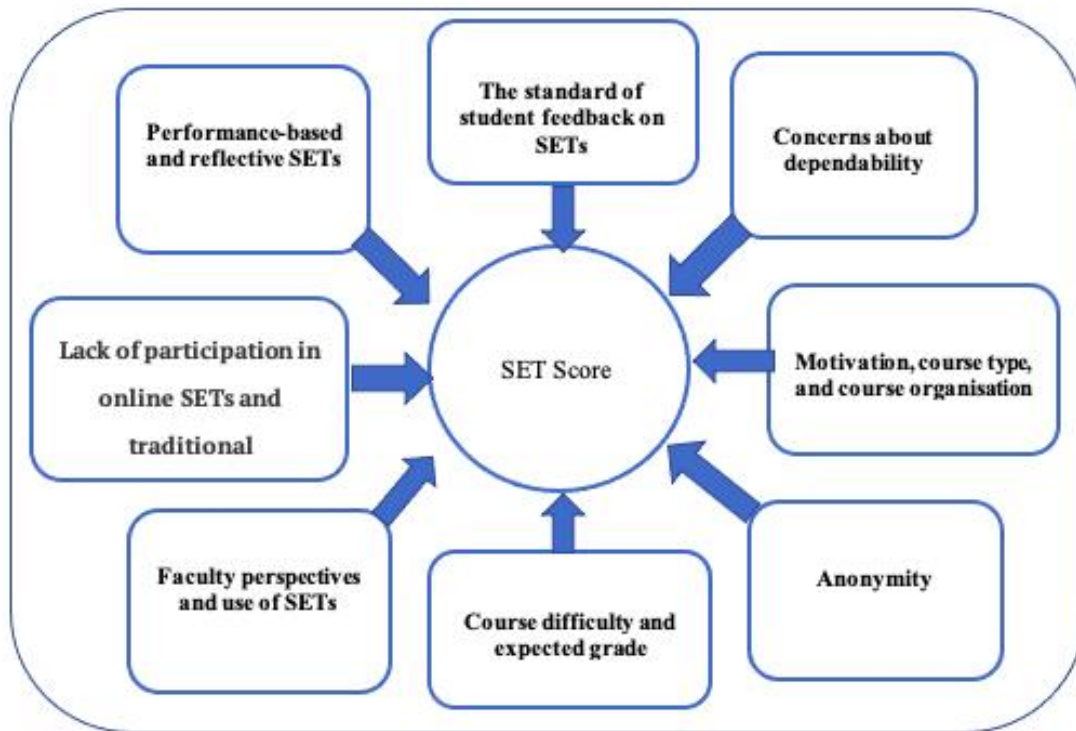

Figure 2. The standard of student feedback on SET

The Standard of Student Feedback on SET

Significant discussions are ongoing regarding the ability of students to accurately provide feedback on the extent to which teaching is effective (Jones, 2012; Husbands,1998; Remedios & Lieberman, 2008; Nasser & Fresko, 2002). It is essential that feedback accurately describes features of the course to the greatest extent possible, be grounded on perceptible behaviour and identify areas in which the instructor can improve the course (Svinicki, 2001).

Simpson and Siguaw (2000) argue that teachers could believe that their students have yet to reach a sufficient level of maturity for their judgements to be sufficiently mature. Nasser and Fresko, 2002) stated that this could call into question the appropriateness of using SET as the basis for decision-making processes related to tenure and promotions. When students are not careful when giving feedback or purposefully providing inaccurate information, this will inevitably harm the outcomes (Nargundkar & Shrikhande, 2012). Researchers such as Clayson and Haley (2011) have found that students' responses to SET must be more consistently truthful.

Educating students about SET may enhance the quality of feedback given by students via SET (Spiller & Ferguson, 2011; Alok, 2011; Svinicki, 2001). According to Keeley, English, Irons and Henslee (2013), minimal guidance is generally given in administering SET. Based on their recommendations, it should be emphasized to students that the SET may influence the teacher's performance in the future. Stowell, Addison, and Smith (2012) also offered analogous recommendations. Moreover, Keeley et al. (2013) suggested that students should be asked to provide separate answers to each item. For example, there should be no interconnection between the ratings given for a teacher's personality and their competence in the particular subject matter. Smith (2012) concluded that asking other teachers to evaluate instruction may not be feasible because most are experts in their specific areas instead of the education field.

Performance-Based and Reflective SET

The administration of SET generally occurs once a semester has ended and provides a measurement of

teaching proficiency (Boysen, 2015; Buchert, Laws, Apperson & Bregman, 2008; M. K. Winchester & T. M. Winchester, 2012; Stowell et al., 2012). In theory, SET give educators feedback regarding improving their teaching (Nargundkar & Shrikhande, 2012). Put differently, the function of SET should be formative and summative (M. K. Winchester & T. M. Winchester, 2014; Alok, 2011; Surgenor, 2013; Kuzmanovic et al., 2012). The contribution of summative SET to continuing lessons is minimal or nonexistent (M. K. Winchester & T. M. Winchester, 2012).

M. K. Winchester and T. M. Winchester (2012)suggested that students' motivation to engage in summative evaluations will be reduced if they know that the results will not be directly relevant to them. Alok (2011) discovered that students disliked SET administered at the end of the term, as they did not help improve the existing curriculum. Formative assessments allow students to perceive their participation as necessary, leading to greater class engagement and enhanced communication between students and teachers (Aultman, 2006). McGowan and Osguthorpe (2011) discovered that most students were open to assessment participation and sincerely wanted to offer feedback. Their motivation could be because they presupposed that constructive feedback would improve the lesson or teaching methodology (Slocombe, Miller & Hite, 2011).

M. K. Winchester and T. M. Winchester (2013) revealed the significance of engaging teachers in formative SET reflection. Participants in their study included members of the faculty of education in a university college in the UK. According to their findings, an increase in SET scores resulted from reflection, which rose in parallel with the degree of reflection.

The UK National Student Survey (NSS) is a famous example of a summative SET. The first administration of the NSS occurred in 2005 and involved undergraduate students during their last year. The popularity of the SET has increased exponentially. For example 2015, the survey was administered to more than 300,000 students in 350 institutions (Higher Education Funding Council for England, 2016). Despite its widespread usage and effectiveness, it still needs improvement. Cocksedge and Taylor (2013) conducted a study comparing the NSS with the Dundee Ready Education Environment Measure (DREEM). Both were primarily administered to a sample comprised of UK medical students (Roff et al., 1997; Roff, 2005). They discovered that in the context of this specific group, the DREEM had increased suitability.

Additionally, researchers criticised the fact that NSS data were aggregated before publication. Gibbons (2012) researched stress, satisfaction with the course and various factors in UK students studying psychology in their first year by employing NSS items and different sources. According to the results, Gibbons suggested that NSS data may be of more significant benefit if factors such as the sense of belonging to an intellectual motivation and learning community were considered.

Low Attention, Insufficient Time, and Low Response Rates

Compared with paper-based SET, online SET have increased in recent years. Nevertheless, research from the literature indicates that students may only complete online SET if they pay specific attention during the completion process. Uijtdehaage and O'Neal (2015) conducted a fascinating study in which they inserted fictional professors into the SET and asked students what they thought about them. Furthermore, the literature suggests that students typically assume that completing SET has no positive effect on education. This might result in a low response rate when combined with students' busy schedules (Spooren & Van Loon, 2012). Suppose the response rate needs to be sufficiently high. In that case, faculty members may need more motivation to consider the findings (Nowell, Gale & Kerkvliet, 2014), even though research indicates that feedback from students who fill in the assessment is similar to the group.

Course Difficulty and Expected Grade

Many researchers (Madichie, 2011; Benton & Cashin, 2013) have demonstrated a link between students' SET scores, examination grade expectations, and exam satisfaction. For example, Stroebe (2020) reported that feedback given in SET to instructors who are lenient in their grading and whose lessons are not tricky is positive; this may encourage teachers to teach less effectively and cause grades to be further inflated.

A recent Looi, Anderson, Bonner, Maguire and Reay (2020) study discovered that higher SET scores were associated with more accessible assessments. Additionally, some worrying research is beginning to emerge, suggesting that SET is inversely linked to subsequent course achievement. This indicates that increased ratings are related to a deterioration in student performance in subsequent studies. A study was conducted by Kornell and Hausman (2016) to determine whether students in higher education institutions obtain more excellent knowledge from instructors with high SET ratings. According to the study's findings, when end-of-course learning assessments were conducted, the teachers whose ratings were the highest also made the most significant contributions to student learning. This was determined by measuring how much students had learnt. Based on performance measurements for the following associated courses, the instructors who received the lowest ratings were the most successful. The researchers hypothesised that if they made the course more challenging in

constructive ways, it would result in lower evaluations while also enhancing learning. Resultantly, the scores that students give on evaluations should not be used to measure the quality of instruction and should be understood in context.

Anonymity

Students should be able to share their thoughts without fear of potential ramifications in their relationships with instructors. However, it should be highlighted that students could exhibit a lack of respect and be malicious towards teachers if their grade falls below their expectations (Tucker, 2014). This is because of the anonymity typically connected with online SET. Research reveals, however, that students rarely provide rude comments and that this can be reduced further by educating students on how to offer feedback constructively (Tucker, 2014). Afonso et al. (2005) found that the students evaluated their teaching performance using anonymous and name-based (open) evaluations. In general, when evaluations were conducted anonymously, the ratings given to faculty were reduced. Based on the study's findings, students responded more sincerely when evaluations were conducted anonymously due to specific obstacles in open evaluations, including the damage caused to the functional relationship with the relevant faculty member.

Motivation, Course Organization, and Course Type

Some studies have found that some courses and how they are organized impact SET rates. Min and Baozhi (1998) discovered in their research that students' SET rates varied significantly according to the subject of their course. For instance, regarding SET scores, clinical sciences were rated higher than basic sciences. According to the authors, this was because medical students aspire to become doctors, which means they are intensely passionate about and highly motivated to partake in clinical science courses instead of introductory science courses (Min & Baozhi,1998). As with electives, students' motivation for the subject positively affects their grades (Benton & Cashin, 2013). Several other studies (Madichie, 2011; Benton & Cashin, 2013) have found that the SET rate may vary according to the type of course taken.

Hessler et al. (2018) conducted an intriguing study that raised concerns about the validity of SET. In their research, they conducted a randomized controlled trial concerning a course on emergency medicine. Medical students in their third year were allocated to one of two groups on a random basis: one was the control, and the other was given unrestricted access to chocolate biscuits (biscuit group). Instruction in both groups was provided by identical teachers, and the course material and educational content were also identical. Following the course, all students had to take the SET. The findings indicated that the biscuit group's teacher evaluations were considerably better than those of the control group, thus raising questions about whether SET are valid (Hessler et al.,2018).

Furthermore, other studies have shown that fewer students within the classroom cause SET rates given to faculty to increase (Madichie, 2011; Min & Baozhi, 1998). This could be explained by fewer students, allowing teachers and students to interact more. Moreover, the environment in which these evaluations are conducted can influence the outcomes. For instance, it seems that both the time and day on which lessons are given can affect SET scores, in addition to the day on which the SET is filled in (e.g., evaluations administered later in the week generally produce increased positive feedback compared with when this is done at the beginning of the week) (Madichie, 2011).

Concerns about Dependability

Student evaluations of teaching produce reliability measures that must be revised to establish validity. As a result, SET may occasionally provide information about students rather than faculty members.

Clayson (2017) demonstrated that despite the widespread belief that SET give accurate answers and that there is stability among class and instructor averages, student responses vary. Resultantly, the findings of this evaluation indicate that students differ on the subject of evaluation.

Learning Outcomes and SET

The most significant reason for applying SET is to improve teaching, thus enhancing student learning. According to Boysen (2015), the results of SET have only a minimal relationship to factors such as student accomplishment. Clayson and Haley (2011) and Surgenor (2013) presented the findings of their research, which suggested that the results of SET may have no relationship to learning or that such a relationship could be adverse. According to Carrell and West (2010), teachers with more excellent experience received lower SET ratings, even though the students' performance in further associated courses increased.

Faculty Perspectives and Use of SET

Faculty members frequently exhibit concern that students who are not satisfied have a greater likelihood of

completing evaluations (Laubsch, 2006). Individuals who obtained better evaluations were more confident in the SET's validity (Nasser & Fresko, 2002). Faculty members may exhibit defensive attitudes towards SET and even be offended (Spiller & Ferguson, 2011). It is even possible that they may become anxious or feel humiliated by negative evaluations (Moore & Kuol, 2005).

Theall and Franklin (2001) found that negative assessments could result in instructors becoming angry towards students and their colleagues and substance abuse students being verbally abused or their grades being lowered (Theall & Franklin, 2001). Yao and Grady (2005) discovered that people whose careers are just starting reported SET-related anxiety and anxiousness more frequently. After receiving SET results, female teachers are more likely to express negative emotions such as fear, sadness, and rage, according to Kogan et al. (2010). This may explain why women are less likely to alter their behaviour in response to SET feedback (Kogan et al., 2010; Nasser & Fresko, 2002).

SET can only be used formatively when teachers use the data for modifying their courses. However, comparatively few teachers may employ SET to improve their teaching (Spiller & Ferguson, 2011; Simpson & Siguaw, 2000; Nasser & Fresko, 2002).

Certain faculty members may need more knowledge regarding the effective use of assessments (Theall & Franklin, 2001), which may discourage them from utilising students' comments to develop the lesson. Integrative Teaching Consultation (ITC) facilitates the process by which teachers can use SET results more effectively (Wibbecke, Kahmann, Pignotti, Altenberger & Kadmon, 2015). In ITC, teachers engage with experts about their SET scores and how their lessons can be modified. Therefore, when higher education institutions expect that instructors will modify their teaching according to SET results, they must incentivise them to achieve this (Nasser & Fresko, 2002).

Lack of Participation in Online Set and Traditional

Low response rates constitute a problem that often arises with SET. Weng, Weng, and Tsai (2014) proposed several methods to encourage people to complete the evaluations, such as informing people about the opportunity to do so and implementing actions to make students think that conducting the evaluations is standard procedure. They were surprised that teachers had a minimal effect on whether students wanted to participate. They stated that teachers should not imply that students would get into trouble if they did not complete the evaluations because that would reduce the likelihood that they would do so. Estelami (2014) told people to be careful when getting everyone to take SET. However, he did propose one method whereby SET completion should be made mandatory for registering for the next term. Alok (2011) noted that shortening SET could encourage more students to participate.

M. K. Winchester and T. M. Winchester (2012) examined students' attitudes regarding SET, how they are administered and how instructors use them with a qualitative approach. They discovered that while students from the United Kingdom and China enjoyed SET, their participation was not always guaranteed. Students cited insufficient time and the monotonous nature of repeatedly providing answers to identical questions as factors causing them not to complete SET. Students preferred the administration of SET to be less frequent, potentially only twice each term, and that such evaluations included both open- and closed-ended questions. Lastly, students expressed confidence that teachers exploited the data collected via SET. One of the factors that motivated students to complete SET was whether the data were used for class improvements.

The administration of SET is increasingly occurring in the online environment (Conn & Norris, 2005; Adams & Umbach, 2011). This makes them more available to every student; conversely, when SETs are administered within the classroom, they may be unable to participate due to absence (Stowell et al., 2012; Estelami, 2014).

SET administered in the online environment are generally cheaper and easier to implement; teachers can receive the results more quickly, and students are given more time to analyse the questions and offer more extensive answers to open-ended questions (Estelami, 2014). The results of such SET can also be tracked in the long term (Conn & Norris, 2005). In comparing SET administered online and in class, Stowell et al. (2012) found reduced student participation in online SET. Another observed difference was that the responses were longer when completing SET online. Otherwise, the outcomes of the two administration strategies were comparable. According to the authors, these conclusions were based on earlier research findings. Conn and Norris (2005) discovered that online administration of SET resulted in fewer student responses. They discovered that advertising the SET once the course had been completed, emphasising the value of feedback from students, and giving students reminders to complete the SET significantly improved online SET participation rates.

Currently, as well as SET, complete courses are now offered online. Over the past two decades, enrolment in online courses has expanded dramatically (Liu, 2012). There has been an increase in the number of courses offered online by higher education institutions; however, it is unclear how these courses should be evaluated

(Bangert, 2006; Jones, 2012; Liu, 2012) argues that as more courses are now given online, it is necessary to re-evaluate SET. SET designed for in-person classes may not be suitable for classes given online, meaning that SET explicitly designed for online courses should be developed (Bangert, 2006). The Student Assessment of Online Teaching Effectiveness is a particular SET for online courses (Bangert, 2006). Jones (2012) stated that when administered in the context of online courses, SET can assist in identifying various activities that online teachers can implement to increase students' satisfaction levels and SET ratings.

The Implementation of Set in Higher Education in Saudi Higher Education

The National Commission for Academic Accreditation and Assessment (NCAAA), an official government organization, was established in 2005 to verify that educational quality and institutional management are at global standards. As student satisfaction is a particularly critical aspect of quality management in higher education, the NCAAA has demanded that students conduct multiple evaluations to attest to the standard of post secondary education in the Kingdom of Saudi Arabia.

Student Learning and Teaching' is considered the most critical of the 11 areas determined by the NCAAA for evaluation based on internationally recognized standards of excellent practice. Requirements include: 'A comprehensive system for evaluating teaching effectiveness, including but not limited to student surveys'(NCAAA, 2005-2007).

The 'Course Evaluation Survey' (CES) of the NCAAA assesses the teaching efficacy of individual courses as a unit. Nevertheless, additional NCAAA standards also exist: first, 'faculty maintain a portfolio of evidence of evaluation and of strategies for improvement'; second, 'analyses and conclusions should be based on valid evidence rather than subjective impressions'; and third, 'benchmarks should include external comparison'. Despite these directions being informative and significant, they still need to be more comprehensive to comprehensively evaluate the strengths and weaknesses of teachers from a professional perspective and their teaching proficiency (NCAAA, 2007).

To comply with NCAAA regulations and enhance the quality of their work, all higher education institutions in Saudi Arabia have established internal quality-assurance systems such that the performance levels of faculty members can be reviewed and evaluated (Al-Ghamdi, Al-Gaid, & Abu-Rasain, 2010).

Students' evaluations of teaching (SET) processes in Saudi universities are considered critically important, as they inform significant decisions (Al-Rubaish et al., 2010; Arun Vijay, 2013). Nevertheless, the demand for data from SET assessments to be used when making staffing decisions has made it more likely that teaching staff will be judged unfairly, lose their tenure or even be wrongfully fired. Also, when students have this much power over their faculty careers and when faculty are expected to increase course loads, student enrolment, student satisfaction, and the overall quality of the programme may decline in the long run.

There is a significant body of literature on student evaluations of faculty (SET). Based on the findings of a study performed in Saudi Arabia (Al-Kuwaiti, 2014), teachers prefer an alternative supervision model that emphasizes an environment in which trust and responsibility are emphasized as opposed to one involving fear and imposition. As noted by Stark (2014), SET should not be implemented at all since they are significantly impacted by aspects that have no relationship to learning, including the ethnicity, gender, and physical appeal of the teacher. According to Chen and Hoshower (2003) and Penny and Coe (2004), when the ratings from SET assessments are employed for evaluating quality and guiding administrative policy decisions, students' feedback is then equated with the quality of teaching. Other scholars have proposed opposing arguments, such as Richardson (2005), Davis (2009) and Seldin (1999), who confirmed the validity of SET as an instrument, as well as Benton and Cashin (2013), who suggested that students' ratings should be combined with supporting evidence such that the validity and reliability criteria can be improved.

Several researchers have examined students' assessments of the teaching efficiency of university professors (Al-Kuwaiti, 2014). Other scholars have alluded to the pervasive fallacies related to the extent to which student ratings are valid and reliable (Al-Rubaish et al., 2010; Al-Kuwaiti, 2014; Burney, 1989; Curby, McKnight, Alexander & Erchov, 2019; Hornstein, 2017). Nevertheless, there needs to be more research on how faculty and administrators view the SET process and how it impacts job chances.

Elzubier and Rizk (2002) argued that assessments are grounded on the theory that, as experts in this context, students are the most qualified to judge the quality of their professors. Nonetheless, the validity of student feedback on classroom instruction is a contentious issue.

Multiple Saudi Arabian higher education institutions have been shown to overly rely on SET findings as a baseline for promotion and tenure decisions, despite divergent perspectives regarding the utility of SET surveys (Al-Rubaish et al., 2010). According to researchers, a more significant proportion of professors are harmful rather

than neutral or positive towards students' evaluations of teaching (Abrami, 2001; Nasser & Fresko, 2002; Theall & Franklin, 2001; Wachtel, 1998). Therefore, instruments like SET, administered unquestionably in higher education, should be scrutinized critically due to their impact on educators' lives. By investigating the perceptions of administrators and teachers regarding the extent to which SET surveys are effective and used within Saudi Arabia and scrutinizing the practice of making decisions regarding career advancement, salaries, and tenure, at least partially, on student assessments conducted anonymously, I aim to draw attention to how SET policy can impact the career trajectories of teachers. This information will benefit faculty and policymakers in evaluating, constructing, and refining their existing SET survey practices and other systems used for assessing teaching and learning.

Strategies Used to Improve Student's Evaluation of Teaching in Higher Education

Given the challenges with the implementation of SET, several studies have demonstrated that it may be possible to improve SET by focusing on certain significant factors that could impact the validity of such assessments.

According to Alok (2011), SET must incorporate particular instances of behaviour associated with the job to which students can respond. Additionally, the study referred to various kinds of items that are recommended to be excluded from SET. Regrettably, numerous SETs are probably designed by the institutions in which they are intended to be employed, and the design process may need to consider the findings of scientific research (Nargundkar & Shrikhande, 2012).

Alok (2011) also reported that the SET must be tailored to the specific institution to satisfy institutional needs optimally and accurately gather pertinent data. Moreover, the SET creation process should involve the participation of both faculty members and students. Vevere and Kozlinskis (2011) proposed that comparing the value of SET through universities could be beneficial and that the use of standard SET that specific universities could tailor to their needs through the addition of further questions could facilitate this comparison. The transfer and evaluation of knowledge, the availability of lecturers, and the characteristics of the lecturer are elements of SET that have specific importance.

According to Palermo (2013), there needs to be more knowledge regarding how and the extent to which educational systems are impacted by SET, and therefore, further studies are required. Several researchers have stressed the significance of enhancing the process by which student feedback is analyzed. For instance, Kuzmanovic et al. (2012) proposed conjoint analysis to improve standard SET, and Socha (2013) demonstrated how hierarchical linear modelling can enhance the analysis and interpretation of SET data. According to Alok (2011) and Nargundkar and Shrikhande (2012), regular updates should be made to SET.

Darwin (2012) presented a discussion on various issues about SET, including that study findings do not consistently demonstrate that the learning of students and performance of teachers is improved. In Australia, he developed and tested a different option to conventional SET. Workshops were held before and after the semester, where an introduction to the designed programme was presented, questions were developed for evaluation, and the findings were discussed among the key features. Throughout the semester, instructors were encouraged to communicate with one another. Students were given open-ended, qualitative questions as the semester approached its completion. Darwin (2012) provided evidence that the programme was largely successful, where the feedback provided was richer than expected.

A specific way to improve the assessment process is to support data from SET with additional sources. Smith, 2012 stated that most UK universities use a combination of peer observation of teaching (POT) and SET. Thampy, Bourke, and Naran (2015) found that the latter can effectively complement SET data, serving both formative and summative functions. The researchers found that although UK medical school faculty members hesitated to use POT, they acknowledged that the positives outweighed the negatives. According to Smith (2012), in the context of UK universities, POT offers benefits that could render them more advantageous compared with SET. There is an increased likelihood that people will be open and criticize their performance if POT does not influence their salaries or tenure decisions. This may mean that teachers do not take the process seriously enough or choose not to participate.

DISCUSSION

According to the literature, SET must be designed to evaluate the effectiveness of higher education teaching. While a large body of research has focused on faculty evaluations in the context of Western higher education, this is not the case for student evaluations in Saudi universities (Royal, Jay Temple, Neel & Nelson, 2018).

The narrative review in the current study was an attempt to address various wide-ranging issues about SET and how they are used in higher education institutions in Saudi Arabia. The study discussed key questions and endeavored to evaluate the topic by providing information to readers about the extant understanding of (a) factors affecting the reliability and validity of students' evaluation of teaching in higher education, (b) the current implementation of SET in Saudi higher education and the specific problems that have been reported. Furthermore, the paper investigated how SET can be improved to develop an SET model with increased effectiveness.

My review has shown that faculty members and administrators should have increased awareness of significant elements that could impact the results when using SET, including students' lack of attention and time, bias resulting from course complexity and expected grades, and the individual motivations of students (Madichie, 2011; Benton & Cashin, 2013; Green et al., 2006). Therefore, each SET should be carefully prepared and revised to account for the elements that influence SET results. Efforts should be made to establish a practical SET to assess teaching effectiveness objectively. The fact that teachers and students have differing perspectives and expectations regarding the scope and aims of SET is one of the barriers frequently encountered in SET. This initial barrier should be overcome by adequately preparing both teachers and students (e.g., through training on the purpose and interpretation of SET) (e.g., on objectives, how the results will be employed, as well as by divulging how feedback from finished SETs has been reviewed and applied).

As well as utilizing SET with increased effectiveness, universities in Saudi Arabia should strive to create an extensive system of evaluation that facilitates the effective measurement of teaching quality, promotes the professional development of teachers in the field of medicine and enhances medical teaching quality. According to evidence in the literature, faculty members and students should be engaged in the SET improvement process to improve how we analyse student feedback. At the same time, other sources should supplement SET data. This kind of wide-ranging programme should be aimed at collecting and triangulating data from various sources, including students, peers, programme administrators, and faculty members, to promote professional development and enhance the teaching standards in higher education institutions in Saudi Arabia. A particular drawback of this type of programme could be that the work and resource requirements of faculty, students, and administrators may be increased, possibly increasing the pressures on all parties. This possible constraint should be considered when designing the programme, and a pilot test should be conducted before the entire programme is implemented. Furthermore, careful consideration should be given to how the findings of different data collection methods can be amalgamated and analyzed, as well as how feedback is communicated with administrators and faculty.

In conclusion, the study's findings indicate that every Saudi university should continually endeavour to enhance students' education quality. To achieve this, the environment in which faculty work must be supportive and caring. The data gathered will have increased objectivity, faculty will receive more accurate feedback, and the mechanisms required to support faculty will be implemented to maximize their teaching potential by developing a holistic evaluation system.

To sum up, one of the essential ways to improve the quality of teaching and learning in Saudi universities is to take comprehensive actions to enhance the quality of SLT. Some measures can be used, including Increasing educational technologies through combining learning management systems into publishers' digital resources and, introducing immediate response systems through mobile phones, customizing faculty workshops to identify areas of concern linked to teaching and learning. Considering the substantial allocation of resources towards SET and the high probability of their ongoing utilization, Saudi universities must enhance their understanding of the best way to design and implement SET. Additionally, it is imperative to implement measures aimed at enhancing the quality of student responses and supporting educators in optimizing the utilization of the data for optimal advantages.

Also, Saudi universities can use other means to complement SET data to improve the quality of teaching and learning. For example, peer observation of teaching can be a valuable supplement to SET data and can serve both formative and functions (Thampy et al., 2015).

This study has some limitations; one is that it exclusively focuses on students' feedback. Future research can extend the illustration by including all other actors affecting the teaching-learning improvement. Researchers might go deeper by considering teachers' experiences and views of student's feedback. Research can also compare student feedback for different disciplines to identify how particular teaching and learning issues differ across different disciplines.

Finally, the researcher recognizes that the current study takes the form of a narrative review, and a specific limitation of this approach is that in attempting to find answers to various broad questions relating to the subject, it may have limited the ability to concentrate in increased detail on studies performed in associated fields.

Research needs to re-examine the primary themes and primary research in greater detail to find more specific answers to the resulting questions.

CONCLUSION

The research implies that higher education institutions such as Saudi universities should not rely solely on SET to assess the effectiveness of their teaching, even though they devote considerable amounts of money, effort, and time to such evaluations. SET should not be reserved only for use by the management when judging the value of individual courses or selecting faculty members (e.g., career advancement decisions, pay rises, and promotion to critical administrative positions). Saudi Universities may continue to employ SET, but they should be adequately planned, given, and evaluated based on the most recent data in the literature and with the participation of students, staff, and school management (Nowell et al., 2014). The ultimate objective for Saudi universities should be to build a complete evaluation system based on evidence. Literature supports using different methods to enhance the implementation of SET in higher education. These methods can help Saudi universities develop new strategies and improve the quality of teaching by implementing a comprehensive evaluation system.

In conclusion, this study recommended that the centers of learning and teaching in Saudi universities should conduct longitudinal and comprehensive analyses of evaluation reports that were written based on student responses collected anonymously via open-ended questions and collected either in an online format or a paper-based, identify persistent issues to transfer strategic decisions into impactful improvements that possess the capacity to affect all stakeholders. It is also critical to construct institutional-level strategic decisions regarding learning and teaching.

ACKNOWLEDGEMENT

FUNDING

REFERENCES

Abaidullah, A.M., Ahmed, N., & Ali, E.A. (2014). Identifying Hidden Patterns in Students' Feedback through Cluster Analysis. *International Journal of Computer Theory and Engineering, 7*, 16-20.

Abrami, P. C. (2001). Improving judgments about teaching effectiveness using teacher rating forms. *New Directions for Institutional Research, 2001*(109), 59-87.

Adams, M. J., & Umbach, P. D. (2011). Non-response and online student evaluations of teaching: Understanding the influence of salience, fatigue, and academic environments. *Research in Higher Education, 53*(5), 576-591.

Afonso, N. M., Cardozo, L. J., Mascarenhas, O. A., Aranha, A. N., & Shah, C. (2005). Are anonymous evaluations a better assessment of faculty teaching performance? A comparative analysis of open and anonymous evaluation processes. *Family medicine, 37*(1), 43–47.

Al-Ghamdi, A., Al-Gaid, A., & Abu-Rasain, M. (2010). Faculty evaluation in Saudi Arabia: A suggested model. *The Saudi Journal of Higher education, 7*(7), 85.

Al-Kuwaiti, A. (2014). Health science students′ attitude towards research training programs in the Kingdom of Saudi Arabia: Reliability and validity of the questionnaire instrument. *Journal of Family and Community Medicine, 21*(2), 134.

Alok, K. (2011). Student Evaluation of Teaching: An Instrument and a Development Process. *The International Journal of Teaching and Learning in Higher Education, 23*, 226-235.

Alshuwaikhat, H. M. & Mohammed, I. (2017). Sustainability Matters in National Development Visions—Evidence from Saudi Arabia's Vision for 2030. *Sustainability, 9*(3), 408. doi: 10.3390/su9030408

Arthur, L. (2009). From performativity to professionalism: Lecturers' responses to student feedback. *Teaching in Higher Education, 14*(4), 441–454.

Arun Vijay, S. (2013). Appraisal of students ratings as a measure to manage the quality of higher education in India: An institutional study using Six Sigma model approach. *International Journal for Quality Research, 7*(3), 3–14.

Aultman, L. P. (2006). An unexpected benefit of formative student evaluations. *College Teaching, 54*(3), 251-285.

Bangert, A. W. (2006). The development of an instrument for assessing online teaching effectiveness. *Journal of Educational Computing Research, 35*(3), 227-244.

Baumeister, R. F. (2012). *The Portable Mentor*. New York, NY: Springer.

Benton, S. L., & Cashin, W. E. (2013). Student ratings of instruction in college and university courses. In *Higher education: Handbook of theory and research: Volume 29* (pp. 279-326). Dordrecht, Netherlands: Springer.

Berk, A. (2005). Survey of 12 strategies to measure teaching effectiveness. *Int J Teach Learn Higher Educ, 17*(1), 48–62.

Boysen, G. A. (2015). Significant interpretation of small mean differences in student evaluations of teaching despite explicit warning to avoid overinterpretation. *Scholarship of Teaching and Learning in Psychology, 1*(2), 150-162.

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.

Burney, F. A. (1989). Students' evaluation of teaching in a Saudi Arabian university. *British Journal of Educational Technology, 20*(3), 200-208.

Carless, D., Salter, D.J., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education, 36*, 395 - 407.

Carrell, S. E., & West, J. E. (2010). Does professor quality matter? Evidence from random assignment of students to professors. *Journal of Political Economy, 118*(3), 409-432.

Carver, C. S., & Scheier, M. F. (1982). Control Theory: A Useful Conceptual Framework for Personality-Social, Clinical, and Health Psychology. *Psychological Bulletin, 92*, 111-135. doi: 10.1037/0033-2909.92.1.111

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. (2017). Student evaluation of teaching and matters of reliability. *Assessment & Evaluation in Higher Education, 43*(4), 666-681.

Clayson, D. E., & Haley, D. A. (2011). Are students telling us the truth? A critical look at the student evaluation of teaching. *Marketing Education Review*, *21*(2), 101-112.

Cocksedge, S. T., & Taylor, D. C. (2013). The National Student Survey: Is it just a bad dream?. *Medical Teacher*, *35*(12).

Coelho, L D A., & Ribeiro, M M D L D O. (2019). Student ratings to evaluate the teaching effectiveness: Factors should be considered. doi: 10.4995/head19.2019.9392

Conn, C., & Norris, J. (2005). Investigating strategies for increasing student response rates to online delivered course evaluations. *2003 Annual Proceedings-Anaheim*, *77*.

Curby, T., McKnight, P., Alexander, L., & Erchov, S. (2019). Sources of variance in end-of-course student evaluations. *Assessment & Evaluation in Higher Education*, *45*(1), 44-53.

D. Royal, K., Jay Temple, L., A. Neel, J., & L. Nelson, L. (2018). Psychometric validation of a medical and Health Professions Course Evaluation Questionnaire. *American Journal of Educational Research*, *6*(1), 38-42.

Darwin, S. (2012). Moving beyond face value: Re-envisioning higher education evaluation as a generator of professional knowledge. *Assessment & Evaluation in Higher Education*, *37*(6), 733-745.

Davis, B. G. (2009). *Tools for teaching*. San Francisco, CA: Jossey-Bass.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627–668. https://doi.org/10.1037/0033-2909.125.6.627.

Elzubeir, M., & Rizk, D. (2002). Evaluating the quality of teaching in medical education: Are we using the evidence for both formative and summative purposes?. *Medical Teacher*, *24*(3), 313-319.

Estelami, H. (2014). The effects of survey timing on student evaluation of teaching measures obtained using online surveys. *Journal of Marketing Education*, *37*(1), 54-64.

Freng, S., & Webber, D J. (2009). Turning up the Heat on Online Teaching Evaluations: Does "Hotness" Matter?. Retrieved from https://scite.ai/reports/10.1080/00986280902959739.

Gibbons, C. (2012). Stress, positive psychology and the National Student Survey. *Psychology Teaching Review*, *18*(2), 22-30.

Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: Secrets of the Trade. *Journal of Chiropractic Medicine*, *5*(3), 101-117.

Hassan, A., Al Ali, A., Mokabel, F., Hegazy, M., Wosornu, L., Al-Rubaish, A., & Abdel Rahim, S. I. (2010). Developing questionnaires for students′ evaluation of individual faculty teaching skills: A Saudi Arabian pilot study. *Journal of Family and Community Medicine*, *17*(2), 91.

Hessler, M., Pöpping, D. M., Hollstein, H., Ohlenburg, H., Arnemann, P. H., Massoth, C., . . . Wenk, M. (2018). Availability of cookies during an academic course session affects evaluation of teaching. *Medical Education*, *52*(10), 1064-1072.

Higher Education Funding Council for England (2016). Retrieved from http://www.hefce.ac.uk/lt/nss/

Hornstein, H. A. (2017). Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education*, *4*(1), 1304016.

Husbands, C. T. (1998). Implications for the assessment of the teaching competence of staff in higher education of some correlates of students' evaluations of different teaching styles. *Assessment & Evaluation in Higher Education*, *23*(2), 117-140.

Jones, S. J. (2012). Reading between the lines of online course evaluations: Predictable actions that improve student perceptions of teaching effectiveness and course value. *Online Learning*, *16*(1).

Keeley, J. W., English, T., Irons, J., & Henslee, A. M. (2013). Investigating halo and ceiling effects in student evaluations of instruction. *Educational and Psychological Measurement*, *73*(3), 440-457.

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*(6), 623-636.

Kornell, N., & Hausman, H. (2016). Do the best teachers get the best ratings?. *Frontiers in Psychology*, *7*, 570.

Kulik, J. A. (2001). Student ratings: Validity, utility, and controversy. *New Directions for Institutional Research*, *2001*(109), 9-25.

Kuzmanovic, M., Savic, G., Popovic, M., & Martic, M. (2012). A new approach to evaluation of university teaching considering heterogeneity of students' preferences. *Higher Education*, *66*(2), 153-171.

Landrum, R. E., & Braitman, K. A. (2008). The effect of decreasing response options on students' evaluation of instruction. *College Teaching*, *56*(4), 215-218.

Laubsch, P. (2006). Online and in-person evaluations: A literature review and exploratory comparison. *MERLOT Journal of Online Learning and Teaching, 2*(2), 62-73.

Liu, O. L. (2011). Student evaluation of instruction: In the new paradigm of Distance Education. *Research in Higher Education*, *53*(4), 471-486.

Looi, J. C., Anderson, K., Bonner, D., Maguire, P., & Reay, R. (2020). Student Evaluations of Teaching (set): Implications for medical education in psychiatry and an approach to evaluating set and student performance. *Australasian Psychiatry*, *28*(4), 463-465.

Madichie, N. O. (2011). Students' evaluation of teaching (SET) in Higher Education: A question of reliability and validity. *The Marketing Review*, *11*(4), 381–391. doi: 10.1362/146934711x13210328715984.

Mahmoud, H G. (2007). Comparison of Clinical Instructors' Teaching Behaviors as Perceived by Themselves with Those Assessed by their Students in the Clinical Setting. Retrieved from https://scite.ai/reports/10.21608/jhiph.2007.22312.

McGowan, W. R., & Osguthorpe, R. T. (2011). 12: Student and faculty perceptions of effects of midcourse evaluation. *To Improve the Academy*, *29*(1), 160-172.

Min, Q., & Baozhi, S. (1998). Factors affecting students' classroom teaching evaluations. *Teaching and Learning in Medicine*, *10*(1), 12-15.

Moore, S., & Kuol, N. (2005). Students evaluating teachers: Exploring the importance of faculty reaction to feedback on teaching. *Teaching in Higher Education*, *10*(1), 57-73.

Nargundkar, S., & Shrikhande, M. (2012). An empirical investigation of student evaluations of instruction-the- the relative importance of factors. *Decision Sciences Journal of Innovative Education*, *10*(1), 117-135.

Nasser, F., & Fresko, B. (2002). Faculty Views of Student Evaluation of College Teaching. *Assessment & Evaluation in Higher Education*, *27*(2), 187-198.

Northcraft, G. B., Schmidt, A. M., & Ashford, S. J. (2011). Feedback and the rationing of time and effort among competing tasks. *Journal of Applied Psychology, 96*(5), 1076–1086. doi: 10.1037/a0023221

Nowell, C., Gale, L. R., & Kerkvliet, J. (2014). Non-response bias in student evaluations of teaching. *International Review of Economics Education*, *17*, 30-38.

Oermann, M. H., Conklin, J. L., Rushton, S., & Bush, M. A. (2018). Student Evaluations of Teaching (set): Guidelines for their use. *Nursing Forum, 53*(3), 280-285.

Păduraru, M., Rihițeanu-Năstase, E., & Stăiculescu, C. (2023). Students' Perceptions Of Feedback In Higher Education. *Proceedings of the 9th International Conference Education Facing Contemporary World Issues (Edu World 2022),* 233-241. doi: 10.15405/epes.23045.36

Palermo, J. (2011). Linking student evaluations to institutional goals: A change story. *Assessment & Evaluation in Higher Education*, *38*(2), 211-223.

Penny, A. R., & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research*, *74*(2), 215-253.

Pepe, J., & Wang, M. (2012). What instructor qualities do students reward?. *College Student Journal, 46*(3), 603-614.

Puteh, M., & Habil, H. (2011). *Student Feedback: The Cornerstone to an Effective Quality Assurance System in Higher Education* (pp. 49-60). Oxford, UK: Chandos Publishing.

Remedios, R., & Lieberman, D. A. (2008). I liked your course because you taught me well The influence of grades, workload, expectations and goals on students' evaluations of teaching. *British Educational Research Journal, 34*(1), 91-115.

Richardson, J. T. (2005). Instruments for obtaining student feedback: A review of the literature. *Assessment & Evaluation in Higher Education*, *30*(4), 387-415.

Roff, S. (2005). The Dundee Ready Educational Environment Measure (dream)—a generic instrument for measuring students' perceptions of undergraduate health professions curricula. *Medical Teacher*, *27*(4), 322-325.

Roff, S., McAleer, S., Harden, R. M., Al-Qahtani, M., Ahmed, A. U., Deza, H., Primparyon, P. (1997). Development and validation of the Dundee Ready Education Environment Measure (Dreem). *Medical Teacher*, *19*(4), 295-299.

Saidi, L G., & Vu, P. (2021). *Students' perspective on higher educators: interpersonal relationship matters*. Retrieved from: https://scite.ai/reports/10.1007/s43545-021-00176-9.

Schiekirka, S., & Raupach, T. (2015). A systematic review of factors influencing student ratings in Undergraduate Medical Education Course Evaluations. *BMC Medical Education*, *15*(1). doi: 10.1186/s12909-015-0311-8

Seldin, P. (1999). *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions*. Bolton, MA: Anker Publishing.

Simpson, P. M., & Siguaw, J. A. (2000). Student Evaluations of Teaching: An Exploratory Study of the Faculty response. *Journal of Marketing Education*, *22*(3), 199-213.

Singh, 1. (2019). *Lecturer's Feedback and Its Impact on Student Learning: A Study of a Public University in Sarawak, Malaysia*. Asia Journal of University Education (AJUE).

Slocombe, T., Miller, D., & Hite, N. (2011). A survey of student perspectives toward Faculty Evaluations. *American Journal of Business Education (AJBE)*, *4*(7), 51-58.

Smith, H. (2012). The unintended consequences of grading teaching. *Teaching in Higher Education*, *17*(6), 747-754.

Socha, A. (2013). A hierarchical approach to students' assessments of instruction. *Assessment & Evaluation in Higher Education*, *38*(1), 94-113.

Spiller, D., & Ferguson, P. B. (2011). Student evaluations: Do lecturers value them and use them to engage with student learning needs?. Retrieved from http://usir.salford.ac.uk/id/eprint/16999/.

Spoorena, P., & Van Loon, F. (2012). Who participates (not)? A non-response analysis on students' evaluations of teaching. *Procedia -Social and Behavioral Sciences*, *69*, 990-996.

Stark, P., & Freishtat, R. (2014). An evaluation of course evaluations. *Science Open Research*, 1-7. doi: 10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1

Stowell, J. R., Addison, W. E., & Smith, J. L. (2012). Comparison of online and classroom-based student evaluations of instruction. *Assessment & Evaluation in Higher Education*, *37*(4), 465-473.

Stroebe, W. (2016). Why Good Teaching Evaluations May Reward Bad Teaching: On Grade Inflation and Other Unintended Consequences of Student Evaluations. *Perspect Psychol Sci, 11*(6), 800−816.

Stroebe, W. (2020). Student evaluations of teaching encourages poor teaching and contributes to grade inflation: A theoretical and empirical analysis. *Basic and Applied Social Psychology*, *42*(4), 276-294.

Surgenor, P. (2013). Obstacles and opportunities: Addressing the growing pains of summative student evaluation of teaching. *Assessment & Evaluation in Higher Education*, *38*(3), 363-376.

Svinicki, M. D. (2001). Encouraging your students to give feedback. *New Directions for Teaching and Learning*, 17-24. doi: 10.1002/tl.24

Thampy, H., Bourke, M., & Naran, P. (2015). Peer-supported review of teaching: An evaluation. *Education for Primary Care*, *26*(5), 306-310.

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? *New Directions for Institutional Research*, *109*, 45-56.

Tucker, B. (2014). Student evaluation surveys: Anonymous comments that offend or are unprofessional. *Higher Education*, *68*(3), 347-358.

Uijtdehaage, S., & O'Neal, C. (2015). A curious case of the Phantom professor: Mindless teaching evaluations by medical students. *Medical Education*, *49*(9), 928-932.

Verve, N., & Kozlinskis, V. (2011). Students' Evaluation of Teaching Quality. *US-China Education Review, 5,* 702–708.

Wachtel, H. K. (1998). Student Evaluation of College Teaching Effectiveness: A brief review. *Assessment & Evaluation in Higher Education*, *23*(2), 191-212.

Weng, C., Weng, A., & Tsai, K. (2014). Online teaching evaluation for higher quality education: Strategies to increase university students' participation. *Turkish Online Journal of Educational Technology-TOJET, 13*(4), 105–114.

Wibbecke, G., Kahmann, J., Pignotti, T., Altenberger, L., & Kadmon, M. (2015). Improving teaching on the basis of student evaluation: integrative teaching consultation. *GMS Zeitschrift fur medizinische Ausbildung*, *32*(1). doi: 10.3205/zma000944

Winchester, M. K., & Winchester, T. M. (2012). If you build it, will they come? Exploring the student perspective of weekly student evaluations of teaching. *Assessment & Evaluation in Higher Education*, *37*(6), 671-682.

Winchester, T. M., & Winchester, M. K. (2013). A longitudinal investigation of the impact of faculty reflective practices on students' evaluations of teaching. *British Journal of Educational Technology*, *45*(1), 112-124.

Yao, Y., & Grady, M. L. (2005). How do faculty make formative use of student evaluation feedback?: A multiple case study. *Journal of Personnel Evaluation in Education, 18*(2), 107-126.