Original Article

Nursing students' perceptions of teaching effectiveness evaluations: A cross-sectional descriptive study

Maliheh Poorkiani [1], Nasser Al Salmi [1], Aisha Alshdefat [2] Zahra Sheikhalipour [3*]

[1]Department of Adult Health and Critical Care, College of Nursing, Sultan Qaboos University. Muscat, Oman
[2]Department of Maternal and Child Health, College of Nursing, Sultan Qaboos University. Muscat, Oman
[3]Department of Medical and Surgical Nursing, Nursing and Midwifery School, Tabriz University of Medical Sciences, Tabriz, Iran

Article info

Article history:
Received 24 Jul. 2024
Accepted 5 Feb. 2025
Published 14 Apr. 2025

***Corresponding author:**
Zahra Sheikhalipour, Department of Medical and Surgical Nursing, Nursing and Midwifery School, Tabriz University of Medical Sciences, Tabriz, Iran
Email: sheikhalipourz@gmail.com

How to cite this article:
Poorkiani M, Al Salmi N, Alshdefat A, Sheikhalipour Z. Nursing students' perceptions of teaching effectiveness evaluations: A cross-sectional descriptive study. J Med Edu Dev. 2025; 18(1): 115-122.

Abstract

Background & Objective: While students' perceptions are rarely discussed in studies on student evaluations of teaching effectiveness, the majority concentrate on the features, validity, and reliability of the questionnaire used, the variables affecting students' ratings, and the proper assessment of teaching quality. Therefore, this study aims to look into how students at Sultan Qaboos University's College of Nursing perceive and behave when evaluating the faculty.

Materials & Methods: This cross-sectional descriptive study was carried out in the academic year 2023–2024 during the summer and fall semesters. Two hundred thirty undergraduate students from Sultan Qaboos University's College of Nursing participated in the study. Convenience sampling was employed to gather data. A self-administered questionnaire was used to learn more about the students' opinions of Teaching Effectiveness (SETE). SPSS version 29 was used to analyze the data. The sample's demographics and the items about teaching effectiveness were described using frequency analysis. Furthermore, a correlational test was used to find relationships between SETE levels and demographic characteristics.

Results: According to the study, the estimated presence of false information was 47.24%, and the mean perceived accuracy of evaluations was 50.78. A sizable percentage acknowledged providing or being aware of assessments that were either higher or lower than appropriate. There were different opinions about the evaluation process; 84.6% of respondents supported student evaluations, while 55.1% thought teachers read comments. With 76.1% of respondents considering false information to be cheating, ethical issues were raised. Statistical analyses revealed no significant relationships between sex and questionnaire responses, nor between evaluation perceptions, demographic characteristics, and Grade Point Average (GPA).

Conclusion: The results show how complicated student opinions are regarding teacher assessments, particularly when it comes to truth and morality. The study calls for improvements in SETE processes to promote more reliable and reasonable evaluations.

Keywords: teaching evaluation, student evaluation of teaching effectiveness, nursing students

Introduction

A common approach to ensure student success and satisfaction with their education is to allow them to share their perspectives on their learning experiences [1]. The primary goal of student ratings is to improve the course experience for students and expand the quality of teaching [2].

The Student Evaluation of Teaching Effectiveness (SETE) is an instrument used to measure students' perspectives when evaluating teaching outcomes. For several reasons, proponents of SETE think that student evaluation of instruction is crucial. These assessments guarantee the caliber of instruction at universities, offer a separate way to measure the efficacy of instructors, direct choices about significant curriculum modifications, support faculty professional development, and aid in creating a framework to measure and incentivize successful teaching outcomes [3] more accurately.

As a result, in many universities, these evaluations are the most important and, in many cases, the only measure for quality assurance, evaluating faculty performance, and judging merit pay, tenure, and promotion [1, 4].

[Downloaded from edujournal.zums.ac.ir on 2026-11-28]

[DOI: 10.61186/edcj.18.1.115]

Copyright © 2025 Zanjan University of Medical Sciences. Published by Zanjan University of Medical Sciences.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International license (https://creativecommons.org/licenses/by-nc/4.0/). Noncommercial uses of the work are permitted, provided the original work is properly cited.

However, a literature review has shown that SETEs may need to be more balanced and fully reflective of teaching effectiveness. Biases that may affect how students assess instruction have been found in empirical research, which could produce unreliable findings [5]. As a result, SETEs run the risk of measuring things other than how well students learn. Furthermore, some research indicates that many teachers believe that a number of uncontrollable elements, like the course's structure and section size, may impact how well they teach, which may affect student assessments [6]. Therefore, the evaluations completed by students address only a few aspects of the course and teaching quality [7]. According to a limited number of studies that have surveyed students, the literature supports the significance of professor evaluations. It indicates that students are capable of assessing their professors. However, these studies also demonstrate that students are not overly hopeful about the overall importance that faculty and administrators place on student opinions [8]. Without serious worries about possible repercussions, students are generally willing to participate in assessments and provide feedback [9].

This paradox demonstrates that while students regularly engage with their instructors, making their opinions and evaluations significant and applicable [8], if students are unaware of the importance of their assessment and fill out the forms carelessly and without consideration, this may affect the validity of the SETE [9]. Therefore, it's critical to comprehend how students view the assessment of instructional effectiveness.

SETE studies rarely discuss students' perceptions, which typically concentrate on the features, validity, and reliability of the questionnaire used, factors influencing student ratings, and appropriate assessment of teaching quality. Thus, this study investigates how Sultan Qaboos University's College of Nursing students view teacher evaluations.

Materials & Methods

Design and setting(s)

This cross-sectional descriptive study was conducted during the summer and fall semesters of the 2023-2024 academic year. Undergraduate students from Sultan Qaboos University's College of Nursing participated in the study. To meet the inclusion requirements, participants had to be currently enrolled undergraduate students who had finished at least one semester of nursing school at Sultan Qaboos University. The study did not include officially postponed or enrolled students in the foundation program.

Participants and sampling

The sample included current undergraduate nursing students at Sultan Qaboos University's College of Nursing in Oman. The students were chosen using convenience sampling techniques. Slovin's formula ($n = N / (1 + Ne^2)$), where (N) is the population size and (e) is the margin of error, was used to determine the sample size.

Based on the college's annual active student records, the population was estimated at 450 students. Using a 5% margin of error and a 95% confidence level, the required sample size was determined to be 208 students. To account for potential attrition (estimated at 10%), the adjusted target sample size was increased to 230 participants. Given an anticipated response rate of 50%, all eligible students were invited to participate in the study by completing the questionnaire.

Tools/Instruments

Part 1**.** Demographic data: Demographic data collected included age, gender, year of study, Grade Point Average (GPA) range, study program, and participation in extracurricular activities.

Part 2**.** Evaluation of teaching effectiveness: The primary tool was intended to capture students' opinions of Student Evaluation of Teaching Effectiveness. The questionnaire had twenty items, including thirteen Yes/No questions, two items with a linear scale from 10% to 100%, and five multiple-choice questions.

Although self-report questionnaires are standard tools for data collection, they are subject to response bias. To mitigate this, responses were anonymized, and participants were assured that their evaluations would not affect their academic standing or grades.

The validity and reliability of the questionnaire were thoroughly evaluated. Ten knowledgeable faculty members with expertise in educational research were asked for their opinions to assess the validity of the qualitative content. The experts' level of agreement was measured using the Content Validity Index (CVI). Since the Kuder-Richardson-20 (KR-20) formula is suitable for instruments with dichotomous items, it was used for reliability assessment. Reliability was tested with a sample of 30 participants. The CVI and KR-20 reliability results are as follows:

Importance

The CVI values for individual item importance meet the acceptable criterion of 0.80, indicating that the experts consider each item necessary.

[Downloaded from edujournal.zums.ac.ir on 2026-11-28]

[DOI: 10.61186/edcj.18.1.115]

Clarity

Every item's CVI value for clarity likewise satisfies the acceptable standard of 0.80, indicating that the items are intelligible and clear.

Suitability

Similarly, every item's suitability CVI value falls within the acceptable range of 0.80, suggesting that the items are appropriate for their intended use.

The Kuder-Richardson-20 (KR-20) formula assesses the internal consistency reliability of scales with dichotomous items. Our scale's KR-20 coefficient value is 0.99.

Data collection methods

The questionnaire was distributed online. Course coordinators for the core courses were contacted to obtain a list of registered students' email addresses. The questionnaire and an informed consent form were sent to the student's emails. The Principal Investigator (PI) also reminded students to complete the survey during class.

Data analysis

Data were analyzed using SPSS/v29 software (SPSS Inc., Chicago, IL, USA). The demographic features of the sample and the items about teaching effectiveness were described using frequency analysis (means, median, percentiles, and standard deviations). Correlational analysis was performed to find correlations between demographic characteristics and SETE levels.

Results

The study disclosed key findings regarding participants' demographics and their opinions regarding teacher evaluations. The average age of participants was 21.58 years (SD = 2.82), and their GPA was 3.03 (SD = 0.42). Most participants (72.3%) were female, and 80.2% had ordinary academic standing. Additionally, 17.5% of participants were in their seventh semester (**Table 1**).

Table 1. Sociodemographic characteristics of the participants (n = 234).

Variable	Categories	Mean ± SD	
Age (year)		21.58 ± 2.82	
Academic GPA (out of 4 points)		3.03 ± 0.42	
Gender		**%**	**n**
	Male	27.7	64
	Female	72.3	167
Academic status	On probation	1.4	3
	Normal	80.2	170
	Extended	18.4	39
Academic semester	Semester 2	11.6	27
	Semester 3	11.6	27
	Semester 4	12.8	30
	Semester 5	12.4	29
	Semester 6	11.5	27
	Semester 7	10.7	25
	Semester 8	11.5	27
	Semester 9	17.5	41

Abbreviations: n, number of participants; SD, standard deviation; GPA, grade point average; %, percentage.

The survey responses shed light on reasons for writing untrue comments in evaluations (**Table 2**). A small percentage of respondents (0.4% and 0.9%) reported knowing someone who wrote false comments to harm or because they disliked a teacher. In contrast, 22.7% and 20.5% indicated knowing someone who wrote false comments to protect a teacher or because of personal liking. Regarding the frequency of teacher evaluations, 17.5% of respondents preferred evaluations for every class each semester, whereas 67.9% preferred evaluations once per semester. Just 9.0% of respondents supported less frequent evaluations, like annual ones. In comparison, smaller percentages suggested evaluations only when issues arise (3.4%) or never (2.1%). Notably, none of the participants favored three-year evaluations.

Additionally, participants talked about their personal experiences of posting false comments (**Table 2**). A sizable portion acknowledged that they had made false remarks because they liked the teacher (50.9%) or to defend the teacher (42.3%). Fewer respondents believed they had written derogatory comments about a teacher because they didn't like them (50.9%) or to hurt the teacher (9.8%).

[Downloaded from edujournal.zums.ac.ir on 2026-11-28]

[DOI: 10.61186/edcj.18.1.115]

Table 2. Students' perspectives and behaviors related to writing untrue comments in evaluations and their opinions on teacher evaluation frequency

	Items	Choices	%	n
1	Do you personally know of a student who wrote something untrue in the written comments because:	They wanted to hurt a teacher	0.4	1
		They disliked a teacher	7.7	18
		Both	66.2	155
		No	25.2	59
2	Do you personally know of a student who wrote something untrue in the written comments because:	They wanted to protect a teacher	0.9	2
		They liked a teacher	22.7	53
		Both	48.5	113
		No	27.5	64
3	How often do you think that teachers should be evaluated?	Every class, every semester	17.5	41
		In one class every semester	67.9	159
		Once a year	9.0	21
		Once every three years	0	0
		Only if there is a problem	3.4	8
		Never	2.1	5
4	Have you written something untrue in the written comments because:	You wanted to protect a teacher	20.5	48
		You liked a teacher	42.3	99
		Both	0	0
		No	37.2	87
5	Have you written something untrue in the written comments because:	You wanted to hurt a teacher	9.8	23
		You disliked a teacher	50.9	119
		Both	0	0
		No	38.9	91

Abbreviations: n, number of participants; SD, standard deviation; GPA, grade point average; %, percentage.

Table 3 summarizes opinions regarding the integrity and correctness of assessments. On a scale of 0 to 100, respondents gave assessments an average perceived accuracy of 50.78 (SD = 27.33), indicating a moderate confidence level in their dependability.

They estimated that, on average, 47.24% of evaluations contained false information (SD = 23.05). A notable proportion of participants reported knowing students who gave evaluations higher than deserved (37.2%) or lower than deserved (35.6%). Additionally, 21.8% admitted to inflating evaluation scores, while 19.2% admitted to giving undeservedly low scores.

The assessment procedure itself uncovered conflicting opinions. While 48.3% of respondents thought administrators reviewed evaluation comments, the majority (55.1%) thought teachers read them. Most respondents (84.6%) favored student assessments of teachers, and 69.8% believed these assessments could aid in teachers' development. There appears to be a gap between the administrative use of evaluations and their perceived value, as only 39.7% of respondents favored using them to decide on teaching assignments. The results show the potential for survey design improvement to better capture student feedback, even though 52.6% of respondents felt the evaluation questions allowed them to express their genuine opinions.

There were also apparent ethical issues. Just 13.7% of respondents said they knew of a teacher who was fired for receiving subpar reviews. Meanwhile, 76.1% of respondents thought it was cheating to write false information in evaluations, and 69.7% thought it was cheating to give unjustly high or low scores. These findings highlight the ethical implications of inaccurate feedback and its potential consequences for teachers.

Statistical analyses (**Table 3**) showed no significant correlations between demographic factors—such as GPA, academic status, and class standing—and perceptions of evaluation accuracy or false information. For instance, GPA had a weak negative correlation with evaluation accuracy ($r = -0.031$, $p = 0.749$) and a weak positive correlation with false information ($r = 0.013$, $p = 0.891$). Similarly, chi-square tests found no significant association between sex and responses to questionnaire items ($p > 0.05$).

[Downloaded from edujournal.zums.ac.ir on 2026-11-28]

[DOI: 10.61186/edcj.18.1.115]

Table 3: The students' insights into the perceived accuracy of evaluations in assessing instructor teaching performance

	Items	Minimum	Maximum	Mean (SD)
6	If you had to give a score from 0 to 100 expressing how accurate these evaluations are in expressing how well an instructor teaches, what score would you give?	0.00	100.00	50.78 (27.33)
7	If you had to guess what percent of all evaluations submitted by students contained scores or information that the student knew was untrue, what percent of all evaluations do you believe that would be?	0.00	100.00	47.24 (23.05)

	Items		%	n
8	Do you personally know a student who has given a teacher a higher-than-deserved evaluation for any reason?	No	62.8	147
		Yes	37.2	87
9	Do you personally know a student who has given a teacher a lower-than-deserved evaluation for any reason?	No	64.4	150
		Yes	35.6	83
10	Have you given a teacher a higher-than-deserved evaluation for any reason?	No	78.2	183
		Yes	21.8	51
11	Have you given a teacher a lower-than-deserved evaluation for any reason?	No	80.8	189
		Yes	19.2	45
12	Do you think that the comments written on the evaluations are read by the teacher?	No	44.9	105
		Yes	55.1	129
13	Do you think that the comments written on the evaluations are read by administrators?	No	51.7	121
		Yes	48.3	113
14	Do you think that students should evaluate their teachers?	No	15.4	36
		Yes	84.6	198
15	Do you think that teachers who are evaluated become better teachers?	No	30.2	70
		Yes	69.8	162
16	Do you think that administrators use evaluations when making teaching assignments?	No	60.3	141
		Yes	39.7	93
17	Do you think that the questions on the evaluation allow you to express what you want to say on the evaluations?	No	52.6	123
		Yes	47.4	111
18	Do you personally know of a teacher who was let go because of bad evaluations?	No	86.3	202
		Yes	13.7	32
19	Do you think that giving an evaluation higher or lower than what a teacher deserves is a form of cheating?	No	30.3	71
		Yes	69.7	163
20	Do you think that writing something untrue on an evaluation about a teacher is a form of cheating?	No	23.9	56
		Yes	76.1	178

Abbreviations: n, number of participants; SD, standard deviation; GPA, grade point average; %, percentage.

Discussion

This study examines various aspects of student evaluations of instructors, focusing on participants' perceptions, behaviors, and preferences. The GPA of the students involved sheds light on their academic performance. Additionally, with an average age of 21.58 years, a predominance of female students, and a significant representation of students from Semester 7, the participant profile reflects diverse academic backgrounds.

The survey findings reveal differing motivations behind submitting inaccurate comments in evaluations. While a small number of respondents reported malicious intent, a larger group indicated their main motivations were to protect or show appreciation for their instructors. These results highlight the complexity of student perceptions and demonstrate how personal biases and relationships can impact evaluation feedback. According to a study by Carless and Winstone, students may give more favorable feedback because they want to protect their instructors or express their gratitude rather than provide accurate assessments [10]. Furthermore, the study of Esarey and Valdes showed that these biases can lead to both overly optimistic and pessimistic feedback, ultimately affecting the overall validity of evaluations [11].

Most respondents preferred evaluations to take place once per semester, indicating a desire for more consistent feedback mechanisms. This inclination implies that more frequent assessments support ongoing enhancements to the caliber of student instruction and learning outcomes. Nonetheless, a minority of participants advocated for less frequent assessments, indicating the range of viewpoints on this matter. Ethical questions concerning the integrity and authenticity of evaluation feedback are raised by the admission of making false comments, especially to

[Downloaded from edujournal.zums.ac.ir on 2026-11-28]

[DOI: 10.61186/edcj.18.1.115]

defend or support preferred instructors. The frequency of this behavior emphasizes the necessity of robust evaluation procedures that minimize biases and conflicts of interest while promoting sincere feedback.

Quinn [12] examined the impact of rating biases, such as personal relationships and favoritism, on the accuracy of instructor evaluations. The researchers emphasized the need for mechanisms to address these biases. Furthermore, Kreitzer and Sweet-Cushman's research from 2021 looked at how students' personal biases and instructors' actions can affect the authenticity and caliber of feedback in assessments. This study emphasized the significance of preserving objectivity in the assessment procedure [13].

Respondents expressed average confidence in the evaluations' capacity to accurately reflect instructor performance, indicating a moderate level of perceived accuracy. The significant standard deviation, however, suggests a great deal of variation in perceptions, which are probably impacted by personal experiences or evaluation biases. Additionally, the study discovered that participants thought a significant percentage of assessments might include false or deceptive information. To guarantee the validity and dependability of feedback, this study emphasizes the necessity of robust quality assurance procedures in the assessment. The subjective nature of evaluations and the possibility of bias resulting from interpersonal relationships or motivations are made clear by looking at individual experiences and behaviors related to evaluation feedback. This highlights how crucial it is to encourage accountability and openness in the assessment procedure to reduce potential biases.

Themes about teacher evaluations have been studied in earlier studies. According to Uttl et al., students typically prefer more frequent assessments, indicating a strong desire for consistent feedback [14]. On the other hand, Boysen noted a notable lack of confidence in the validity of assessments, suggesting that opinions may vary depending on the educational setting [15]. Different findings about the frequency of inaccurate information in assessments were reported by Baniasadi et al., highlighting the need for more studies to identify the causes of disparities in student feedback [16]. Furthermore, Clayson estimated that about 30 percent of evaluations contained false information and found that most students were aware of instances where assessments had been fabricated [17]. The validity of student evaluations in enhancing the performance of individual instructors, changing curricula, and creating comparative scales for faculty assessment is called into question by these findings.

Correlation analyses were performed to investigate possible relationships between variables like GPA, questionnaire answers, and demographic characteristics. The absence of statistically significant correlations indicates that there may not be many linear relationships between these variables. The complexity of evaluation results, which are impacted by a number of variables beyond straightforward linear correlations, must be taken into account.

Numerous studies have examined the connection between student grades and SETE. According to most of these studies, higher expected grades and evaluation ratings are positively correlated; students with higher GPAs also typically rate their teachers higher. However, Stroebe et al. discovered that student evaluations were not substantially impacted by teachers' grading practices [18]. Furthermore, other researchers found that students' perceptions of their teachers were not significantly affected by their knowledge of their grades [19, 20].

According to Berezvai et al., an increase of one point in a student's GPA was linked to an increase of 0.2 to 0.4 in the instructor's evaluation score. They also discovered a correlation between GPA and the average Student Evaluation of Teaching (SET) score for instructors [21]. Examining questionnaire responses and gender provides information about possible evaluation-related attitudes and action variations. The absence of significant correlations raises the possibility that attitudes toward evaluation feedback and preferences for evaluation frequency are not primarily influenced by gender.

Chávez and Mitchell found that students' language when evaluating male professors differed significantly from that used for female professors [22]. Additionally, a male instructor who taught the same course received higher ordinal scores than a female instructor, even for questions that were not specific to the instructor.

The current study has various restrictions. First, Sultan Qaboos University nursing students are the only ones included. Second, students from other disciplines should be surveyed for the study. Finally, the results may not represent students' opinions in different university colleges because they are unique to nursing students at Sultan Qaboos University's College of Nursing in Oman.

Conclusion

The study's conclusions help us better understand how students view teacher evaluations and point out important areas where these procedures need to be

[Downloaded from edujournal.zums.ac.ir on 2026-11-28]

[DOI: 10.61186/edcj.18.1.115]

improved. Addressing ethical issues, enhancing the genuineness of feedback, and adapting assessment techniques to students' varied preferences are all crucial. These actions are essential to maximize the efficiency and equity of teacher evaluation systems. Future studies should examine what influences students' attitudes and actions, how evaluation feedback affects instructional strategies, and how well interventions that enhance evaluation procedures work. Furthermore, longitudinal studies that monitor how students' attitudes and behaviors change over time may provide important new information about how teacher-student feedback interactions are changing.

Ethical considerations

The College of Nursing's ethics committee approved this study (ethical code: Ref. No. CON/DF/2023/8). Furthermore, the Assistant Dean of the College of Nursing's Undergraduate Office approved the data collection. The study's goals were explained in detail to the participants, who were also informed of their freedom to discontinue participation without incurring fees. They were assured that their decision to participate would not affect their academic standing or grades. The study maintained participant anonymity by refraining from collecting personally identifiable information, such as names or identification numbers. This procedure was clearly outlined in the informed consent form. Surveys were conducted using secure, password-protected electronic platforms and only accessible to the Principal Investigator (PI).

Artificial intelligence utilization for article writing

The authors state that artificial intelligence was utilized to paraphrase the paragraphs and edit the article's final version, particularly for English writing.

Acknowledgment

The authors thank the participants for participating sincerely during the research process.

Conflict of interest statement

The authors declare that no financial or other types of conflicts of interest exist.

Author contributions

MPK conceived the study. AA participated in data acquisition and sampling. NAS and ZSh participated in the study's design and performed the statistical analysis. ZSh interpreted the data. MPK obtained ethical clearance and permission to conduct the study. AA performed the tests. ZSh and MPK drafted or revised the article critically for important intellectual content. All authors read and approved the final manuscript.

Funding

Sultan Qaboos University provided financial support for this study.

Data availability statement

Data is available on request due to privacy/ethical restrictions.

References

1. Larry Crumbley D, Reichelt KJ. Teaching effectiveness, impression management, and dysfunctional behavior: Student evaluation of teaching control data. Quality Assurance in Education. 2009;17(4):377-92. [https://doi.org/10.1108/09684880910992340]
2. Oermann MH, Conklin JL, Rushton S, Bush MA. Student evaluations of teaching (SET): guidelines for their use. Nursing Forum. 2018;53(3):280-285. [https://doi.org/10.1111/nuf.12249]
3. Spencer KJ, Schmelkin LP. Student perspectives on teaching and its evaluation. Assessment & Evaluation in Higher Education. 2002;27(5):397-409. [https://doi.org/10.1080/0260293022000009285]
4. Otani K, Kim BJ, Cho JI. Student evaluation of teaching (SET) in higher education: how to use SET more effectively and efficiently in public affairs education. Journal of Public Affairs Education. 2012;18(3):531-44. [https://doi.org/10.1080/15236803.2012.12001698]
5. Jones J, Gaffney-Rhys R, Jones E. Handle with care! An exploration of the potential risks associated with the publication and summative usage of student evaluation of teaching (SET) results. Journal of Further and Higher Education. 2014;38(1):37-56. [https://doi.org/10.1080/0309877X.2012.699514]
6. Mohammed TA, Pandhiani SM. Analysis of factors affecting student evaluation of teaching effectiveness in Saudi higher education: the case of Jubail University College. American Journal of Educational Research. 2017;5(5):464-75. [https://doi.org/10.12691/education-5-5-2]
7. Uttl B, White CA, Gonzalez DW. Meta-analysis of faculty's teaching effectiveness: student evaluation of teaching ratings and student learning are not related. Studies in Educational Evaluation. 2017;54:22-42. [https://doi.org/10.1016/j.stueduc.2016.08.007]
8. Vilchez JA, Kruse J, Puffer M, Dudovitz RN. Teachers and school health leaders' perspectives on distance learning physical education during the COVID-

[Downloaded from edujournal.zums.ac.ir on 2026-11-28]

[DOI: 10.61186/edcj.18.1.115]

19 pandemic. Journal of School Health. 2021;91(7):541-9. [https://doi.org/10.1111/josh.13030]
9. Paufler NA, Sloat EF. Using standards to evaluate accountability policy in context: school administrator and teacher perceptions of a teacher evaluation system. Studies in Educational Evaluation. 2020;64:100806. [https://doi.org/10.1016/j.stueduc.2019.07.007]
10. Carless D, Winstone N. Teacher feedback literacy and its interplay with student feedback literacy. Teaching in Higher Education. 2023;28(1):150-63. [https://doi.org/10.1080/13562517.2020.1782372]
11. Esarey J, Valdes N. Unbiased, reliable, and valid student evaluations can still be unfair. Assessment & Evaluation in Higher Education. 2020;45(8):1106-20. [https://doi.org/10.1080/02602938.2020.1724875]
12. Quinn DM. Experimental evidence on teachers' racial bias in student evaluation: the role of grading scales. Educational Evaluation and Policy Analysis. 2020;42(3):375-92. [https://doi.org/10.3102/016237372093218]
13. Kreitzer RJ, Sweet-Cushman J. Evaluating student evaluations of teaching: a review of measurement and equity bias in SETs and recommendations for ethical reform. Journal of Academic Ethics. 2021;20(3–4):1–12. [https://doi.org/10.1007/s10805-021-09400-w]
14. Rollett W, Bijlsma H, Röhl S. Student feedback on teaching in schools: using student perceptions for the development of teaching and teachers. Springer, Cham. 2021: 237-56. [https://doi.org/10.1007/978-3-030-75150-0]
15. Boysen GA. The multidimensional nature of teaching and student evaluations: commentary on students' judgments of learning and teaching effectiveness. Journal of Applied Research in Memory and Cognition. 2020:9(2);152–156. [https://doi.org/10.1016/j.jarmac.2020.02.001]
16. Baniasadi A, Salehi K, Khodaie E, Bagheri Noaparast K, Izanloo B. Fairness in classroom assessment: a systematic review. The Asia-Pacific Education Researcher. 2022:1-9. [https://doi.org/10.1007/s40299-021-00636-z]
17. Clayson DE. A comprehensive critique of student evaluation of teaching: Critical perspectives on validity, reliability, and impartiality. Routledge; 2020. [https://doi.org/10.4324/9781003091462]
18. Stroebe W. Student evaluations of teaching encourage poor teaching and contribute to grade inflation: a theoretical and empirical analysis. Basic and Applied Social Psychology. 2020;42(4):276-94. [https://doi.org/10.1080/01973533.2020.1756817]
19. Carpenter SK, Witherby AE, Tauber SK. On students'(mis) judgments of learning and teaching effectiveness. Journal of Applied Research in Memory and Cognition. 2020;9(2):137-51. [https://doi.org/10.1016/j.jarmac.2019.12.009]
20. Chamberlin K, Yasué M, Chiang I-CA. The impact of grades on student motivation. Active Learning in Higher Education. 2023;24(2):109-24. [https://doi.org/10.1177/146978741881972]
21. Berezvai Z, Lukáts GD, Molontay R. Can professors buy better evaluation with lenient grading? The effect of grade inflation on student evaluation of teaching. Assessment & Evaluation in Higher Education. 2021;46(5):793-808. [https://doi.org/10.1080/02602938.2020.1821866]
22. Chávez K, Mitchell KM. Exploring bias in student evaluations: gender, race, and ethnicity. Political Science & Politics. 2020;53(2):270-4. [https://doi.org/10.1017/S1049096519001744]

[Downloaded from edujournal.zums.ac.ir on 2026-11-28]

[DOI: 10.61186/edcj.18.1.115]