

# Student evaluation of teaching (SET) in dental faculty: Is it influenced by grouping the students according to their averages?

Hooman Sadrhaghighi<sup>1</sup>  
Afsaneh Zarghami<sup>2</sup>  
Mahdi Irani<sup>3</sup>  
Mahsa Eskandarinezhad<sup>4</sup>  
Shahrzad Sadrhaghighi<sup>5</sup>  
Mahdi Parvaresh<sup>6</sup>

<sup>1</sup> Associate professor, Department of Orthodontics, Tabriz University of Medical Sciences, Faculty of Dentistry, Tabriz, Iran

<sup>2</sup> Assistant professor, Department of Orthodontics, Tabriz University of Medical Sciences, Faculty of Dentistry, Tabriz, Iran

<sup>3</sup> D.D.S., Tabriz University of Medical Sciences, Faculty of Dentistry, Tabriz, Iran.

<sup>4</sup> Associate professor, Department of Endodontics, Tabriz University of Medical Sciences, Faculty of Dentistry, Tabriz, Iran

<sup>5</sup> Prosthodontist, Islamic Azad University, Tehran, Iran

<sup>6</sup> D.D.S Tabriz University of Medical Sciences, Faculty of Dentistry, Tabriz, Iran

**Corresponding author:** Mahdi Parvaresh  
e-mail: mahdi.parvaresh89@gmail.com

## Abstract

**Purpose/objectives:** Student evaluation of teaching (SET) is a crucial method for assessing teaching performance in higher education. This study evaluated the impact of grouping students by their academic averages on SET results in dental education.

**Methods:** In this descriptive-analytical study, 109 dental students were divided into five groups based on their academic averages: 48 in the top 30%, 15 in the top 10%, 17 in the middle 10%, 16 in the lower 10%, and 48 randomly selected from the bottom 30%. These groups completed questionnaires evaluating five university professors across four aspects: teaching methods, academic proficiency, communication and behavior, and organization and discipline in theoretical lessons. Two weeks later, internal reliability was assessed by redistributing the questionnaires to 10% of participants. Statistical analyses were performed using Chi-Square and ANOVA with SPSS 18, considering results significant at  $P < 0.05$ .

**Results:** The internal reliability of the statements was moderate. Significant differences in mean scores among rater groups were observed for teaching methods ( $P = 0.001$ ), academic proficiency ( $P = 0.043$ ), and communication and behavior ( $P = 0.012$ ). However, no significant difference was found for organization and discipline ( $P = 0.855$ ).

**Conclusion:** Grouping students according to their averages may influence SET in certain teaching aspects. However, this approach only affects professors' organization and discipline evaluation.

**Keywords:** SET, Teaching, Education, Dental education.

## Introduction

Improvement of teaching performance is a significant priority in medical education institutes (1-3). Constructive feedback facilitates learning and growth, enables performance evaluation and targeted improvement, supports competence and intrinsic motivation, and contributes to knowledge-building and enhanced clinical skills (4). Evaluation of teaching includes different categories such as formalized self-appraisal of teaching, review of teaching portfolios, interviews with samples of students, student evaluation of teaching (SET), reviews by teaching experts, and assessment by alumni (5-7). Providing feedback by obtaining students' attitudes toward teaching effectiveness has been used worldwide (1, 3, 7, 8). It has been reported that 81% of the 36 U.S. dental

## Authors

Hooman Sadrhaghighi - Associate professor, Department of Orthodontics, Tabriz University of Medical Sciences, Faculty of Dentistry, Tabriz, Iran

Afsaneh Zarghami - Assistant professor, Department of Orthodontics, Tabriz University of Medical Sciences, Faculty of Dentistry, Tabriz, Iran

Mahdi Irani - D.D.S., Tabriz University of Medical Sciences, Faculty of Dentistry, Tabriz, Iran

Mahsa Eskandarinezhad - Associate professor, Department of Endodontics, Tabriz University of Medical Sciences, Faculty of Dentistry, Tabriz, Iran

Shahrzad Sadrhaghighi - Prosthodontist, Islamic Azad University, Tehran, Iran

Mahdi Parvaresh - D.D.S Tabriz University of Medical Sciences, Faculty of Dentistry, Tabriz, Iran



## License

This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Authors contributing to Oral and Implantology agree to publish their articles under the [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/), which allows third parties to copy and redistribute the material providing appropriate credit and a link to the license but does not allow to use the material for commercial purposes and to use the material if it has been remixed, transformed or built upon.

## How to Cite

R. Sadrhaghighi, A. Zarghami, M. Irani, M. Eskandarinezhad, S. Sadrhaghighi, Mahdi Parvaresh.

Student evaluation of teaching (SET) in dental faculty: Is it influenced by grouping the students according to their averages?.

Oral and Implantology  
Vol. 16 No. 3 (2024), 156-161.

schools used student evaluations, also known as student evaluation of teaching (SET) (9, 10). In this method, standardized questionnaires assess instructional quality, and the results are used both for summative and formative purposes (2, 3, 11-17). There is no consensus that student-oriented approaches for evaluation lead to notable improvements in teaching communication and behavior (6, 12, 14, 15, 18-21). However, investigations suggest that changes in teachers' self-perception and providing diagnostic evidence for administrative decision-making for promotions and tenure and informative sources for future students are expected (3, 7, 22-24). Although there is much debate about the proper use of SET, this survey's fast and easy process has made it a standard instrument in many higher education institutes (3, 15, 18). It is important to note that none of these assumptions have been confirmed by empirical research (25). Furthermore, there is a lack of information on the validity and reliability of questionnaires available, which directly affects outcomes and accountability of the results (7, 26). SET has been linked to higher grade point averages, as faculty often reduce workloads and grade more leniently to boost ratings. Skilled professors who challenge students may receive lower ratings, highlighting a disconnect between SET scores and teaching quality (24, 27). Moreover, teachers and institutions have shown some resistance to SET, primarily due to concerns over how feedback is interpreted (12, 13, 28). SET results are influenced more by individual professors than specific courses. Factors such as teacher popularity, communication skills, evaluator gender, workload, exam difficulty, class timing, and student grades can all impact evaluations (23, 26, 29, 30). Some researchers have explored the correlation between students' overall averages and SET, distinct from the potential influence of grade satisfaction when students know their final grades (31). Additionally, the impact of smaller student subsets on SET has been studied, though findings remain inconclusive (31-33). In this respect, we aimed to investigate the effect of aggregating dental students based on their averages on mean evaluation scores of SET for the teachers of the same theoretical course unit in dental faculty to provide evidence for better feedback and interpretation of this test.

## Methods and Materials

### Study Design

This descriptive-analytical study was conducted at a dental school to investigate the impact of student grouping on teaching evaluation practices.

### Participants

A total of 109 dental students participated in the study, comprising 52 males and 57 females, with a mean age of  $22 \pm 1.8$  years. Participants were categorized into five groups based on their academic performance from the previous semester, as obtained from the faculty's education office (Table 1). All participants had completed the relevant course unit.

### Data Collection

A questionnaire consisting of 25 items was developed using a 4-point Likert scale (never, sometimes, often, almost always) to assess students' satisfaction and perceptions across four domains of teaching quality:

1. Teaching Method: 10 statements
2. Academic Proficiency: 6 statements
3. Communication and Behavior: 5 statements
4. Organization and Discipline: 4 statements
5. An instructor distributed the questionnaire anonymously to the participants after the final exam.

### Reliability Assessment

To evaluate the reliability of the questionnaire, it was administered a second time to 10% of randomly selected participants two weeks after the initial distribution. The reliability was quantified using Cronbach's alpha, and the kappa statistic was calculated for each questionnaire item.

### Statistical Analysis

545 completed questionnaires were analyzed using nonparametric Chi-Square tests and ANOVA, employing SPSS version 18 (Microsoft, IL, USA). A p-value of  $<0.05$  was considered statistically significant.

## Results

Internal reliability calculating *Cronbach's alpha* coefficients for each statement ranged from 0.42 to 0.60. Using the *Pearson* Chi-Square test, the differences in frequency of Likert-scale answers (never, sometimes,

**Table 1.** Characteristics of groups of the study.

Class	Number of students	Group 1 Top 30%	Group 2 Top 10%	Group 3 Middle 10%	Group 4 Lower 10%	Group 5 Random 30%	Actual number of participants with regard to overlaps
1	50	15	5	5	5	15	29
2	55	17	5	6	6	17	42
3	52	16	5	6	5	16	38
Total	157	48	15	17	16	48	109

often, almost always) to each questionnaire statement were obtained among five groups of raters (Table 2). The mean evaluation scores in each aspect of teaching were compared among groups of students (Table 3). Results are categorized into four elements of the study.

**Table 2.** The differences in frequency of Likert-scale answers to each questionnaire statement among five groups of raters using the Pearson Chi-Square test.

Aspect of teaching	Statement number	P*
Teaching method	1	0.000
	2	0.000
	3	0.020
	4	0.003
	5	0.000
	6	0.000
	7	0.041
	8	0.001
	9	0.022
	10	0.544
Academic proficiency	11	0.223
	12	0.150
	13	0.335
	14	0.922
	15	0.027
	16	0.000
Communication and behavior	17	0.011
	18	0.000
	19	0.505
	20	0.000
	21	0.004
Organization and discipline	22	0.001
	23	0.000
	24	0.003
	25	0.027

\*P value; P<0.05 was considered statistically significant.

### Teaching method

The difference in mean scores regarding teaching methods was statistically significant among five groups of students ( $P=0.001$ ). Furthermore, the frequency of answers to nine out of ten statements differed significantly

( $P<0.05$ ). It seems that grouping students based on their averages may affect the teaching evaluation results in this respect.

### Academic proficiency

The mean scores given to academic proficiency were statistically significant among five groups of students ( $P=0.043$ ). However, the frequency of Likert scale answers to only two out of six statements significantly differed ( $P<0.05$ .) It can be concluded that grouping students based on their averages may affect the results of SET in this respect.

### Communication and behavior

The mean scores regarding communication and behavior were statistically significant among five groups of participants ( $P=0.012$ ), and the frequency of answers to four out of five statements was significantly different ( $P<0.05$ ). Grouping students based on their averages may affect the teaching evaluation results.

### Organization and discipline

The mean scores regarding communication and behavior were not statistically significant among the five groups ( $P=0.855$ ). However, the frequency of answers to four out of four statements was significantly different ( $P<0.05$ ). It seems that grouping the students based on their averages may not affect the results of SET in organization and discipline.

### Discussion

Many higher education institutes have developed various procedures and instruments for collecting, analyzing, and interpreting SET as the primary source of teaching quality evaluation (3, 12, 25). The key focus is the ease of data collection and the presentation and interpretation of results (5, 25). The study revealed significant effects of student grouping on teaching evaluations across several key areas. Mean scores for teaching methods, academic proficiency, and communication and behavior differed significantly among the five groups. Conversely, mean scores for organization and discipline did not show significant differences. These findings indicate that grouping students based on academic performance influences teaching evaluation results, particularly in teaching methods, academic proficiency, and communication and behavior. At the same time, organization and discipline appear to be less affected. No prior research exists on the effect of grouping students on SET results. Valadez et al. (20) concluded that grouping students positively impacts academic performance and specific aspects of creativity. Griffin et al. (34) found a moderate correlation between students' GPAs and SET results. In their review, Contantinou et al. (35) found that smaller classroom sizes correlate with higher SET rates for faculty. This trend may be attributed to the increased opportunities for interaction between faculty and students in smaller classes.

Royal et al. (36) highlighted the limitations of using SETs to evaluate courses and faculty in medical and healthcare programs. They noted that SETs designed for general higher education do not fit the medical curriculum, as questions about instructors and courses are often combined, making it difficult for students to differentiate between them. Furthermore, the involvement of multiple

**Table 3.** The mean evaluation scores in each aspect of teaching.

	Groups of participants	n	Mean	SD	95% CI		Min	Max	P*
					Lower	Upper			
Teaching Method	1	240	28.04	6.18	27.26	28.83	13	40	0.001
	2	75	29.80	4.59	28.74	30.86	15	40	
	3	85	26.26	5.61	25.05	27.47	14	40	
	4	80	28.29	7.47	26.62	29.95	13	40	
	5	240	29.20	6.18	28.42	29.99	10	40	
	Total	720	28.43	6.19	27.98	28.88	10	40	
Academic Proficiency	1	240	17.63	3.88	17.14	18.12	74	24	0.043
	2	75	18.73	3.37	17.96	19.51	7	24	
	3	85	17.00	3.87	16.17	17.83	8	24	
	4	80	17.26	4.16	16.34	18.19	9	24	
	5	240	17.79	3.84	17.30	18.28	6	24	
	Total	720	17.68	3.86	17.40	17.97	6	24	
Communication and behavior	1	240	15.21	3.34	14.78	15.63	5	20	0.012
	2	75	16.17	2.55	15.59	16.76	7	20	
	3	85	14.45	3.99	13.59	15.31	5	20	
	4	80	15.11	3.03	14.44	15.79	8	20	
	5	240	14.84	3.33	14.41	15.26	5	20	
	Total	720	15.08	3.34	14.84	15.33	5	20	
Organization and discipline	1	240	12.34	2.54	12.01	12.66	4	16	0.855
	2	75	12.39	2.07	11.91	12.86	4	16	
	3	85	12.54	2.57	11.99	13.09	4	16	
	4	80	12.08	2.91	11.43	12.72	6	16	
	5	240	12.33	2.78	11.98	12.68	4	16	
	Total	720	12.33	2.62	12.14	12.53	4	16	

\*P value; P<0.05 was considered statistically significant.

N: number; SD: standard deviation; CI: confidence interval; Min: minimum; Max: maximum.

instructors in medical courses complicates the evaluation process. In the present study, SETs were administered to evaluate the performance of five university professors teaching the same theoretical course unit. This approach addresses the limitations identified by Royal et al. (36) by eliminating potential confounding factors, such as individual characteristics of professors (e.g., attraction, sex, age), course difficulty, and student grades. Administering SETs in this manner depersonalizes the relationship between teachers and students, allowing for a more focused evaluation of teaching effectiveness. In line with the current study, Almakadma et al. (24) assessed medical students' perceptions of SET after final exam results were released, noting that students worry tutors aware of SET submissions might adjust exam difficulty or grading. However, this timing could introduce

bias, as students may rate courses based on exam difficulty, favoring faculty with more straightforward exams over actual teaching quality. Furthermore, simpler exams were shown to positively impact faculty evaluations. The internal reliability of the questionnaire in this study was moderate. In contrast, Emdadi S et al. (37) reported low internal reliability of their SET questionnaire. Nazir et al. (38) used SET to evaluate teaching effectiveness in a nonclinical dental course, found high internal reliability, and assessed the survey's psychometric properties. Grouping students according to their averages might mitigate some inherent biases and issues that Rowan et al. (19) highlights in their study. They reviewed two viewpoints on the effectiveness of the SET survey in assessing dental faculty. This approach could reduce tendencies like "straight-lining" or the inclination to

default to higher ratings without critical assessment by categorizing students based on their academic standing. Higher-performing students may provide more reflective and insightful evaluations, reducing monotonic response patterns and potentially enhancing the validity of SETs. Additionally, this method could counteract nonresponse bias by ensuring that feedback represents a diverse academic spectrum rather than skewing towards students with higher grades.

In contrast to the current study, Beran et al. (31) found that various student characteristics accounted for 7% of the total variance in SET scores. Spooen et al. (32) reported a small local effect size of 6.3% for students' grades—additionally, Smith et al. (39) noted statistically significant effects of both student and instructor sex on SET scores. However, these factors contributed to less than 1% of the explained variance.

This study adds to the literature by examining the impact of grouping students by academic performance on SET results, potentially reducing biases related to exam difficulty and classroom dynamics. Limitations include its focus on a single course and institution, which may restrict generalizability. Future research should explore this approach across diverse course types and institutions to validate the consistency of bias reduction.

## Conclusion

1. SET aims to evaluate teaching methods, academic proficiency, communication, and behavior that may be affected by grouping the students according to their averages.
2. When evaluating the organization and discipline of professors, grouping the students based on their averages may not affect the SET results.

## Conflict of interest

None

## References

1. Zabaleta, F., The use and misuse of student evaluations of teaching. *Teaching in Higher Education*, 2007. 12(1): p. 55-76.
2. Yao, Y. and M.L. Grady, How do faculty make formative use of student evaluation feedback?: A multiple case study. *Journal of Personnel Evaluation in Education*, 2005. 18(2): p. 107.
3. Spooen, P., B. Brockx, and D. Mortelmans, On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 2013. 83(4): p. 598-642.
4. Orsini, C., et al., Common models and approaches for the clinical educator to plan effective feedback encounters. *Journal of Educational Evaluation for Health Professions*, 2022. 19.
5. Marsh, H.W., Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International journal of educational research*, 1987. 11(3): p. 253-388.
6. Pounder, J.S., Is student evaluation of teaching worthwhile? An analytical framework for answering the question. *Quality Assurance in Education*, 2007. 15(2): p. 178-191.
7. Clayson, D.E., Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature. *Journal of Marketing Education*, 2009. 31(1): p. 16-30.
8. Shevlin, M., et al., The validity of student evaluation of teaching in higher education: love me, love my lectures? *Assessment & Evaluation in Higher Education*, 2000. 25(4): p. 397-405.
9. Grillo, A.C., et al., Student evaluations of teaching: dental and dental hygiene students' and faculty members' perspectives. *Journal of Dental Education*, 2016. 80(4): p. 439-451.
10. Jahangiri, L., et al., Assessment of teaching effectiveness in US dental schools and the value of triangulation. *Journal of dental education*, 2008. 72(6): p. 707-718.
11. Amini, M. and M. Honardar, The view of faculties and medical students about evaluation of faculty teaching experiences. *koomesh*, 2008. 9(3): p. 171-178.
12. Ory, J.C., Faculty thoughts and concerns about student ratings. *New directions for teaching and learning*, 2001. 2001(87): p. 3-15.
13. Penny, A.R. and R. Coe, Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research*, 2004. 74(2): p. 215-253.
14. Najafipour, S. and M. Amini, A Survey of teachersview points of Jahrom Medical School Towards Teachers Evaluation By Student. *Iranian Journal of Medical Education*, 2002. 2: p. 41-42.
15. Marzabadi, A., et al., Comparison between academic staffs and students' views about the teacher's evaluation form in theoretical lessons. *Education strategies in medical sciences*, 2013. 6(3): p. 135-139.
16. Nemec II, E.C., et al., Development of valid and reliable tools for student evaluation of teaching. *Currents in Pharmacy Teaching and Learning*, 2018. 10(5): p. 549-557.
17. Sadrhaghighi, H., et al., Esthetic perception of smile components by orthodontists, general dentists, dental students, artists, and laypersons. *Journal of investigative and clinical dentistry*, 2017. 8(4): p. e12235.
18. Spooen, P., D. Mortelmans, and J. Denekens, Student evaluation of teaching quality in higher education: development of an instrument based on 10 Likert-scales. *Assessment & Evaluation in Higher Education*, 2007. 32(6): p. 667-679.
19. Rowan, S., et al., Should student evaluation of teaching play a significant role in the formal assessment of dental faculty? Two viewpoints: Viewpoint 1: Formal faculty assessment should include student evaluation of teaching and viewpoint 2: Student evaluation of teaching should not be part of formal faculty assessment. *Journal of dental education*, 2017. 81(11): p. 1362-1372.
20. Valadez, D., et al., Evaluation of the effects of grouping high capacity students in academic achievement and creativity. *Sustainability*, 2020. 12(11): p. 4513.
21. Sadrhaghighi, A.H., et al., Esthetic preferences of laypersons of different cultures and races with regard to smile attractiveness. *Indian Journal of Dental Research*, 2017. 28(2): p. 156-161.
22. Chen, Y. and L.B. Hoshower, Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment & evaluation in higher education*, 2003. 28(1): p. 71-88.
23. Vahabi, A., et al., Factors affecting teacher evaluation scores: the studentsviewpoints of Kurdistan University of Medical Sciences. *Iranian Journal of Medical Education*, 2015. 15: p. 111-121.
24. Almakadma, A.S., et al., Perceptions and attitudes of medical students towards student evaluation of teaching: A cross-sectional study. *Medical Education Online*, 2023. 28(1): p. 2220175.
25. Penny, A.R., Changing the agenda for research into students' views about university teaching: Four shortcomings of SRT research. *Teaching in higher education*, 2003. 8(3): p. 399-411.
26. Richardson, J.T.E., Instruments for obtaining student feedback: A review of the literature. *Assessment & Evaluation in Higher Education*, 2005. 30(4): p. 387-415.
27. Stroebe, W., Student evaluations of teaching encourages poor teaching and contributes to grade inflation: A theoretical and empirical analysis. *Basic and applied social psychology*, 2020. 42(4): p. 276-294.
28. Arthur, L., From performativity to professionalism: lecturers' responses to student feedback. *Teaching in Higher Education*, 2009. 14(4): p. 441-454.
29. Charkazi, A., et al., The Viewpoint of Students about The

- Influential Factors of Teacher's Performance in Theoretical Courses in Golestan University of Medical Sciences' Students. *Journal of Research Development in Nursing and Midwifery*, 2013. 9(2): p. 56-64.
30. Vakili, A., S. HajiAghajani, and A. Rashidy-Pour, An investigation of factors influencing student evaluation of teacher performance: A comprehensive study in Semnan University of Medical Sciences. *Koomesh*, 2011. 12(2): p. 93-103.
  31. Beran, T. and C. Violato, Ratings of university teacher instruction: How much do student and course characteristics really matter? *Assessment & Evaluation in Higher Education*, 2005. 30(6): p. 593-601.
  32. Spooren, P., On the credibility of the judge: A cross-classified multilevel analysis on students' evaluation of teaching. *Studies in educational evaluation*, 2010. 36(4): p. 121-131.
  33. Kuzmanovic, M., et al., A new approach to evaluation of university teaching considering heterogeneity of students' preferences. *Higher Education*, 2013. 66: p. 153-171.
  34. Griffin, T.J., et al., Correlation between grade point averages and student evaluation of teaching scores: Taking a closer look. *Assessment & Evaluation in Higher Education*, 2014. 39(3): p. 339-348.
  35. Constantinou, C. and M. Wijnen-Meijer, Student evaluations of teaching and the development of a comprehensive measure of teaching effectiveness for medical schools. *BMC Medical Education*, 2022. 22(1): p. 113.
  36. Royal, K.D., et al., Psychometric validation of a medical and health professions course evaluation questionnaire. *American J Educ Res*, 2018. 6(1): p. 38-42.
  37. Emdadi S, A.F., Soltanian AR, Imani B, Maghsoud A, Shojaeei S, A Study of Reliability and Validity of the Teacher Evaluation Form and Factors Affecting Student's Evaluation of Teachers. *Strides in Development of Medical Education*, 2013. 10(1): p. 8.
  38. Nazir, M., et al., Use of student evaluation of teaching (SET) survey to evaluate effectiveness of teaching in a leadership course among dental students over three years. *The Scientific World Journal*, 2020. 2020(1): p. 6436102.
  39. Smith, S.W., et al., The influence of student sex and instructor sex on student ratings of instructors: Results from a college of communication. *Women's Studies in Communication*, 2007. 30(1): p. 64-77.