Research Article

Self-assessment and Teacher-Assessment During an OSCE in Undergraduate Dental Students-Application of a Visual Metaphor

Gerhard Schmalz [1, *], Deborah Kreher [1], Dirk Ziebolz [1], Maria Strauss [2], Stefan Buechi [3]

[1] Department of Cariology, Endodontology and Periodontology, Leipzig University, Leipzig, Germany
[2] Department of Psychiatry and Psychotherapy, Leipzig University, Leipzig, Germany
[3] MediX Gruppenpraxis, Zürich, Switzerland

[*] *Corresponding author:* Department of Cariology, Endodontology and Periodontology, Leipzig University, Leipzig, Germany. Email: gerhard.schmalz@medizin.uni-leipzig.de

Abstract

Background: Reflection on both student and teacher perspectives is crucial for effective communication and professional relationships during education.

Objectives: This observational cohort study aimed to compare students' self-assessment with teacher assessments, as well as with estimated self-assessment and estimated teacher-assessment, using the pictorial representation of illness and self-measure (PRISM) during an objective structured clinical examination (OSCE). Additionally, it sought to compare self-assessment and teacher-assessment with OSCE scores.

Methods: Fourth-year dental students (n = 44) were included at the beginning of their clinical course. Three tasks were selected for the OSCE exams: Oral examination on a model (task 1), matrix placement (task 2), and endodontic radiograph evaluation (task 3). Objective structured clinical examination scores were rated by an independent rater. Students and one of three calibrated teachers used PRISM to evaluate their respective assessments independently and blinded from each other. The relationships between the different assessments were determined using the Pearson correlation coefficient.

Results: For task 1, a moderate correlation was found between students' self-assessment and estimated self-assessment (r = 0.44, P < 0.01). For task 2, moderate correlations were observed between self-assessment and teacher-assessment, estimated teacher-assessment and teacher-assessment, as well as between self-assessment and estimated self-assessment (P ≤ 0.01). For task 3, moderate correlations were found between self-assessment and teacher-assessment, and between self-assessment and estimated self-assessment (P < 0.01). A moderate negative correlation between self-assessment and the OSCE score was observed only for task 2 (r = -0.41, P = 0.01). Moderate negative correlations between teacher-assessment in PRISM and the OSCE score were found for all three tasks (P < 0.01).

Conclusions: Self-assessment and teacher-assessment using PRISM exhibited task-dependent correlations, while results for estimated assessments varied. PRISM may serve as a promising tool for feedback and discussion in the future, as it seems capable of highlighting different views and expectations in the teaching context. Further studies are needed to confirm these findings.

Keywords: Dental Education, PRISM, Visual Metaphor, Undergraduate Education, Self-Reflection, Communication

1. Background

Teaching in undergraduate dental education is challenging due to the vast array of competencies required for graduates to practice autonomously immediately after their studies (1). In this context, self-assessment and self-directed learning are critical skills for dental students and practitioners, making these topics essential in dental curricula (1, 2). Self-assessment has been shown to enhance motivation, learning, and academic performance (3). Recent systematic reviews emphasize the importance of self-assessment and strategies to support it by increasing attention to self-assessment and developing and testing relevant approaches (4, 5).

In addition to students' self-assessment, teacher and peer assessments are also crucial in dental education. Peer assessment has been identified as a valuable tool to

support the training progress of dental students ([6](#)). Specifically, in objective structured clinical exams (OSCE), the combination of self- and peer-assessment has been found to be beneficial for dental students ([7](#), [8](#)). Furthermore, understanding the relationship between teacher-assessment and students' self-assessment is of significant interest in both dental and overall medical education. A recent study from Japan indicated that while the assessments by students and teachers in medical bedside teaching were generally consistent, some students tended to over- or underestimate their performance ([9](#)). This variation in perceptions between teacher-assessment and self-assessment is crucial for providing substantive feedback in dental education ([10](#)). It also impacts the student-teacher relationship, a key element in creating a positive learning environment and facilitating academic progress ([11](#)). Thus, examining the relationship between students' self-assessment and teacher-assessment is valuable in the educational context.

In this context, various approaches can be used to perform and visualize self-assessment and teacher assessment. Tools such as rubrics, digital methods, or scanners have been reported to support self-assessment ([6](#), [9](#), [12](#), [13](#)). In addition to these strategies, a novel visual tool has recently been introduced to illustrate the learning progress of dental students. The pictorial representation of illness and self-measure (PRISM) is a visual metaphor that has already been used in dental education to visualize students' self-perceived competencies ([14](#)). Pictorial representation of illness and self-measure has been found to be superior to numeric scales for self-assessment ([15](#)), suggesting it could be a promising approach for depicting both self-assessment and teacher-assessment in a dental education context.

2. Objectives

In summary, this study was designed to visualize self-assessment and teacher-assessment within a dental education setting. To achieve this, different practical tasks within an OSCE were utilized. Additionally, students' estimated self-assessment from the teacher's perspective (how teachers would expect students to rate their own performance) and the estimated teacher-assessment from the students' perspective (how students would expect teachers to rate their performance) were evaluated using PRISM. The study aimed to compare students' self-assessment with teacher-assessment, as well as with estimated self-assessment and estimated teacher-assessment using

PRISM. Furthermore, self-assessment and teacher-assessment were compared with OSCE scores. It was hypothesized that self-assessment and teacher-assessment would correlate with each other.

Additionally, it was hypothesized that the estimated self-assessment (from the teacher's perspective) and the estimated teacher-assessment (from the students' perspective) would show limited correlations with actual self-assessment and teacher-assessment. The third hypothesis assumed that both self-assessment and teacher-assessment would correlate with the OSCE results, as reflected by the sum score.

3. Methods

3.1. Study Design

This observational cohort study included a group of fourth-year undergraduate dental students before the beginning of their first clinical patient course. The ethics committee of the Medical Faculty of the University of Leipzig, Germany, reviewed and approved the protocol (No: 117/20-ek). Written informed consent was obtained from all participants to ensure voluntary participation.

3.2. Participants

Prior to the beginning of the clinical patient treatment course in October 2022 for fourth-year dental students, an OSCE was conducted. The participants of this OSCE, a cohort of fourth-year undergraduate dental students, were invited to participate voluntarily and were subsequently included in the study. Only students who were undergoing the OSCE for the first time were considered for participation. No further inclusion or exclusion criteria were applied.

3.3. Objective Structured Clinical Examination Tasks

The OSCE comprised a series of tasks. For the current study, three tasks were selected ([Figure 1](#)). First, students performed an oral examination on a dental model featuring healthy, carious, and restored teeth, and were required to record the findings accurately. In the second task, students selected and placed a suitable matrix for an occlusal-distal defect. In the third task, students evaluated a dental radiograph of endodontically treated teeth to assess the adequacy of the endodontic therapy. After completing each OSCE task, students self-assessed their performance as good, moderate, or poor. Additionally, an independent rater assigned an OSCE score based on a set of criteria and points, using a
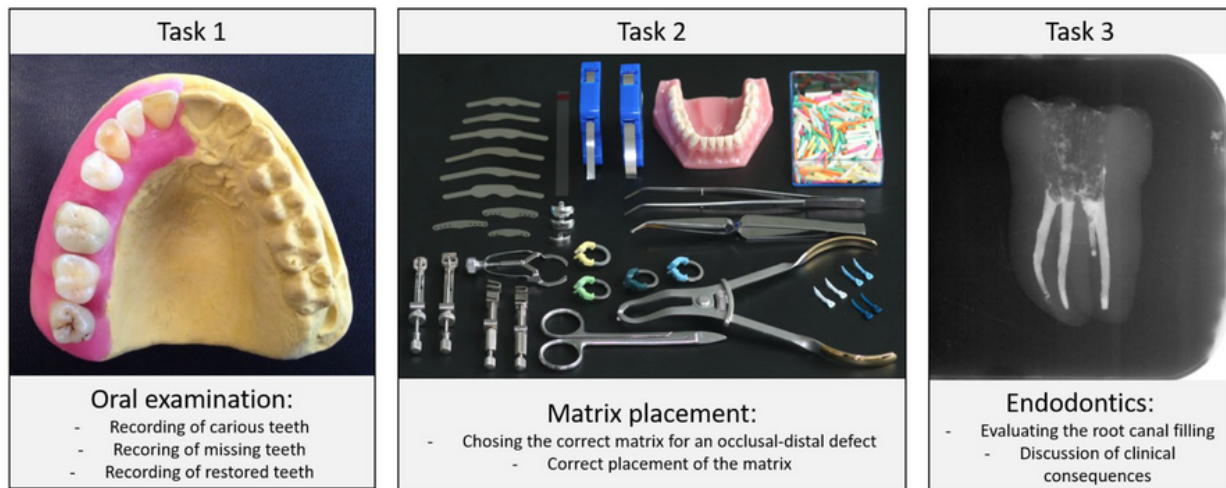
Figure 1. Tasks within the objective structured clinical examination (OSCE), which were used in the current study. Task 1 included a dental examination at a model. Task 2 was the correct choice and application of a matrix for an oral-distal cavity. Task 3 was the assessment and evaluation of an endodontic radiograph.

provided protocol to ensure consistent and reproducible ratings.

3.4. Pictorial Representation of Illness And Self-Measure (PRISM) Tasks

The general approach for applying PRISM in dental education has been previously described in detail ([14](#)). Originally developed in the field of psychology/psychosomatic medicine ([16](#)), PRISM consists of a white metal board (297 × 210 mm; "context") with a fixed yellow circle (7 cm in diameter) located in the bottom right corner ("subject"). Different colored discs (5 cm in diameter, "objects") are placed relative to the yellow circle to represent their relationships.

In the current study, PRISM was adapted to visualize both the students' self-assessment and the teacher's assessment, as well as their perceptions of each other's assessments. Both the teacher and the students used their own PRISM boards. After completing each task, teachers placed two objects on the board: Their evaluation of the student's performance (teacher-assessment) and their estimate of the students' self-assessment (how they thought students would rate their own performance) ([Figure 2](#)). Independently, students placed two objects: Their evaluation of their own performance (student's self-assessment) and their estimate of the teacher's assessment (how they thought teachers would rate their performance) ([Figure 3](#)).

Teachers and students performed the PRISM evaluation independently and without knowledge of each other's placements. PRISM evaluations were conducted for each of the three tasks separately. Distances between the centers of the objects and the subject were measured in centimeters using a template (ranging from 0 to 25 cm).

3.5. Study Flow

The flow of the current study is depicted in [Figure 4](#). Due to organizational reasons, the OSCE was conducted over three consecutive days. The entire group of 44 students was randomly assigned to the respective days. Students from each group completed all three OSCE tasks on the same day.

Students first underwent the OSCE tasks. After completing these tasks, both students and teachers performed the PRISM evaluation independently. Students then assessed their own performance (good, moderate, poor) and received their OSCE scores from an independent rater. Three calibrated and trained teachers conducted the PRISM evaluations. Calibration was achieved by performing repeated independent PRISM evaluations of constructed cases until there was at least 80% agreement among the values. To minimize potential teacher effects on the results, teachers rotated among the three tasks throughout the days. Experienced dentists acted as raters for the OSCE tasks and used standardized rating protocols for this purpose. The detailed study flow can be seen in [Figure 4](#).
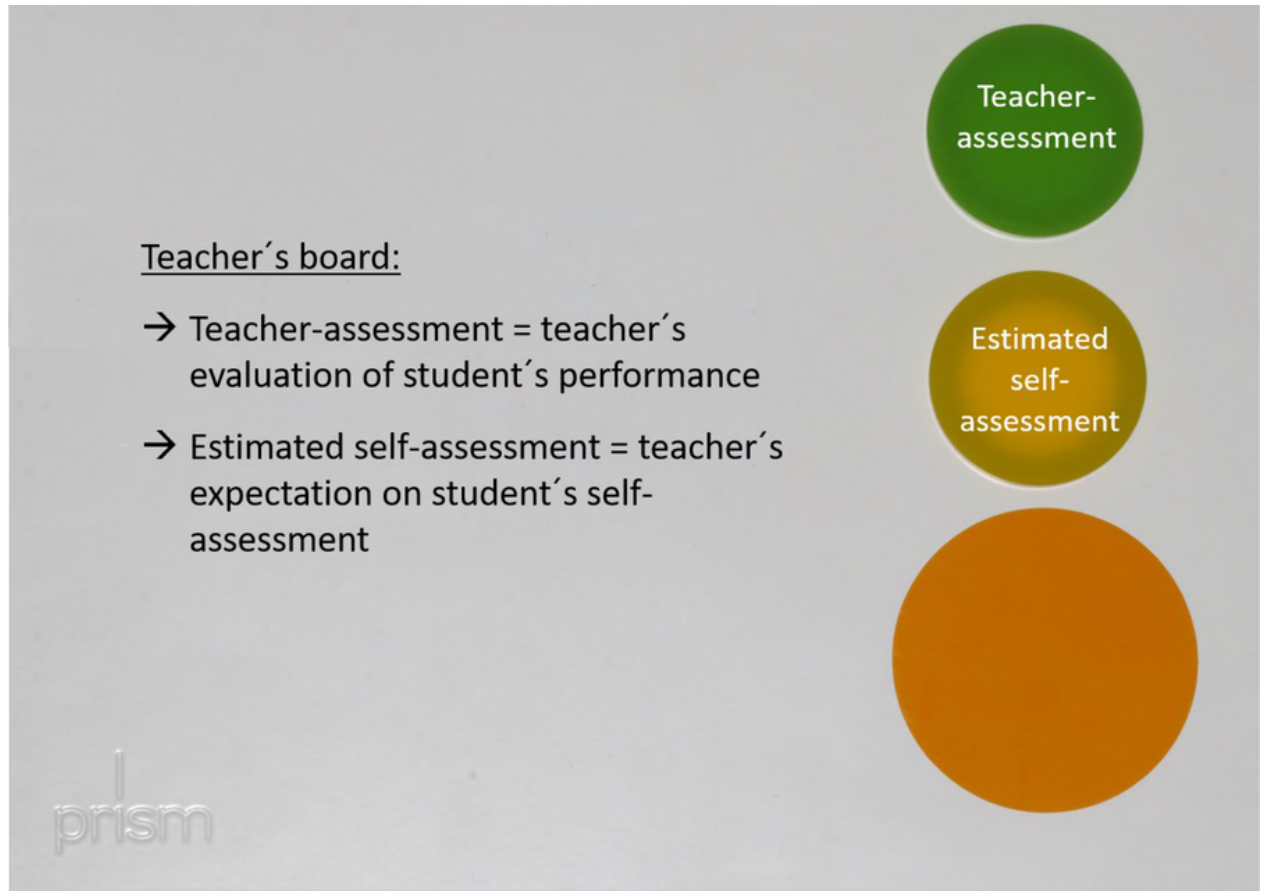
Figure 2. Principle of pictorial representation of illness and self-measure (PRISM) as used in the current study. The teacher and the student had an own PRISM board each. This figure shows the teacher board, whereby teacher-assessment and estimated self-assessment of the student were evaluated for each task.

3.6. Statistical Analysis

The basis for statistical analysis was the distance (in centimeters) between the objects representing the student's self-assessment, estimated teacher-assessment, teacher-assessment, and estimated self-assessment, and the subject circle at the bottom right-hand corner of the PRISM board. Statistical analysis was performed using SPSS for Windows, version 25.0 (SPSS Inc., U.S.A.).

First, normal distribution was assessed using the Shapiro-Wilk test. Correlations between the different assessments were evaluated using the Pearson correlation coefficient. The following interpretation was used: 0.1 - 0.39 = weak correlation, 0.4 - 0.69 = moderate correlation, 0.7 - 0.89 = strong correlation, and 0.9–1 = very strong correlation ([17]). A potential influence of the teacher on the different assessments was tested using linear regression. Two-sided significance testing was conducted for all values, with the significance level set at $P < 0.05$.

4. Results

4.1. Participants

In the current study, 44 fourth-year undergraduate dental students, with a mean age of 23.5 ± 2.6 years and comprising 64% female participants, were included. The OSCE results revealed that students achieved the lowest mean score in the first task, which involved an oral examination at a model. The OSCE results are presented in Table 1. The distances in PRISM for the different assessments and tasks are detailed in Appendix 1.
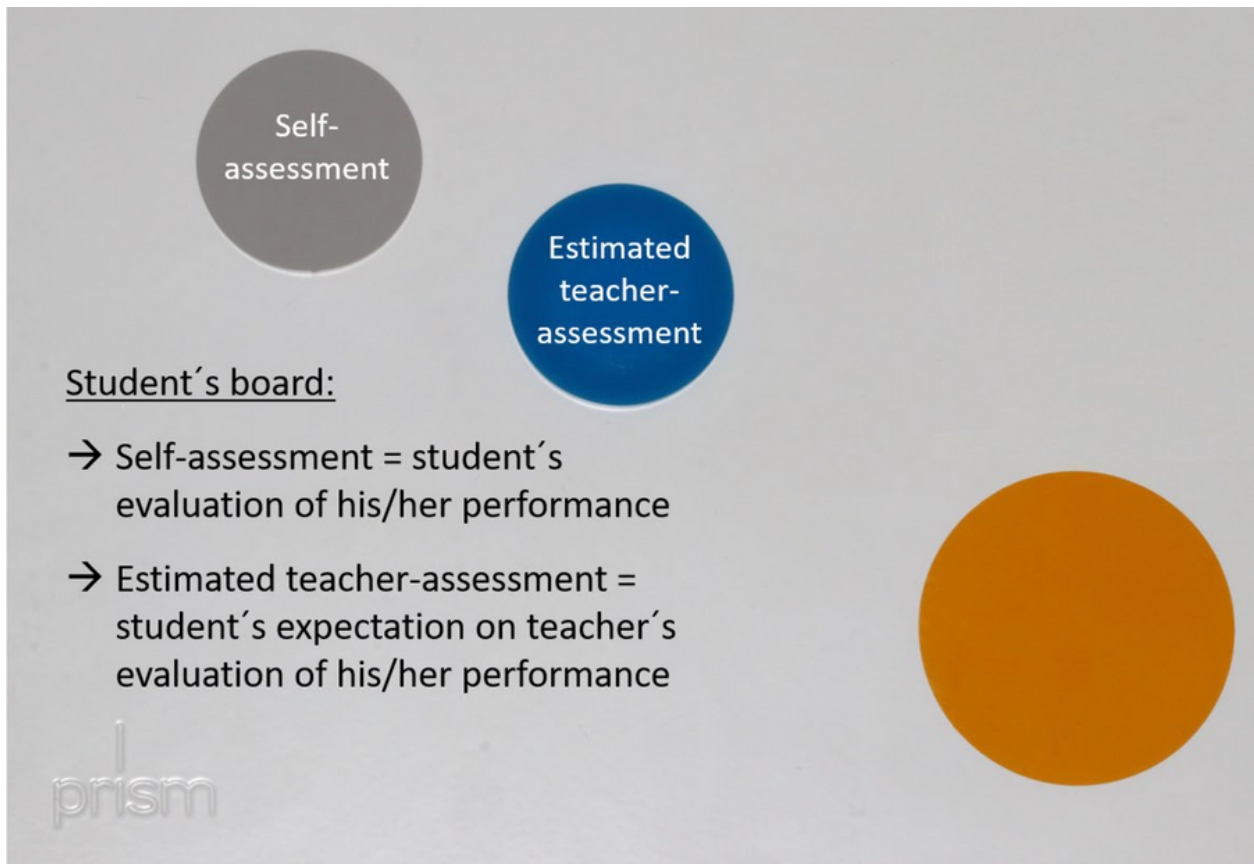
Figure 3. Principle of pictorial representation of illness and self-measure (PRISM) as used in the current study. The teacher and the student had an own PRISM board each. This figure shows the student board, whereby student´s self-assessment and estimated teacher-assessment were evaluated for each task.

4.2. Correlation Between Students' Self-Assessment and Teacher-Assessment Using Pictorial Representation Of Illness And Self-Measure (PRISM)

In Task 1, a moderate correlation between students' self-assessment and estimated self-assessment was found (r = 0.44, P < 0.01). In Task 2, moderate correlations were observed between self-assessment and teacher-assessment, estimated teacher-assessment and teacher-assessment, as well as between self-assessment and estimated self-assessment (P ≤ 0.01). In Task 3, a moderate correlation was found between self-assessment and teacher-assessment, as well as between self-assessment and estimated self-assessment (P < 0.01, Table 2).

4.3. Correlation Between Pictorial Representation of Illness And Self-Measure-Based Assessment and Objective

Structured Clinical Examination Results

The self-assessment in PRISM and the OSCE self-assessment (good, moderate, or poor) were significantly correlated across all three tasks (P < 0.01). Notably, in Task 3, a strong correlation was observed (r = 0.72, P < 0.01). A moderate negative correlation between self-assessment and the OSCE score was found only in Task 2 (r = -0.41, P = 0.01). For teacher-assessment in PRISM and the OSCE score, moderate negative correlations were found in all three tasks (P < 0.01; Table 2).

4.4. Influence of the Teacher

In the regression analysis, the teacher was found to be an influential factor on estimated self-assessment (β = 0.26, P < 0.01; Appendix 2). When using self-assessment and estimated teacher-assessment as the dependent
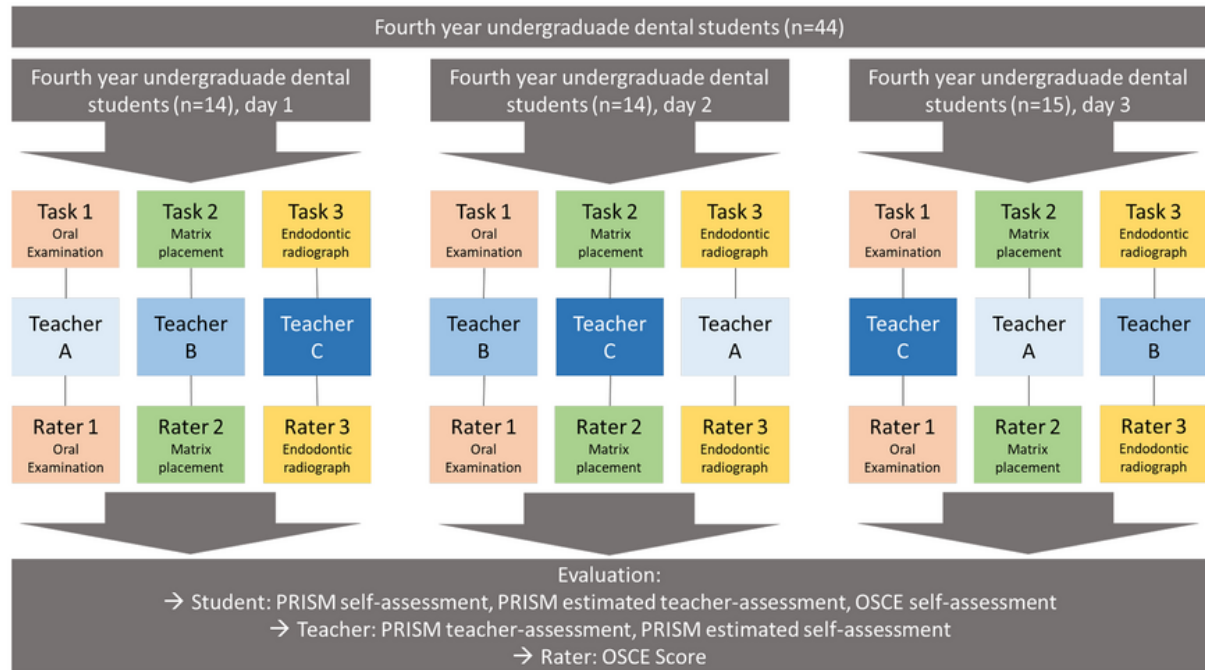
Figure 4. Study flow; at three different days, the randomly assigned students underwent all of the three objective structured clinical examination (OSCE) tasks and performed pictorial representation of illness and self-measure (PRISM) with a respective teacher. An independent rater provided the OSCE score for the respective task.

variables, no significant associations were found (Appendix 2).

5. Discussion

This current study used PRISM, a visual metaphor, to depict students' self-assessment and teacher-assessment in an OSCE setting. The underlying method originated from the field of psychology and was conceived to visualize suffering, and was subsequently applied to several other fields in psychology and medicine (18). Previous studies successfully modified the original PRISM task for application in dental education, showing reasonable reproducibility, sensitivity to changes, and a perceived superiority over numeric scales for self-reflection (14, 15). This methodological approach has now been applied in the context of the current study to test three different hypotheses.

First, it was hypothesized that self-assessment and teacher-assessment depicted with PRISM would correlate with each other. As the results show, a moderate positive correlation was found in this respect, although this was only observed for task 2 (matrix placement) and task 3 (endodontics). The OSCE scores

revealed that the average points for task 1 were the lowest among the three tasks, potentially indicating that task 1 (oral examination at a model) was the most challenging, which might have influenced the difference between students' self-assessment and the teacher-assessment provided by the respective teacher. A Japanese study, which compared self-assessment and teacher-assessment in medical interview performance using an assessment tool, showed good agreement between self-assessment and the teachers' ratings (9).

In contrast, another hospital-based study of medical graduates showed a wide range between self-assessment and assessment by an expert (19). Interestingly, this range increased in cases of inadequate performance (19). This could explain the lack of correlation in task 1, which appeared to be the most challenging task with the largest number of students with inappropriate results. Another study, which compared mini clinical evaluation exercises of clinical students with their supervisors in gynecology, internal medicine, pediatrics, psychiatry, and surgery, found correlations between student-assessment and teacher-assessment ranging from 0.29 to 0.51, depending on the task (20). The scale of correlation was similar to that in the current study

Table 1. Participants Characteristics and Objective Structured Clinical Examination (OSCE) Scores in the Included Patients [a]

Variables	Students (n = 44)
Age	23.46 ± 2.59
Gender (female)	64
Objective structured clinical examination scores	
Task 1 (examination)	14.3 ± 5.2
Task 2 (matrix placement)	19.7 ± 4.0
Task 3 (evaluation of endodontic radiograph)	16.1 ± 3.2
Objective structured clinical examination self-estimation task 1	
Good	7 (15.9)
Moderate	32 (72.7)
Poor	5 (11.4)
Objective structured clinical examination self-estimation task 2	
Good	12 (27.3)
Moderate	25 (56.8)
Poor	7 (15.9)
Objective structured clinical examination self-estimation task 3	
Good	13 (29.5)
Moderate	30 (68.2)
Poor	1 (2.3)

[a] Values are expressed as No. (%) or mean ± SD.

Table 2. Correlations Between Pictorial Representation of Illness And Self-Measure (PRISM) Based Self-Assessment, Teacher-Assessment and Estimated Assessments in the Three Different Tasks. Furthermore, Correlations Between PRISM Based Self-Assessment and Teacher-Assessment with the Objective Structured Clinical Examination (OSCE) Results in the Three Different Tasks Are Presented

Variables	Task 1		Task 2		Task 3	
	r	P-Value	r	P-Value	r	P-Value
Self-assessment vs. teacher-assessment	0.07	0.66	0.54	< 0.01	0.52	< 0.01
Estimated teacher-assessment vs. teacher-assessment	0.02	0.88	0.41	0.01	0.35	0.02
Self-assessment vs. estimated self-assessment	0.44	< 0.01	0.52	< 0.01	0.49	< 0.01
Self-assessment vs. objective structured clinical examination self-assessment	0.59	< 0.01	0.69	< 0.01	0.72	< 0.01
Self-assessment vs. objective structured clinical examination score	-0.16	0.30	-0.41	0.01	-0.39	0.01
Teacher-assessment vs. objective structured clinical examination score	-0.63	< 0.01	-0.53	< 0.01	-0.61	< 0.01

and supports the assumption that the match between self-assessment and teacher-assessment depends on the task and possibly on how challenging it is for the student. Besides the tasks on which students and teachers assess performance, several influential factors appear relevant in the context of the overlap or discrepancy between self-assessment and teacher-assessment. It has been reported that students' strategies to self-assess their performance are complex and influenced by both internal and external factors (21). Potential issues include the purpose of assessment, students' self-efficacy, the context in which self-assessment is performed, and gender (21-24). Therefore, this study also considered the estimated assessment of the teacher (estimated self-assessment and estimated

teacher-assessment) to gain a deeper understanding of the results, which led to the second hypothesis of this study.

This second hypothesis assumed that the estimated self-assessment (teacher perspective) and the estimated teacher-assessment (student perspective) would show limited correlations with the real self-assessment and teacher-assessment. Therefore, PRISM was used as a novel approach to depict the physical distance as a visual metaphor. For the relation between estimated teacher-assessment and real teacher-assessment, which was again task-dependent, almost weak or no correlations were found. In contrast, a moderate correlation was found between estimated self-assessment and the real self-assessment across all tasks.

This is interesting, as it suggests that students' estimates of how the teacher would evaluate their work differ from the teacher's perspective on the student's self-assessment. On the one hand, this finding supports a type of strategic self-assessment, as mentioned previously (21). Students appear to anticipate a certain evaluation from the teacher, which might influence their self-assessment. In this context, the allocation of roles between the teacher and the student, along with a related imbalance of power and potential sources of conflict, has been a topic in pedagogy for several decades (25, 26). This role constellation and the associated internal or social tensions might influence how students expect to be rated by the teacher, who are often seen more as adversaries than as supporters.

Based on the current study's findings, PRISM could serve as a basis to discuss the discrepancy between students' and teachers' ratings, potentially helping to resolve this perceived mismatch. A strength of PRISM is that the visual metaphor can facilitate discussions about the problem using the metal board, thereby transferring the subject of discussion. This allows students and teachers to perceive the problem more objectively, which might support their collaborative validation. Given the significant impact of a good student-teacher relationship in medical teaching (11), this might be a novel and promising approach. However, this study did not include a discussion of the results between students and teachers based on the PRISM findings; this is planned for a subsequent project to confirm the potential benefits of using PRISM in this respect. Additionally, regression analysis showed that the estimated self-assessment was somewhat teacher-dependent. Therefore, teachers should be aware of how and under what circumstances they assess their students.

The third and last hypothesis of this study assumed that both self-assessment and teacher-assessment would correlate with the OSCE results, as reflected by a sum score. This was primarily evaluated to assess whether PRISM would be suitable for depicting a situation in accordance with objective scoring criteria. The teacher-assessment showed, as expected, a moderate negative correlation with the OSCE score of an independent rater. Since the principle of PRISM indicates that a lower distance reflects a better result, the negative correlation is plausible. These results support previous findings that PRISM can quantify competencies (14). In addition, students' self-assessment in PRISM correlated with their assessment of the OSCE task as good, moderate, or poor, further supporting the potential of PRISM to facilitate self-reflection and learning progress. Interestingly,

students' self-assessment showed a task-dependent relationship with the OSCE score of an independent rater.

Thereby, no correlation (task 1), a weak negative correlation (task 3), and a moderate negative correlation (task 2) were found. These results indicate a mismatch between the students' self-assessment and their actual performance, supporting the discussion of the second hypothesis. In principle, the ability to self-reflect on competencies appears essential in medical education, as it is related to performance, skills, and empathy (1, 2, 27, 28). Age and study progress are relevant influential factors on the self-assessment abilities of students (29). The current sample of beginning fourth-year students appears to have some limitations in their self-assessment skills. This might be due to their limited experience in a clinical context, as they start treating real patients in the fourth year. The visualization of this discrepancy using PRISM could be helpful, especially since a previously performed study showed that PRISM was perceived as a good tool for self-reflection among fourth-year dental students (15).

5.1. Strengths and limitations

Overall, this study employs a promising approach by using PRISM as a visual metaphor to illustrate students' and teachers' perspectives in an OSCE setting. Objective structured clinical examinations have long been recognized as reliable and valid procedures in dental education (30). Despite their widespread use and evaluation, there remains a continued need for improvement (30-32). While the reliability and validity of OSCE procedures are subjects of debate, this study sets the scoring of the OSCE and its ability to depict students' competencies in relation to PRISM-based reflection. The results suggest that PRISM has the potential to enhance student-teacher communication, their relationship, and cooperation, which aligns with the idea of teachers acting more as coaches than judges (33, 34).

However, there are several limitations that need to be considered. Firstly, the sample consists of only 44 volunteers from a single center and was studied on a one-time basis. This sample size is too small to draw generalizable conclusions. Future studies should adopt longitudinal approaches with a larger sample size, possibly in a multicentric setting, to increase the robustness of the results. While the study execution appears robust, with measures such as independent OSCE raters, teacher rotation, independent and blinded PRISM tasks, varied tasks, and reproducible settings, this stringent setting limited the more practical application

of PRISM. It has been documented that self-assessment should be combined with feedback or discussion of results ([35]). This will be a necessary step for future research in this field. The chosen approach represents a novel use of PRISM and thus has a pilot character. Further studies are needed to provide more evidence regarding the effects of PRISM in dental and medical education.

5.2. Conclusions

Within the limitations of this study, PRISM successfully visualized students' self-assessment, the corresponding teacher-assessment, and their relationship to the OSCE results for fourth-year undergraduate dental students. The estimated teacher-assessment and the real teacher-assessment often differed, potentially reflecting individual perspectives on performance and varying expectations within the teaching context. Using PRISM to illustrate these conflicting views could provide a solid foundation for addressing problems, defining learning objectives, and thereby enhancing the quality of feedback in dental and medical education.

5.3. Lay Summary

This study examined how dental students assess their own performance compared to their teachers during practical exams, using a visual tool called the PRISM. Forty-four fourth-year dental students participated in OSCEs involving three tasks. The findings revealed that students' self-assessments generally aligned with teacher assessments, although variations were noted, particularly in more challenging tasks. The study suggests that using PRISM can enhance communication about performance between students and teachers, potentially improving educational outcomes in dental training.

5.4. Highlights

Enhancing self-assessment skills: The study emphasizes the importance of self-assessment in dental education, suggesting that integrating tools like PRISM can help students better understand their competencies and improve their learning outcomes.

Improving teacher-student communication: By visualizing assessments, PRISM fosters clearer communication between students and teachers, allowing for more effective feedback and discussions about performance.

Modernizing assessment methods: The findings advocate for the adoption of innovative assessment methods in medical education, highlighting the need to move beyond traditional numeric scales to more engaging and informative visual tools.

Supplementary Material

Supplementary material(s) is available here [To read supplementary materials, please refer to the journal website and open PDF/HTML].

Footnotes

References

1. Burrows RS. Understanding self-assessment in undergraduate dental education. *Br Dent J*. 2018;**224**(11):897-900. [PubMed ID: 29855594]. https://doi.org/10.1038/sj.bdj.2018.437.

2. Leisnert L. Self-Directed Learning, Teamwork, Holistic View and Oral Health. *Swed Dent J Suppl*. 2014;(235):9-65. [PubMed ID: 26685633].

3. Sharma R, Jain A, Gupta N, Garg S, Batta M, Dhir SK. Impact of self-assessment by students on their learning. *Int J Appl Basic Med Res*. 2016;**6**(3):226-9. [PubMed ID: 27563593]. [PubMed Central ID: PMC4979309]. https://doi.org/10.4103/2229-516X.186961.

4. Mays KA, Branch-Mays GL. A Systematic Review of the Use of Self-Assessment in Preclinical and Clinical Dental Education. *J Dent Educ*. 2016;**80**(8):902-13. [PubMed ID: 27480701].

5. Gadbury-Amyot CC, Woldt JL, Siruta-Austin KJ. Self-Assessment: A Review of the Literature and Pedagogical Strategies for Its Promotion in Dental Education. *J Dent Hyg*. 2015;**89**(6):357-64. [PubMed ID: 26684992].

6. Ellakany P, El Tantawi M, Al-Eraky DM, Aly NM, Al-Eraky MM. Self-assessment and peer-assessment of the psychomotor skills in preclinical prosthodontics: A prospective study. *Eur J Dent Educ*. 2023;**27**(3):601-9. [PubMed ID: 35920079]. https://doi.org/10.1111/eje.12846.

7. Quick KK. The Role of Self- and Peer Assessment in Dental Students' Reflective Practice Using Standardized Patient Encounters. *J Dent Educ*. 2016;**80**(8):924-9. [PubMed ID: 27480703].

8. Larsen T, Jeppe-Jensen D. The introduction and perception of an OSCE with an element of self- and peer-assessment. *Eur J Dent Educ*. 2008;**12**(1):2-7. [PubMed ID: 18257758]. https://doi.org/10.1111/j.1600-0579.2007.00449.x.

9. Tang Z, Shikama Y, Otani K. Comparison of student self-assessment and teacher assessment of medical interview performance during bedside learning. *Fukushima J Med Sci*. 2023;**69**(1):21-8. [PubMed ID: 36990787]. [PubMed Central ID: PMC10122972]. https://doi.org/10.5387/fms.2023-03.

10. Leung A, Fine P, Blizard R, Tonni I, Louca C. Teacher feedback and student learning: A quantitative study. *Eur J Dent Educ*. 2021;**25**(3):600-6. [PubMed ID: 33222381]. https://doi.org/10.1111/eje.12637.

11. Haidet P, Stein HF. The role of the student-teacher relationship in the formation of physicians. The hidden curriculum as process. *J Gen Intern Med*. 2006;**21 Suppl 1**(Suppl 1):S16-20. [PubMed ID: 16405704]. [PubMed Central ID: PMC1484835]. https://doi.org/10.1111/j.1525-1497.2006.00304.x.

12. Meglioli M, Mergoni G, Artioli F, Ghezzi B, Manfredi M, Macaluso GM, et al. A Novel Self-Assessment Method for Training Access Cavity on 3D Printed Endodontic Models. *Dent J (Basel)*. 2023;**11**(6). [PubMed ID: 37366675]. [PubMed Central ID: PMC10297574]. https://doi.org/10.3390/dj11060152.

13. Stoilov M, Trebess L, Klemmer M, Stark H, Enkling N, Kraus D. Comparison of Digital Self-Assessment Systems and Faculty Feedback for Tooth Preparation in a Preclinical Simulation. *Int J Environ Res Public Health*. 2021;**18**(24). [PubMed ID: 34948828]. [PubMed Central ID: PMC8703803]. https://doi.org/10.3390/ijerph182413218.

14. Schmalz G, Kullmann H, Sensky T, Kreher D, Haak R, Buchi S, et al. Pilot study to evaluate a novel measure of self-perceived competencies among dental students. *BMC Med Educ*. 2022;**22**(1):582. [PubMed ID: 35906646]. [PubMed Central ID: PMC9336052]. https://doi.org/10.1186/s12909-022-03663-6.

15. Schmalz G, Kreher D, Sensky T, Buchi S, Ziebolz D. Comparison of PRISM and numeric scale for self-assessment of learning progress during a clinical course in undergraduate dental students. *BMC Med Educ*. 2022;**22**(1):894. [PubMed ID: 36564754]. [PubMed Central ID: PMC9789606]. https://doi.org/10.1186/s12909-022-03967-7.

16. Buchi S, Sensky T, Sharpe L, Timberlake N. Graphic representation of illness: a novel method of measuring patients' perceptions of the impact of illness. *Psychother Psychosom*. 1998;**67**(4-5):222-5. [PubMed ID: 9693349]. https://doi.org/10.1159/000012284.

17. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth Analg*. 2018;**126**(5):1763-8. [PubMed ID: 29481436]. https://doi.org/10.1213/ANE.0000000000002864.

18. Sensky T, Buchi S. PRISM, a Novel Visual Metaphor Measuring Personally Salient Appraisals, Attitudes and Decision-Making: Qualitative Evidence Synthesis. *PLoS One*. 2016;**11**(5). e0156284. [PubMed ID: 27214024]. [PubMed Central ID: PMC4877057]. https://doi.org/10.1371/journal.pone.0156284.

19. Abadel FT, Hattab AS. How does the medical graduates' self-assessment of their clinical competency differ from experts' assessment? *BMC Med Educ*. 2013;**13**:24. [PubMed ID: 23402221]. [PubMed Central ID: PMC3576227]. https://doi.org/10.1186/1472-6920-13-24.

20. Berendonk C, Rogausch A, Gemperli A, Himmel W. Variability and dimensionality of students' and supervisors' mini-CEX scores in undergraduate medical clerkships - a multilevel factor analysis. *BMC Med Educ*. 2018;**18**(1):100. [PubMed ID: 29739387]. [PubMed Central ID: PMC5941409]. https://doi.org/10.1186/s12909-018-1207-1.

21. Sargeant J, Armson H, Chesluk B, Dornan T, Eva K, Holmboe E, et al. The processes and dimensions of informed self-assessment: a conceptual model. *Acad Med*. 2010;**85**(7):1212-20. [PubMed ID: 20375832]. https://doi.org/10.1097/ACM.0b013e3181d85a4e.

22. Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. *Acad Med*. 2005;**80**(10 Suppl):S46-54. [PubMed ID: 16199457]. https://doi.org/10.1097/00001888-200510001-00015.

23. Boor K, Scheele F, van der Vleuten CP, Teunissen PW, den Breejen EM, Scherpbier AJ. How undergraduate clinical learning climates differ: a multi-method case study. *Med Educ*. 2008;**42**(10):1029-36. [PubMed ID: 18823522]. https://doi.org/10.1111/j.1365-2923.2008.03149.x.

24. Riese A, Rappaport L, Alverson B, Park S, Rockney RM. Clinical Performance Evaluations of Third-Year Medical Students and Association With Student and Evaluator Gender. *Acad Med*. 2017;**92**(6):835-40. [PubMed ID: 28099178]. https://doi.org/10.1097/ACM.0000000000001565.

25. Jamieson DW, Thomas KW. Power and Conflict in the Student-Teacher Relationship. *J Appl Behav Sci*. 2016;**10**(3):321-36. https://doi.org/10.1177/002188637401000304.

26. Carless D. Differing perceptions in the feedback process. *Stud High Educ*. 2006;**31**(2):219-33. https://doi.org/10.1080/03075070600572132.

27. Paloniemi E, Mikkola I, Vatjus R, Jokelainen J, Timonen M, Hagnas M. Measures of empathy and the capacity for self-reflection in dental and medical students. *BMC Med Educ*. 2021;**21**(1):114. [PubMed ID: 33602204]. [PubMed Central ID: PMC7890956]. https://doi.org/10.1186/s12909-021-02549-3.

28. Ratelle JT, Bonnes SL, Wang AT, Mahapatra S, Schleck CD, Mandrekar JN, et al. Associations between teaching effectiveness and participant self-reflection in continuing medical education. *Med Teach*. 2017;**39**(7):697-703. [PubMed ID: 28301975]. https://doi.org/10.1080/0142159X.2017.1301655.

29. Achenbach J, Schafer T. Modelling the Effect of Age, Semester of Study and Its Interaction on Self-Reflection of Competencies in Medical Students. *Int J Environ Res Public Health*. 2022;**19**(15). [PubMed ID: 35954936]. [PubMed Central ID: PMC9367769]. https://doi.org/10.3390/ijerph19159579.

30. Brown G, Manogue M, Martin M. The validity and reliability of an OSCE in dentistry. *Eur J Dent Educ*. 1999;**3**(3):117-25. [PubMed ID: 10865346]. https://doi.org/10.1111/j.1600-0579.1999.tb00077.x.

31. Napankangas R, Karaharju-Suvanto T, Pyorala E, Harila V, Ollila P, Lahdesmaki R, et al. Can the results of the OSCE predict the results of clinical assessment in dental education? *Eur J Dent Educ*. 2016;**20**(1):3-8. [PubMed ID: 25470560]. https://doi.org/10.1111/eje.12126.

32. Abd-Rahman ANA, Baharuddin IH, Abu-Hassan MI, Davies SJ. A comparison of different standard-setting methods for professional qualifying dental examination. *J Dent Educ*. 2021;**85**(7):1210-6. [PubMed ID: 33792052]. https://doi.org/10.1002/jdd.12600.

33. Cavalcanti RB, Detsky AS. The education and training of future physicians: why coaches can't be judges. *JAMA*. 2011;**306**(9):993-4. [PubMed ID: 21900141]. https://doi.org/10.1001/jama.2011.1232.

34. Deiorio NM, Carney PA, Kahl LE, Bonura EM, Juve AM. Coaching: a new model for academic and career achievement. *Med Educ Online*.

2016;**21**:33480. [PubMed ID: 27914193]. [PubMed Central ID: PMC5136126]. https://doi.org/10.3402/meo.v21.33480.

35. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA*. 2006;**296**(9):1094-102. [PubMed ID: 16954489]. https://doi.org/10.1001/jama.296.9.1094.