

# Correlation between theoretical knowledge and the difference between self- and teaching doctor-assessment

Janosch Goob Dr<sup>1</sup> | Kurt Erdelt Dr Dipl-Ing<sup>1</sup> | Ephraim Nold Dr MSc<sup>1</sup> |  
Anja Liebermann Univ Prof Dr MSc<sup>2</sup>

<sup>1</sup>Department of Prosthetic Dentistry,  
University Hospital, LMU Munich,  
Munich, Germany

<sup>2</sup>Department of Prosthetic Dentistry,  
University of Cologne, Faculty of  
Medicine and University Hospital  
Cologne, Cologne, Germany

## Correspondence

Janosch Goob, Department of Prosthetic  
Dentistry, University Hospital, LMU  
Munich, Goethestrasse 70, 80336 Munich,  
Germany.

Email:

[Janosch.Goob@med.uni-muenchen.de](mailto:Janosch.Goob@med.uni-muenchen.de)

## Abstract

**Objectives:** The right intuition for self-assessment without overestimating or underestimating oneself compared to theoretical knowledge is a skill that requires conscious practice and is very often paradoxically opposed to the level of knowledge. Self-assessment is an essential skill for dental professionals for lifelong learning and improvement. The objective of our study was to assess the correlation between theoretical knowledge and the difference between self- and teaching doctor-assessment.

**Methods:** The app “digital course organizer” for organization and assessment was used for the self- and teaching doctor-assessment of students for each day of patient treatment during the clinical courses at a university hospital. The difference between the self- and teaching doctor-assessment was compared to the score of an initial theoretical written test at the beginning of the eighth semester to assess the correlation between overestimation or underestimation and theoretical knowledge.

**Results:** A total of 309 dental students over four semesters in the clinical study phase (fourth and fifth years; eighth and 10th semesters) participated in the investigation. The overall view of all values showed very low correlations (<0.2) of the assessment difference values of both practical courses for the initial written test.

**Conclusion:** There were very low correlations between the initial written test (theoretical knowledge) and the difference between self- and teaching doctor-assessment as well as no evidence of overestimation and underestimation.

## KEYWORDS

app-based assessment, assessment, prosthodontic treatment, self-assessment, self-evaluation, teacher's evaluation teaching assessment

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Journal of Dental Education* published by Wiley Periodicals LLC on behalf of American Dental Education Association.

## 1 | INTRODUCTION

It is part of human nature that self-perception of competence often deviates from the truth.<sup>1-3</sup> When it comes to self-assessment in a variety of contexts, it is a common occurrence that people do not assess themselves correctly in terms of their performance, knowledge, and skills.<sup>4,5</sup> Investigations have already specified that low-achieving people, for example, with the lowest score on a test, tend to overestimate their performance the most because they cannot recognize the qualitative difference between their own performance and the performance of others.<sup>3-5</sup> Whereas the high-achieving people self-assess themselves more cautiously and critically and show less overestimation with a general tendency to underestimate their performance.<sup>3,4,6-8</sup> This cognitive bias that people with low ability paradoxically tend to overestimate themselves is known as the Dunning-Kruger effect (DKE).<sup>3</sup> DKE can help to explain paradoxical phenomena in society. The greater a person's ignorance is, the more confident they seem to be in their own knowledge and label evidence-based scientific knowledge as flawed.<sup>9,10</sup> This impact puts a double burden on those affected individuals. On the one hand, their incomplete and misguided knowledge leads them to erroneous conclusions and unfortunate choices; on the other hand, the same deficits rob them of the metacognitive ability and prevent them from recognizing their own mistakes or from adequately assessing wiser decisions.<sup>3,11</sup>

The knowledge about this cognitive bias is of great interest for medical educators in general and especially in the field of dentistry as students start to treat their own patients at an early stage, supervised by teaching doctors, in the clinical section of their studies. This suggests that those with the lowest abilities are most at risk of misjudging their abilities.<sup>12</sup> Therefore, dental students need to learn to transfer their knowledge into applied theory and practice. As a dentist, it is essential to properly assess one's own capabilities and adequately evaluate the quality of treatment. It is also important that teaching doctors have a high competence level at their jobs to be able to assess the students accurately and constructively. However, there are reports that visual inspection, which is a standard analysis method in dentistry, is flawed due to a variety of human factors.<sup>13,14</sup> Besides, it is reasonable to assume that the same self-assessment difficulty can be found among physicians.<sup>12</sup> An investigation showed that a significant proportion of dental students may receive incomplete or inappropriate feedback and invalid grades for their performance.<sup>15</sup> Studies have compared the level of agreement between student self- and teaching doctor-assessment and analyzed it in terms of overestimation and underestimation.<sup>16-18</sup>

The present investigation introduces the results of a dual assessment app ("digital course organizer" [DCO]) and compares the dual assessment results with the outcome of an initial theoretical written test. The app "DCO" provides a daily dual assessment of individual treatment steps, consisting of self-assessment by the students and expert assessment by the teaching doctor, for direct feedback in the clinical section of prosthetic dentistry.

The investigation aimed to identify deficits and deviations in self-perception compared to the theoretical knowledge of dentistry students. To the best of the authors' knowledge, no app-based dual assessment of daily clinical treatment performance has been evaluated and compared in terms of its correlation with theoretical knowledge.

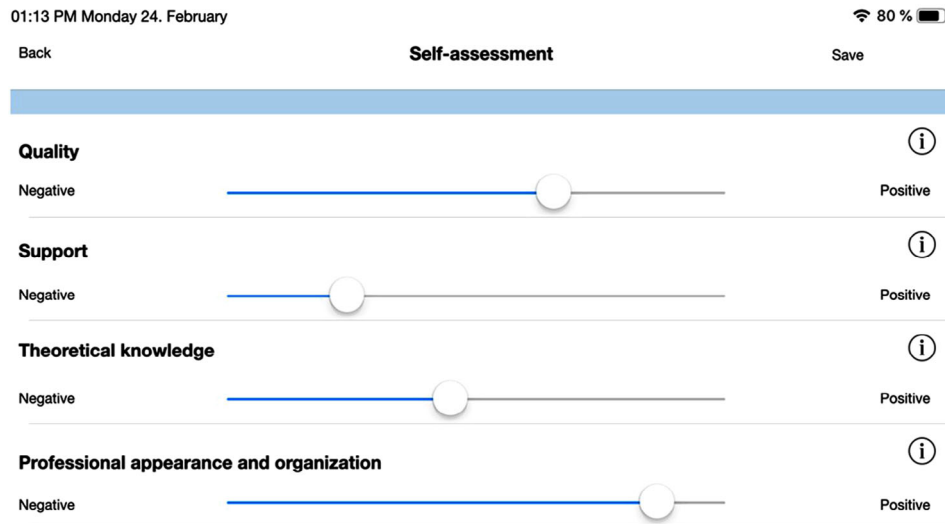
The first hypothesis of this investigation is that there is no correlation between the theoretical knowledge and the difference between self- and teaching doctor-assessment. The second hypothesis is that there is no difference in self-assessment skills among dental students regardless of their academic year. The third hypothesis is that there is no difference in self-assessment skills between low and high scorers of theory exam.

## 2 | MATERIALS AND METHODS

The investigation was approved by the ethics committee of the Medical School (project no. 21-0313) and declared harmless.

In the clinical section of dental education, students rotate through the various disciplines of dentistry. Out of 10 semesters, the eighth and 10th semesters (fourth and fifth years) are devoted to prosthetic dentistry with a focus on fixed and removable dental prosthesis. Dental students must pass a written test at the beginning of the first practical prosthetic semesters (eighth semester) in the respective university in order to objectively assess theoretical knowledge before participation in dental clinical courses. This test consisted of 20 multiple-choice questions and must be answered within 30 min. The test includes questions on the topics: fixed and removable dental prosthesis, functional therapy, temporomandibular disorders, implant prosthodontics, dental materials, conventional cementation, adhesive luting, general treatment procedures, hygiene, and organization in dental clinical courses. The minimum number of points required to pass the exam and to participate in the clinical practical course is 12 points, that is, 60%, with a maximum of 20 points.

Dental students from the eighth and 10th semesters treat patients with prosthetic treatment needs under the supervision of a teaching doctor (graduated dentist who is practicing the profession and also playing a teaching role at dental college) and senior dentists (graduated dentist



**FIGURE 1** Self- and teaching doctor-assessment features. The information button explains question in detail.

who is practicing the profession and managing the clinical courses as supervision as well as the assisting teaching doctors). Each semester was supervised by four teaching doctors whereby each teaching doctor accompanied the same group of students for the whole semester. An app “DCO” was developed in the Department of Prosthetic Dentistry in cooperation with an external software engineer to continuously assess the performance and progress in dental clinical course. The app can be obtained for free via a download link and was accessible only to students and teaching doctors of this specific dental school. The app “DCO” for organization and assessment in dental clinical courses was used for the self- and teaching doctor-assessment for each day of treatment. The dual assessment questionnaire consisted of four questions (Q) that were answered using a visual analog scale (VAS) (Figure 1). The VAS ranged from 1 to 10 (1%–100%). To ensure that the assessment was uniform as possible, students were familiarized with the app and assessment system at the beginning of the semester and were instructed and calibrated by the teaching doctors for each assessment category (Figure 1). Additional instructional videos in screencast format for interface use of the app and the dual assessment feature were made available to all students via an already existing online learning platform (Moodle, Moodle Pty Ltd.). Supervising teaching doctors were also familiarized with the assessment procedures and calibrated by senior dentists. After each treatment session the following Q were assessed: (i) quality of treatment; (ii) support from the teaching doctor; (iii) theoretical knowledge preparation; and (iv) professional appearance and organization (Figure 1 and Table 1).

The answers for the VAS questions were marked with a scroll bar on the line, which reflected the range from 1 to 10 (0%–100%). The evaluation items were consciously chosen

**TABLE 1** Definition of each category.

Quality of treatment	How would you rate the quality of your treatment today?
Support from the teaching doctor	How much assistance/support did you need from the teaching doctor?
Theoretical knowledge preparation	How were you theoretically prepared for the treatment day?
Professional appearance and organization	How do you evaluate your appearance in front of the patient and the organization of your treatment?

in order to critically assess the clinical process afterwards. Within 24 h after patient treatment, students finished self-assessment before receiving any feedback and comments from their supervising teaching doctor to avoid bias. Since self-assessment by students was completed, teaching doctors assessed the students using the same criteria without knowing self-assessment by students to avoid bias. Only after completed assessment by both parties, students and teaching doctors have access to the results. The discrepancy between self- and teaching doctor-assessment was graphically illustrated as underestimation and overestimation in the app individually for each student (Figure 2). All students were encouraged to compare the difference between self- and teaching doctor-assessment to evaluate their self-assessment accuracy. In the case of a large assessment discrepancy, the treatment session was reviewed with the student in order to use the feedback to reflect students’ weaknesses and strengths. The assessment data were collected for each student and each day of treatment over one semester (approximately 3 months) and compared with the outcome of the initial written test.

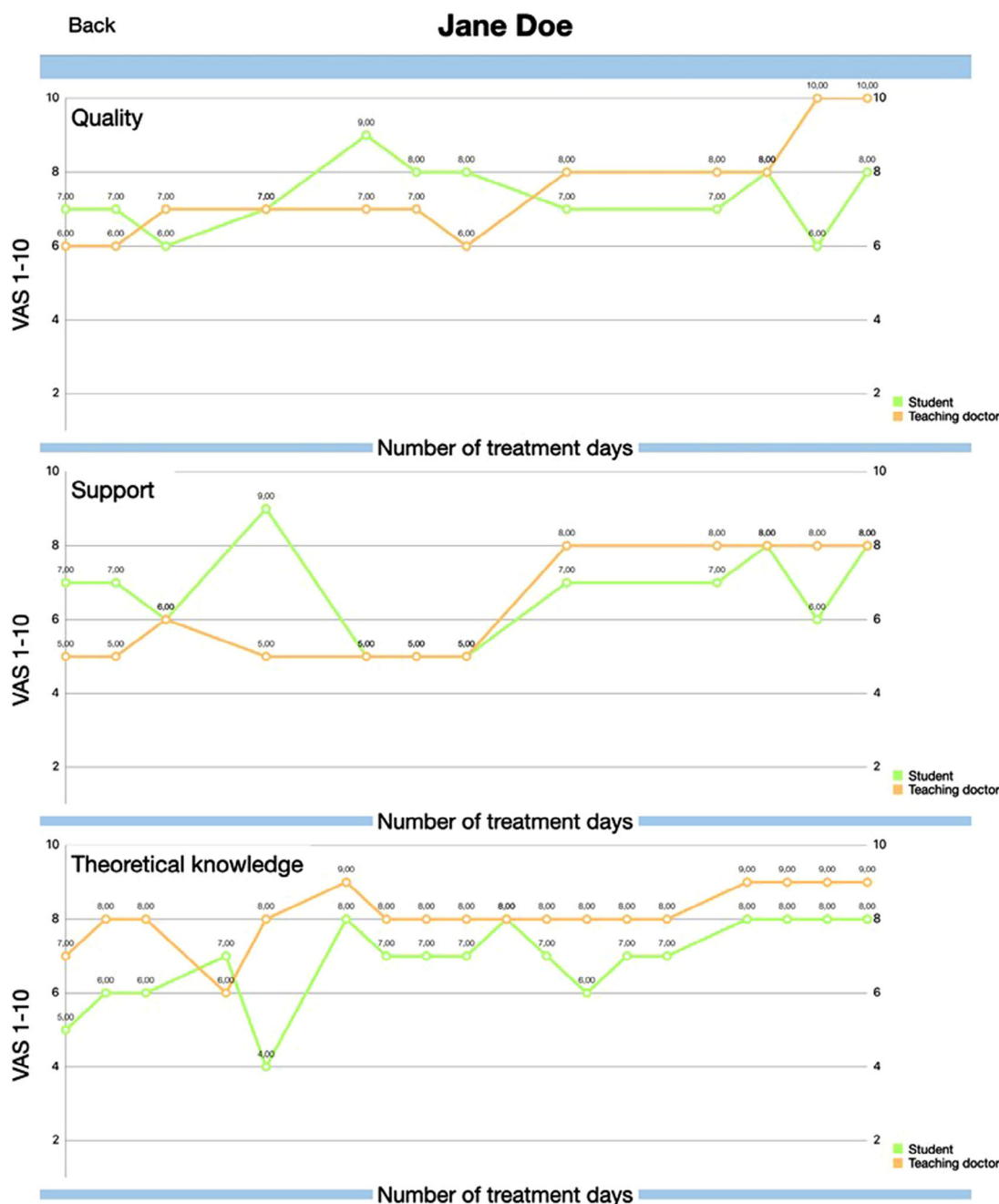


FIGURE 2 Exemplary graphical illustration of self-assessment (green) and teaching doctor-assessment (orange) for each day of treatment in dental clinical courses. Differences between the green and orange graphs reflect overestimation and underestimation.

## 2.1 | Statistical analysis

Data were analyzed using the statistical program SPSS 28.0 (IBM). The difference scores of students' self- and teaching doctor-assessment were calculated and stored in the variable "assessment difference." The difference scores of the students over the investigation period were tested for normal distribution using the Kolmogorov–Smirnov

test. Non-parametric tests were used to analyze the data.

To calculate the correlation values (0: no correlation; 1: maximum correlation—functionally describable), variable scores and assessment difference, the Kendall-Tau-b test was used. For further statistical analyses, the Kruskal–Wallis test and the Wilcoxon test were applied. The significance level was set at  $p \leq 0.05$ .

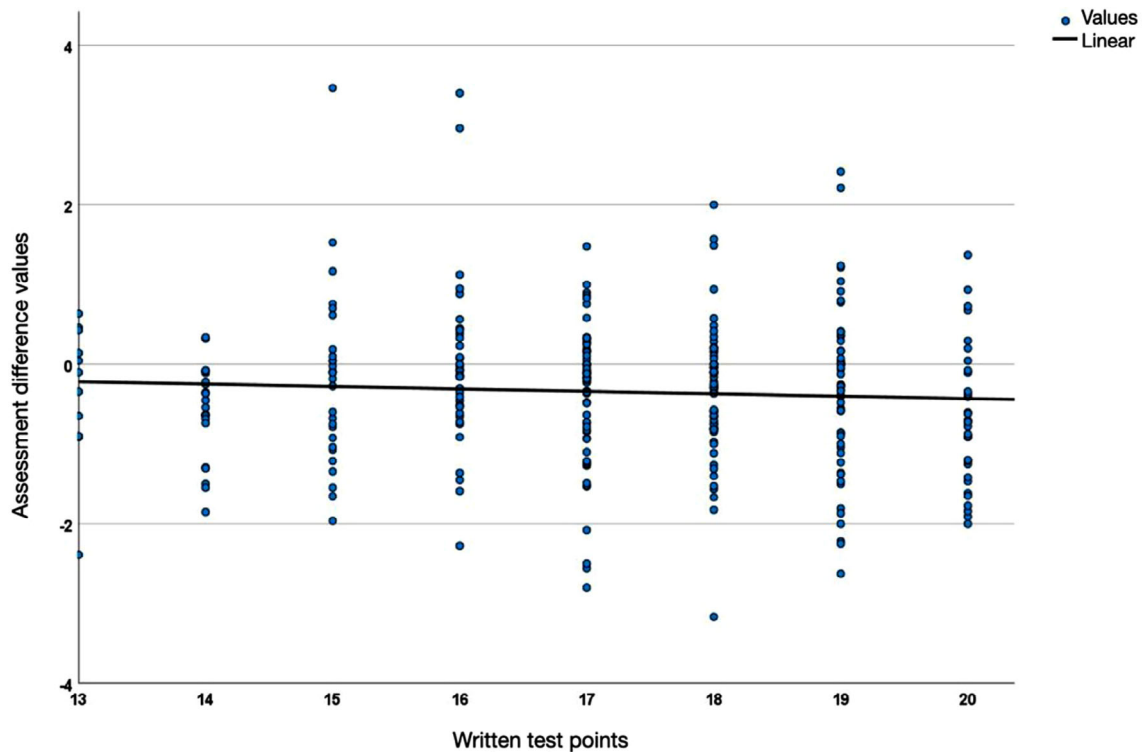


FIGURE 3 Correlation between assessment difference values of both practical courses and initial written test points.

### 3 | RESULTS

A total of 309 dental students over four semesters in the clinical investigation phase (eighth and 10th semesters) of a department of prosthetic dentistry at a university hospital participated.

In total, 51.9% of the assessment difference values of the 309 students involved did not show a normal distribution.

The overall view of all values, a total of 1448 self- and teaching doctor-assessment, showed low correlations ( $<0.2$ : very low correlation) of the assessment difference values of both practical courses for the written test at the beginning of the first practical prosthetic semester (Figure 3).

By analyzing the assessment difference values from the initial written test points group (groups from “written test points 13–20”; see Figure 3), there was no statistical difference within each of these groups (Kruskal–Wallis:  $p > 0.085$ ). Consequently, the assessment difference scores for the study groups could be seen as one group.

The assessment categories “support from the teaching doctor” and “professional appearance and organization” showed significant differences between the eighth and 10th semesters ( $p \leq 0.038$ ). The other two evaluation points, “quality of treatment” and “theoretical knowledge preparation,” presented no significant differences ( $p \geq 0.288$ ).

### 4 | DISCUSSION

The present study analyzed the difference between self- and teaching doctor-assessment after each day of treatment over one semester and compared the results to an initial theoretical written test. As hypothesized, the study found a very low correlation ( $<0.2$ ) between theoretical knowledge (initial written test) and the difference between self- and teaching doctor-assessment (first hypothesis), regardless of the academic year/eighth or 10th semester (second hypothesis). With a correlation value of 0.8 or more, a correlation of the values could be presumed. Additionally, there was no evidence that students with lower scores on the initial test had a higher discrepancy in their assessments, or vice versa (third hypothesis). Students with the highest number of incorrect answers (lowest test results) on the initial written test showed no significant tendency with regard to the highest degree of erroneous assessment during clinical courses. Conversely, for students with the lowest number of incorrect answers (highest test results), which was considered as proficient theoretical preparation in this investigation, no correlation was found either. Furthermore, the assessment categories “support from the teaching doctor” and “professional appearance and organization” showed significant differences between the eighth and 10th semesters ( $p \leq 0.038$ ). Students from the 10th semester are potentially



more experienced than those from the eighth semester and have already treated prosthetic patients. They had already gained practical experience under the supervision and help of experienced dentists. They possibly need less help from their teaching doctor, are confident demeanor in front of patients and are organized because procedures are known.

Obviously, these results raise interesting questions for further research in dental education since the hypothesis was confirmed and the results were not consistent with similar investigations.<sup>8,19,20</sup> In other investigations, students tend to overrate their performance when compared with teaching doctor assessments.<sup>16,18</sup> It is important to note that the teaching doctors had different levels of experience, which may have caused bias in the assessment data, as the same teaching doctor always assessed the same group of students. Given this limitation, it would be valuable to examine the internal variation of the assessments between the different teaching doctors for the same student. The present investigation has chosen the initial written test as reference value since it is uniform and covers the theoretical knowledge about the demands of prosthetic clinical courses. Practical course grades are difficult to use as a reference due to the fact that prosthetic workflows vary too much in their degree of difficulty depending on the type of work, patient compliance, and patient-specific factors. In clinical dentistry, there is no such thing as a standardized patient. Although a clinical procedure may seem straightforward during didactic training but when the student gets to clinic many other variables enter the picture that may affect the achievable level of quality. In addition to practical variables, personal factors can also play a role and influence the results. These factors might include fatigue, interpersonal dynamics, operator nerves, time allowed for inspection, and many other factors. Additionally, it should be recognized that it is difficult for students to adequately assess themselves correctly for dental steps they have partly never or rarely done on the patient before. A further limitation of the study is that students and teaching doctors had 24 h to complete the assessment. Immediately after a treatment day, a student's emotional state may be more heightened, which may affect their performance on an assessment. For example, if someone is feeling anxious or stressed after a treatment, they may not perform as well on an assessment as they would when they are calmer. After 24 h, a person's emotional state may have stabilized, which may improve their performance and gain a deeper understanding of it on assessment.

Despite these limitations, from an educational perspective, it can still provide a valuable contribution to the research on assessment due to the significant amount of data it offers. The findings and insights derived from these data can inform further research and improve our

understanding of the topic. The results show that a large number of students from the clinical courses do not follow the pattern of DKE. This does not mean that they assess themselves correctly but it shows that there is no tendency of overestimation or underestimation with regard to their theoretical knowledge. This may be seen as an advantage, as professions such as dentistry are largely self-regulating professions after graduation. The ability to accurately self-assess and self-criticize is necessary in dentistry where lifelong acquisition of new skills is mandatory.<sup>8,20,21</sup> Deficits can only be identified through accurate self-assessment, which can sustainably improve practice.<sup>12</sup> Being able to accurately self-assess is a skill that requires intentional practice.<sup>19</sup>

Overall, it is challenging to generalize the results to other universities because educational design and didactics may vary from one university to another. However, this is not the only question that remains unanswered and needs further exploration. Thus, conducting a multicenter investigation across different universities with similar assessment studies and more variables, as described above, is necessary to compare and gather more information and clarity.

## 5 | CONCLUSION

There were very low correlations between the initial written test results and the difference between self- and teaching doctor-assessment. There is no difference in self-assessment skills among dental students regardless of the academic year, and no evidence of overestimation and underestimation.

## ACKNOWLEDGMENTS

The authors would like to thank all teaching doctors and students of the Department of Prosthetic Dentistry of the University Hospital (LMU) of Munich for supporting this study.

Open access funding enabled and organized by Projekt DEAL.

## CONFLICT OF INTEREST STATEMENT

The authors declare they have no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The data will be provided on request.

## REFERENCES

1. Dunning D. The Dunning–Kruger effect: on being ignorant of one's own ignorance. *Adv Exp Soc Psychol*. 2011;44:247-296.
2. Isaacs D. The illusion of superiority: the Dunning–Kruger effect and COVID-19. *J Paediatr Child Health*. 2022;58(2):224-225.

3. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol.* 1999;77(6):1121-1134.
4. Dunning D, Johnson K, Ehrlinger J, Kruger J. Why people fail to recognize their own incompetence. *Curr Dir Psychol Sci.* 2003;12(3):83-87.
5. Sanchez C, Dunning D. Overconfidence among beginners: is a little learning a dangerous thing? *J Pers Soc Psychol.* 2018;114(1):10-28.
6. Pennycook G, Ross RM, Koehler DJ, Fugelsang JA. Dunning-Kruger effects in reasoning: theoretical implications of the failure to recognize incompetence. *Psychon Bull Rev.* 2017;24(6):1774-1784.
7. Moore DA, Healy PJ. The trouble with overconfidence. *Psychol Rev.* 2008;115(2):502-517.
8. Cho GC, Chee WW, Tan DT. Dental students' ability to evaluate themselves in fixed prosthodontics. *J Dent Educ.* 2010;74(11):1237-1242.
9. Motta M, Callaghan T, Sylvester S. Knowing less but presuming more: Dunning-Kruger effects and the endorsement of anti-vaccine policy attitudes. *Soc Sci Med.* 2018;211:274-281.
10. Nyhan B, Reifler J, Richey S, Freed GL. Effective messages in vaccine promotion: a randomized trial. *Pediatrics.* 2014;133(4):e835-e842.
11. Coutinho MVC, Thomas J, Alsuwaidi ASM, Couchman JJ. Dunning-Kruger effect: intuitive errors predict overconfidence on the cognitive reflection test. *Front Psychol.* 2021;12:603225.
12. Hodges B, Regehr G, Martin D. Difficulties in recognizing one's own incompetence: novice physicians who are unskilled and unaware of it. *Acad Med.* 2001;76(10):S87-S89.
13. McPherson KR, Mennito AS, Vuthiganon J, et al. Utilizing self-assessment software to evaluate student wax-ups in dental morphology. *J Dent Educ.* 2015;79(6):697-704.
14. Kwon SR, Restrepo-Kennedy N, Dawson DV, et al. Dental anatomy grading: comparison between conventional visual and a novel digital assessment technique. *J Dent Educ.* 2014;78(12):1655-1662.
15. Greany TJ, Yassin A, Lewis KC. Developing an all-digital workflow for dental skills assessment: part I, visual inspection exhibits low precision and accuracy. *J Dent Educ.* 2019;83(11):1304-1313.
16. San Diego J, Newton T, Quinn B, Cox M, Woolford M. Levels of agreement between student and staff assessments of clinical skills in performing cavity preparation in artificial teeth. *Eur J Dent Educ.* 2014;18(1):58-64.
17. Huth K, Baumann M, Kollmuss M, Hickel R, Fischer M, Paschos E. Assessment of practical tasks in the Phantom course of Conservative Dentistry by pre-defined criteria: a comparison between self-assessment by students and assessment by instructors. *Eur J Dent Educ.* 2017;21(1):37-45.
18. Awooda EM, Ibrahim HM, Ahmed EA, Abdelhlim EM. Comparison between undergraduate dental students' self-assessment of pre-clinical crown preparation and assessment by two tutors. *J Oral Res.* 2018;7(7):227-231.
19. Morgan MK, Irby DM. *Evaluating Clinical Competence in the Health Professions.* Mosby; 1978.
20. Gordon MJ. A review of the validity and accuracy of self-assessments in health professions training. *Acad Med.* 1991;66(12):762-769.
21. Bayne SC, Schmalz G. Reprinting the classic article on USPHS evaluation methods for measuring the clinical research performance of restorative materials. *Clin Oral Investig.* 2005;9(4):209-214.

**How to cite this article:** Goob J, Erdelt K, Nold E, Liebermann A. Correlation between theoretical knowledge and the difference between self- and teaching doctor-assessment. *J Dent Educ.* 2023;1-7. <https://doi.org/10.1002/jdd.13428>