Çerçi Akçay and Üstün *BMC Medical Education* (2026) 26:980
https://doi.org/10.1186/s12909-026-09338-w

BMC Medical Education

RESEARCH Open Access

Cognitive alignment and assessment validity in a high-stakes dental specialty examination: a Bloom's taxonomy-based analysis

Hülya Çerçi Akçay[1*] and Nilüfer Üstün[2]

Abstract

Background In high-stakes postgraduate examinations, the cognitive complexity of assessment items is central to evaluating advanced clinical reasoning and decision-making competencies. Alignment between examination content, cognitive demand, and competency-based educational frameworks is essential for assessment validity. This study evaluated the cognitive structure of pediatric dentistry questions in the Turkish Dental Specialty Examination (DUS) using Bloom's revised taxonomy and examined their alignment with curricular expectations.

Methods A retrospective cross-sectional analysis was conducted on 127 officially released pediatric dentistry questions administered between 2012 and 2021. Each item was independently classified according to Bloom's revised cognitive levels. Curriculum relevance and scientific accuracy were rated using a 5-point Likert scale. Inter-rater reliability was assessed using weighted Cohen's kappa. Associations between cognitive level and curriculum relevance were analyzed, and temporal trends across examination years were explored.

Results Questions were predominantly concentrated at the Understand and Apply levels, with fewer items categorized at the Analyze level. No questions were classified at the Evaluate or Create levels. Although lower- and higher-order cognitive skills appeared proportionally balanced when dichotomized, higher-order items largely reflected procedural application rather than advanced analytical or evaluative reasoning. No significant temporal progression toward greater cognitive complexity was observed. Curriculum relevance ratings were high overall but showed no significant association with cognitive level.

Conclusions This high-stakes specialty examination predominantly assesses lower- and intermediate-level cognitive processes, with limited representation of advanced higher-order thinking. The findings indicate potential blueprint misalignment with postgraduate competency expectations and underscore the need for deliberate integration of higher cognitive-level items to strengthen assessment validity.

Keywords Bloom's revised taxonomy, Dental specialty examination, Pediatric dentistry education, Assessment validity, Higher-order thinking skills

*Correspondence:
Hülya Çerçi Akçay
hulya.akcay@kocaelisaglik.edu.tr
[1]Department of Pedodontics, Faculty of Dentistry, Kocaeli Health and Technology University, Kocaeli, Turkey
[2]Department of Pedodontics, Faculty of Dentistry, Istanbul Health and Technology University, Istanbul, Turkey

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Assessment is a fundamental pillar of dental education, influencing not only the evaluation of learning outcomes but also students' approaches to studying and clinical reasoning [1, 2]. In postgraduate dental education, high-stakes examinations are expected to assess advanced cognitive skills that go beyond factual knowledge to include application, analysis, and judgment [3, 4]. However, multiple-choice questions (MCQs), despite their widespread use, are often criticized for emphasizing lower-level cognitive skills [5].

Bloom's revised taxonomy provides a solid framework for evaluating the cognitive demands of assessment items [6]. By organizing learning objectives into hierarchical levels, this taxonomy enables objective analysis of whether exam questions adequately reflect the higher-order thinking skills required for clinical competence [7]. In health professions education, aligning assessment design with cognitive goals is crucial to maintaining validity and educational effectiveness [8].

Evidence from dental and medical education shows that written exams frequently rely on recall and comprehension questions, even at advanced training levels [9, 10]. This mismatch can limit assessments' ability to identify higher-level competencies and may unintentionally promote surface learning strategies [11].

In Türkiye, the Dental Specialty Examination (DUS) is a nationally standardized, high-stakes examination that determines admission to postgraduate dental programs [1]. The DUS consists of 120 multiple-choice questions, including 40 basic science and 80 clinical science items, with each dental specialty branch represented by a fixed number of questions. Pediatric dentistry constitutes one of the core clinical disciplines within this structure. Given its competitive nature and its decisive role in specialty placement, the DUS functions not only as a knowledge assessment tool but also as a gatekeeping mechanism that shapes candidates' learning priorities and preparation strategies.

Pediatric dentistry, as a clinical discipline, requires integration of preventive principles, growth and development knowledge, behavior management strategies, and complex treatment planning in medically compromised or anxious children. Effective pediatric dental practice therefore relies heavily on diagnostic reasoning, risk assessment, and clinical judgment rather than factual recall alone. Consequently, the cognitive profile of examination questions in this field should ideally reflect the multidimensional and decision-oriented nature of pediatric dental care.

Furthermore, postgraduate dental education in Türkiye is structured according to nationally defined competency frameworks, including the National Core Dental Education Program (DUÇEP) and the Medical Specialization Board Curriculum Development System (TUKMOS) [12–14]. These frameworks define expected learning outcomes and clinical competency levels, emphasizing progressive development from knowledge acquisition to independent clinical management. Theoretically, Bloom's revised taxonomy provides a complementary cognitive lens through which these competency expectations can be examined [15]. Analyzing DUS questions within this dual framework allows exploration not only of cognitive level distribution but also of the degree to which assessment practices align with nationally defined educational standards.

An objective, taxonomy-based analysis of pediatric dentistry questions can therefore provide valuable insights into the extent to which the DUS aligns with postgraduate educational goals and competency-based frameworks [16]. Understanding the cognitive structure of exam items can guide future improvements in assessment quality, blueprint design, and evidence-based educational policy decisions [17].

Therefore, this study aimed to analyze questions in pediatric dentistry in the Turkish Dental Specialty Examination using Bloom's revised taxonomy, focusing on their cognitive-level distribution and educational relevance.

It was hypothesized that no statistically significant association would be observed between cognitive level and curriculum relevance, and that higher-order cognitive levels would be underrepresented.

Materials and methods

Study design

This study was designed as a retrospective, descriptive cross-sectional educational analysis examining the cognitive demands of pediatric dentistry questions in the Turkish Dental Specialty Examination (DUS). Cognitive complexity was evaluated using Bloom's revised taxonomy as described by Anderson and Krathwohl (2001) [15]. The study focused exclusively on the structural and cognitive characteristics of examination items and did not involve candidate performance data, examination scores, or outcome measures.

Data source and study material

The dataset consisted of officially released DUS question booklets obtained from the website of the Turkish Assessment, Selection and Placement Center (ÖSYM), Ankara, Türkiye (https://www.osym.gov.tr). Pediatric dentistry (pedodontics) questions administered between 2012 and 2021 were systematically reviewed. All materials were publicly accessible and retrieved on 15 January 2026.

Initially, 130 pediatric dentistry questions were identified. Following screening, three items were excluded

due to official cancellation, resulting in a final analytical sample of 127 questions. The DUS is a nationally administered, high-stakes examination that determines admission to postgraduate dental specialty training programs and is designed to assess advanced theoretical knowledge and clinical reasoning skills.

Eligibility criteria

Inclusion Criteria

- Pediatric dentistry questions published in the DUS
- Single-best-answer multiple-choice questions
- Items assessing theoretical knowledge and/or clinical decision-making in pediatric dentistry

Exclusion Criteria

- Officially cancelled or duplicated items
- Questions belonging to other dental specialties
- Items with insufficient information to determine the dominant cognitive process

Cognitive classification using Bloom's revised taxonomy

All included items were independently classified according to Bloom's revised taxonomy, which categorizes cognitive processes into six hierarchical levels: Remember, Understand, Apply, Analyze, Evaluate, and Create [15].

Each question was assigned a single dominant cognitive level corresponding to the highest level of cognitive processing required for a correct response. Classification decisions were guided by the action verbs used in the question stem, contextual complexity, and the degree of reasoning and information integration required, consistent with established principles of assessment design [15].

Classification decisions were further guided by predefined operational indicators, including the type of cognitive processing required (e.g., recall, interpretation, application, or clinical reasoning), the degree of information integration, and whether the item required context-based decision-making.

Rater training and evaluation process

Two independent pediatric dentistry specialists with experience in undergraduate and postgraduate dental education conducted the classification process. Prior to evaluation, both raters reviewed the theoretical definitions and operational descriptors of Bloom's revised taxonomy to ensure conceptual consistency.

Raters independently evaluated all items and were blinded to each other's classifications. Inter-rater reliability was quantified using the intraclass correlation coefficient (ICC), reflecting agreement between two independent raters. Disagreements were not resolved through forced consensus; consensus discussions were conducted only for interpretative clarification.

Assessment of curriculum relevance and scientific accuracy

In addition to cognitive classification, each item was independently evaluated for curriculum relevance and scientific accuracy using a 5-point Likert scale (1 = very poor; 5 = excellent).

Curriculum relevance was operationalized as alignment with nationally defined undergraduate pediatric dentistry learning outcomes outlined in the National Core Curriculum for Undergraduate Dental Education (DUÇEP-2024) [12]. Scientific accuracy referred to consistency with current evidence-based knowledge and accepted clinical guidelines.

This structured scoring approach enabled quantitative assessment of content alignment and allowed examination of the relationship between cognitive complexity and curricular representation.

Theoretical framework for cognitive–competency alignment

To strengthen construct interpretation, Bloom's revised taxonomy [15] was conceptually integrated with Miller's Pyramid of Clinical Competence [16] and the national postgraduate competency framework (TUKMOS-2024) [13].

Bloom's taxonomy conceptualizes hierarchical cognitive processes, whereas Miller's Pyramid describes progression from knowledge acquisition ("Knows") to independent clinical performance ("Does") [16]. TUKMOS defines postgraduate specialty competencies in terms of increasing levels of clinical autonomy and responsibility [13].

A conceptual alignment model was therefore developed to map cognitive complexity (Bloom) to observable clinical competence (Miller) and nationally defined postgraduate expectations (TUKMOS). This integration was used solely for interpretative purposes and does not imply formal equivalence between frameworks; rather, it provides a structured lens for evaluating blueprint coherence and construct representation within the examination.

Statistical analysis

Statistical analyses were performed using Python (version 3.11). Data management and statistical computations were conducted using the NumPy, pandas, and SciPy libraries.

Descriptive statistics were used to summarize the distribution of questions across Bloom's cognitive levels. Categorical variables were expressed as frequencies and percentages.

Inter-rater reliability for Bloom's taxonomy classification was assessed using the intraclass correlation

Table 1 Descriptive statistics of pooled suitability by Bloom level

Bloom level	Code	(*n*)	Mean	SD	Median	IQR	Percentage(%)
Remember	1	7	4.43	0.98	5.00	1.00	4.7
Understand	2	59	4.40	0.91	5.00	1.00	47.2
Apply	3	37	4.46	0.79	5.00	1.00	26.8
Analyze	4	24	4.31	1.01	4.75	1.13	21.3
Evaluate	5	0	0.0	0.0	0.0	0.0	0.0
Create	6	0	0.0	0.0	0.0	0.0	0.0
Total		**127**					**100**

Bloom levels were classified according to the revised Bloom's taxonomy. Scores represent pooled mean ratings of two independent reviewers on a 5-point Likert scale (1–5). *n* indicates the number of questions per level. Median and IQR are reported. Percentage values show the distribution of questions across Bloom levels. No items were classified as Evaluate or Create

Table 2 Summary of Bloom's cognitive level grouping

Cognitive Group	Bloom Levels	(*n*)	Percentage (%)
Lower-order thinking skills	Remember, Understand	66	51.9
Higher-order thinking skills	Apply, Analyze, Evaluate, Create	61	48.1
Total		**127**	**100**

Lower- and higher-order cognitive groupings were defined in accordance with Bloom's revised taxonomy

coefficient (ICC), two-way mixed-effects model, absolute agreement definition (ICC [2, 1]). This model was selected because it evaluates agreement between fixed raters rating the same set of items and is appropriate for hierarchical ordinal classifications treated as interval-scaled ratings.

Bloom's cognitive levels were further dichotomized into lower-order (Remember, Understand) and higher-order (Apply, Analyze, Evaluate, Create) categories.

Associations between cognitive level (lower vs. higher order) and curriculum relevance were examined using the chi-square test. Correlation between ordinal Bloom levels and curriculum relevance scores was analyzed using Spearman's rank correlation coefficient. A p-value < 0.05 was considered statistically significant.

Ethical considerations

Ethical approval was not required, as the study analyzed publicly available examination materials and did not involve human participants or identifiable data.

Use of artificial intelligence

Artificial intelligence tools were used solely for language editing and grammar refinement. They were not used for study design, data analysis, or interpretation. All scientific decisions and conclusions were made exclusively by the authors.

Results

A total of 127 pediatric dentistry specialty examination questions were included in the analysis. All items were independently evaluated by two reviewers and classified according to Bloom's revised taxonomy, curriculum relevance, and overall educational suitability.

Cognitive level distribution

As shown in Table 1, questions were predominantly distributed across the Understand, Apply, and Analyze levels, with a limited representation of Remember-level items. No questions addressed the Evaluate or Create levels. When cognitive levels were dichotomized (Table 2), lower-order (Remember–Understand) and higher-order (Apply–Analyze) cognitive demands were represented in nearly equal proportions.

Educational suitability across bloom levels

Overall, pooled suitability scores were high across all cognitive levels (Table 1). Median scores consistently fell within the upper range of the Likert scale (1–5), indicating strong perceived educational appropriateness across cognitive categories. Variability in scores was more pronounced in higher-order levels, as visualized in Fig. 1, whereas lower-order levels demonstrated more homogeneous ratings.

Inter-rater agreement

Inter-rater reliability for Bloom-level classification demonstrated good agreement (ICC [2, 1] = 0.85; Table 3), supporting the consistency of cognitive-level judgments between reviewers.

Temporal trends

The distribution of cognitive levels across examination years is presented in Fig. 2. Across all years, Understand and Apply levels remained predominant. No progressive shift toward higher-order cognitive levels was observed over time.

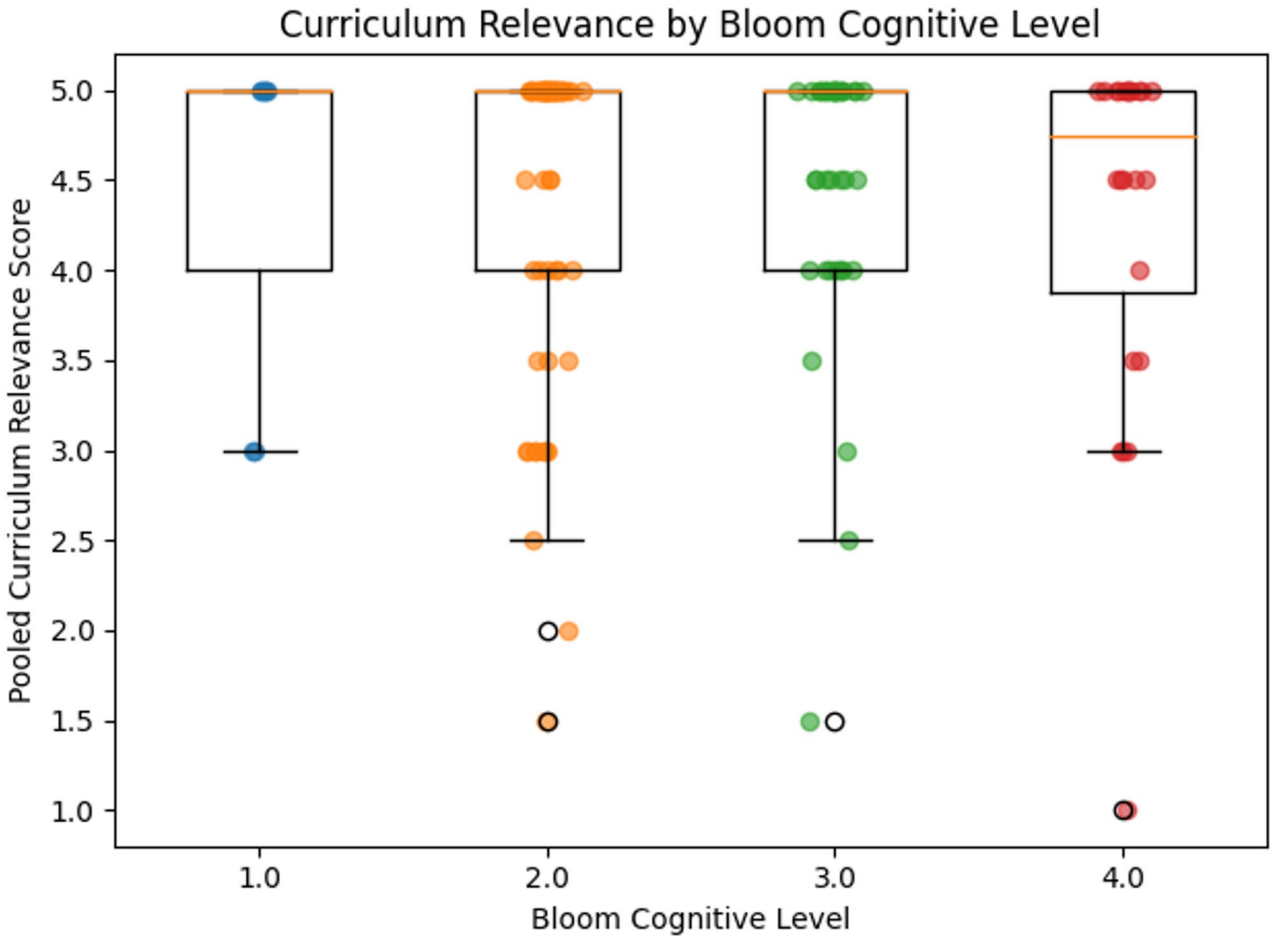

Fig. 1 Box-and-jitter plot showing curriculum relevance scores across Bloom's cognitive levels. The x-axis represents cognitive levels (1 = Remember, 2 = Understand, 3 = Apply, 4 = Analyze), and the y-axis shows Likert-based relevance scores. Points indicate individual questions, and boxplots display median and variability

Table 3 Inter-rater reliability for Bloom's taxonomy classification

Assessment Domain	ICC Model	ICC (2,1)	Level of Agreement
Bloom's cognitive level	Two-way mixed	**0.85**	Good

ICC values were interpreted as follows: <0.50 = poor, 0.50–0.75 = moderate, 0.75–0.90 = good, > 0.90 = excellent agreement

Relationship between cognitive level and curriculum relevance

Spearman correlation analysis revealed no significant association between Bloom's cognitive level and curriculum relevance scores ($\rho = -0.047$, $p = 0.599$; Table 4), suggesting that cognitive complexity and perceived curriculum alignment functioned as independent dimensions in question design.

Conceptual cognitive–competency mapping

When Bloom's cognitive levels were conceptually aligned with Miller's Pyramid and DUÇEP/TUKMOS competency descriptors (Table 5), most pediatric dentistry questions corresponded to the "Knows How" tier of Miller's framework. Items classified at the Understand and Apply levels accounted for the majority of questions (74.0%), whereas Analyze-level questions represented 21.3%. No items were categorized at the Evaluate or Create levels.

From a competency-mapping perspective, lower and mid-level cognitive items primarily aligned with DUÇEP knowledge acquisition and supervised clinical application expectations. In contrast, limited representation was observed for levels corresponding to independent clinical management and advanced decision-making descriptors.

This interpretative framework illustrates the hierarchical alignment between cognitive complexity (Bloom), observable clinical competence (Miller), and nationally defined postgraduate training expectations (DUÇEP/TUKMOS). The model was developed to support construct interpretation and blueprint coherence analysis within high-stakes specialty assessment. It does not imply formal regulatory equivalence between frameworks but

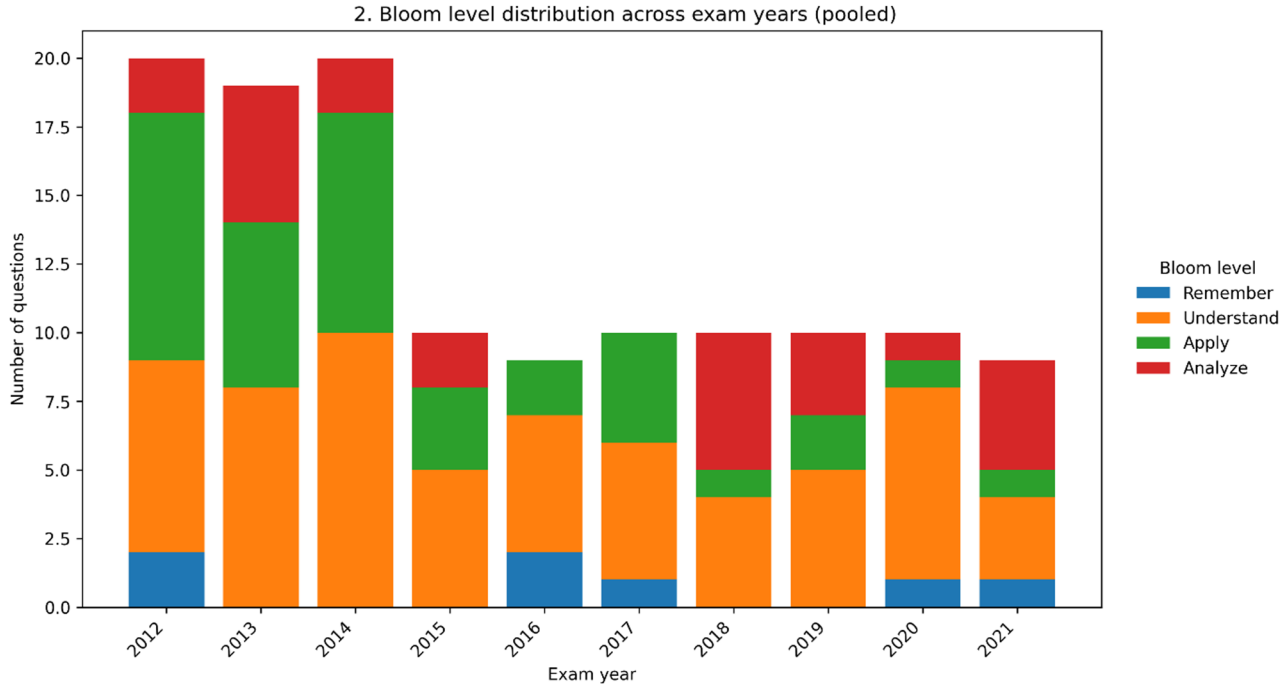

Fig. 2 Distribution of pediatric dentistry questions across Bloom's cognitive levels by examination year (2012–2021). Stacked bars represent the number of questions at each cognitive level per year

Table 4 Correlation between Bloom's cognitive level and curriculum relevance

Variables	Spearman's r_s	*p*-value
Bloom's cognitive level vs. curriculum relevance	−0.047	0.599

The association between Bloom's cognitive level and curriculum relevance was evaluated using Spearman's rank correlation analysis

provides a theoretical lens for evaluating representational balance in examination design.

Discussion

The present study provides a systematic and longitudinal evaluation of the cognitive demands of pediatric dentistry questions in the Turkish Dental Specialty Examination (DUS) between 2012 and 2021, using Bloom's revised taxonomy as a structured analytical framework for educational assessment. Beyond a descriptive mapping of examination items, the findings raise important questions regarding assessment validity, particularly the alignment between cognitive complexity and the advanced competencies expected in postgraduate pediatric

Table 5 Conceptual cognitive–competency alignment framework integrating Bloom's revised taxonomy, Miller's Pyramid of Clinical Competence, and National Postgraduate Competency Standards (DUÇEP-2024 and TUKMOS-2024)

Bloom Level	Cognitive Process	Miller's Pyramid	DUÇEP Competency Orientation	TUKMOS Clinical Level	Expected Clinical Autonomy
Remember	Recall of facts	Knows	Knowledge acquisition	B	Recognition and referral
Understand	Explain concepts	Knows How	Theoretical integration	T	Diagnosis with guidance
Apply	Use knowledge in routine cases	Knows How / Shows How	Supervised clinical application	T / K	Structured case management
Analyze	Differential diagnosis and reasoning	Shows How	Integrated clinical reasoning	TT	Independent diagnostic planning
Evaluate	Critical appraisal and prioritization	Shows How / Does	Advanced treatment decision-making	TT + A	Independent and emergency-level decisions
Create	Novel synthesis and complex planning	Does	Specialist-level independent management	Advanced TT	Full autonomous specialist competence

dentistry training [1–4]. In high-stakes specialty entrance examinations, cognitive demand is not merely a technical feature of item construction; it is a central component of construct representation.

The analysis demonstrated that the majority of questions were concentrated at the Understand and Apply levels, with a smaller proportion classified at the Analyze level. Notably, no items were categorized at the Evaluate or Create levels, which represent the highest tiers of Bloom's cognitive hierarchy. This absence is theoretically significant. Within contemporary validity frameworks, insufficient sampling of higher-order reasoning processes may indicate construct underrepresentation, whereby critical aspects of clinical competence—such as reflective judgment, prioritization of competing treatment options, and synthesis of multidimensional clinical information—are not adequately captured [5, 6]. From a construct validity perspective, such underrepresentation may limit the examination's ability to discriminate candidates based on advanced reasoning skills that are central to autonomous pediatric dental practice [7, 8].

These findings are consistent with prior research in medical and dental education reporting a persistent predominance of lower- and mid-level cognitive processes in written assessments, even in high-stakes contexts [1–3, 9]. Although multiple-choice questions (MCQs) are sometimes criticized for favoring recall-based testing, empirical evidence indicates that scenario-based and context-rich MCQs can be intentionally designed to assess higher-order cognitive processes when cognitive intent is explicitly embedded during item construction [5, 11]. Therefore, the complete absence of Evaluate and Create-level items in this examination appears less attributable to inherent format limitations and more suggestive of blueprint design priorities within the assessment system.

The longitudinal stability observed across the ten-year study period further strengthens this interpretation. Despite ongoing international shifts toward competency-based medical and dental education, and increasing emphasis on analytical reasoning and evidence-based decision-making within curricula [6, 8], the cognitive distribution of DUS pediatric dentistry questions showed no meaningful progression toward higher levels. This temporal stagnation raises concerns regarding alignment between evolving educational paradigms and assessment practice. From the perspective of consequential validity, examinations exert powerful influence on learning behaviors; sustained emphasis on lower cognitive levels may inadvertently encourage surface learning strategies and signal that procedural recall and routine application are sufficient for specialty admission [10–12].

An additional important finding is the absence of a statistically significant association between Bloom's cognitive level and curriculum relevance scores. However, the lack of statistical significance does not necessarily indicate the absence of a true relationship. Although most questions were rated as highly aligned with the undergraduate pediatric dentistry curriculum, increased cognitive complexity did not correspond to higher perceived curricular relevance. This dissociation highlights a fundamental conceptual distinction between content alignment and cognitive rigor. Alignment with curriculum content does not necessarily equate to assessment of higher-order thinking skills [9, 10]. From an assessment design standpoint, this finding suggests that integration of higher cognitive levels can be achieved without altering curricular scope, primarily through intentional modification of item framing, contextual complexity, and reasoning demands.

Inter-rater reliability analysis demonstrated good agreement between reviewers, supporting the methodological robustness of taxonomy-based classification when applied by trained domain experts. This finding aligns with previous studies validating Bloom's taxonomy as a reliable tool for systematic examination analysis in health professions education [11, 18, 19]. Nevertheless, it is important to acknowledge that cognitive classification inherently involves expert judgment and that Bloom's framework captures only the cognitive dimension of competence.

When cognitive levels were dichotomized into lower- and higher-order categories, a near-equal distribution was observed. However, higher-order cognition was predominantly represented by Apply-level items rather than Analyze or above. This distinction carries educational significance. Apply-level questions frequently assess algorithmic or routine clinical decision-making, whereas Analyze, Evaluate, and Create levels are more closely associated with complex diagnostic reasoning, prioritization, critical appraisal, and integration of multiple information streams [13]. The limited representation of these advanced processes may therefore constrain the examination's capacity to assess the full spectrum of competencies required for postgraduate specialization.

The conceptual integration of Bloom's taxonomy with Miller's Pyramid and national competency frameworks provides further interpretative depth. While the majority of items corresponded to the "Knows How" tier of Miller's model, representation of the "Shows How" and "Does" tiers was limited. From a construct validity perspective, this pattern may indicate a possible partial misalignment with postgraduate competency expectations; however, this interpretation should be considered cautiously given that the analysis was limited to publicly available examination questions and did not include internal blueprint data. Strengthening representation of higher cognitive tiers may therefore enhance blueprint

coherence and improve alignment between examination design and competency-based postgraduate training objectives.

Several implications emerge from these findings. From an assessment validity standpoint, limited sampling of advanced cognitive processes may reduce discriminatory power among high-performing candidates and constrain accurate ranking in competitive specialty selection contexts. From a policy perspective, deliberate and theory-informed blueprint revision—explicitly incorporating higher cognitive-level targets—may improve both construct representation and educational impact. Integrating cognitive intent alongside content coverage and curricular alignment represents a critical step toward strengthening the validity argument underlying high-stakes specialty examinations [5, 8, 11]. To enhance practical relevance, higher-order cognitive levels may be incorporated into item design through increased contextual complexity and decision-oriented question framing.

This study has limitations. The analysis was restricted to publicly available pediatric dentistry questions and may not fully reflect the complete item pool or internal blueprint specifications of the DUS. In particular, the lack of access to internal examination blueprints limits the ability to draw definitive conclusions regarding structural alignment. Bloom's revised taxonomy captures only the cognitive dimension of assessment and does not encompass psychomotor performance, affective competencies, or real-time clinical judgment [1, 5, 6, 8]. Additionally, cognitive classification inherently involves expert interpretation, introducing potential subjectivity despite satisfactory inter-rater agreement. Finally, the study did not analyze candidate performance data; therefore, no conclusions can be drawn regarding the relationship between cognitive demand and examination outcomes.

Conclusions

This study demonstrates that pediatric dentistry questions in a high-stakes dental specialty examination predominantly assess lower- and intermediate-level cognitive processes, with limited representation of advanced higher-order thinking. Despite general alignment with nationally defined curricular outcomes, cognitive complexity appears independent of curricular relevance, indicating that content alignment alone does not ensure assessment of advanced reasoning skills.

The absence of Evaluate- and Create-level items, together with the longitudinal stability of cognitive distribution, may suggest a possible misalignment, which should be interpreted cautiously given the scope of the available data, with postgraduate competency expectations that emphasize independent clinical judgment and complex decision-making. From a construct validity perspective, more deliberate integration of higher cognitive-level items may enhance representation of advanced professional competence and improve discriminatory capacity in competitive specialty selection contexts.

Future examination redesign efforts should explicitly incorporate cognitive-intent mapping within blueprint development to strengthen alignment between assessment structure and competency-based postgraduate training objectives. Future research should aim to replicate these findings across different specialties and examination systems to strengthen generalizability.

Abbreviations

DUS	Dental Specialty Examination
DUÇEP	National Core Curriculum for Undergraduate Dental Education
ICC	Intraclass Correlation Coefficient
IQR	Interquartile Range
MCQ	Multiple-Choice Question
ÖSYM	Assessment, Selection and Placement Center
SD	Standard Deviation
TUKMOS	Medical Specialty Curriculum Development and Standardization System

Acknowledgements

During the preparation of this manuscript, the authors used ChatGPT (OpenAI, version GPT-5.2) for language editing and grammar refinement purposes. The authors reviewed, edited, and take full responsibility for the final content of this publication.

Authors' contributions

Conceptualization: H.C.A. ; Methodology: H.C.A. ; Data curation: H.C.A. & N.U. ; Formal analysis: H.C.A. & N.U. ; Writing – original draft: H.C.A. & N.U. ; Writing – review & editing: H.C.A. & N.U.

Funding

This research received no external funding.

Data availability

The data analyzed in this study consist of officially released pediatric dentistry questions published by the Turkish Assessment, Selection and Placement Center (ÖSYM). These materials are publicly available through the official ÖSYM website. No new experimental data were generated. The data extraction and classification procedures are described in the Methods section to ensure transparency and reproducibility.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 26 February 2026 / Accepted: 27 April 2026
Published online: 30 April 2026

References

1. Turunc-Oguzman R, Yurdabakan ZZ. Performance of Chat Generative Pre-trained Transformer and Bard on the questions asked in the dental specialty entrance examination in Türkiye regarding Bloom's revised taxonomy. Curr Res Dent Sci. 2024;34(1):25–34. https://doi.org/10.5152/CRDS.2024.23261.

2. Turan Gökduman C, Arılı Öztürk E, Aktaş Ş, Çanakçi BC. Comparison of chatbots' accuracy in endodontics questions in dentistry specialization exam in Türkiye: ChatGPT-4o, Gemini Advanced, Copilot, and Claude. BMC Oral Health. 2026;26:28. https://doi.org/10.1186/s12903-025-07346-8.
3. Vieno-Corbett K, Campbell N, Pecyna A, Rogers KA. A scoping review of teaching approaches and learning objectives for anatomical variation in gross anatomy courses across degree programs. Anat Sci Educ. 2025;18:937–47. https://doi.org/10.1002/ase.70072.
4. Schüler IM, Bach K, Schädlich P, Rabe U. Developing and evaluating a new preclinical curriculum with focus on prevention in dentistry. Dent J. 2025;13:81. https://doi.org/10.3390/dj13020081.
5. Stringer JK, Santen SA, Lee E, Rawls M, Bailey J, Richards A, et al. Examining Bloom's taxonomy in multiple-choice questions: Students' approach to questions. Med Sci Educ. 2021;31:1305–14. https://doi.org/10.1007/s40670-021-01305-y.
6. Mukhopadhyay A, Mukhopadhyay S, Biswas R. Evaluation of large language models in pediatric dentistry: A Bloom's taxonomy-based analysis. Folia Med (Plovdiv). 2025;67(4):e154338. https://doi.org/10.3897/folmed.67.e154338.
7. Gonzalez-Cabezas C, Anderson OS, Wright MC, Fontana M. Association between dental student-developed exam questions and learning at higher cognitive levels. J Dent Educ. 2015;79(11):1295–306.
8. Wei Y, Peng Z. Application of the flipped classroom model based on Bloom's taxonomy of educational objectives in endodontics education for undergraduate dental students. PeerJ. 2025;13:e18843. https://doi.org/10.7717/peerj.18843.
9. Thompson AR, Lake LPO. Relationship between learning approach, Bloom's taxonomy, and student performance in an undergraduate human anatomy course. Adv Health Sci Educ Theory Pract. 2023;28(4):1115–30. https://doi.org/10.1007/s10459-023-10208-z.
10. Herrmann-Werner A, Festl-Wietek T, Holderried F, et al. Assessing ChatGPT's mastery of Bloom's taxonomy using psychosomatic medicine exam questions: A mixed-methods study. J Med Internet Res. 2024;26:e52113. https://doi.org/10.2196/52113.
11. Bharatha A, Ojeh N, Fazle Rabbi AM, et al. Comparing the performance of ChatGPT-4 and medical students on MCQs at varied levels of Bloom's taxonomy. Adv Med Educ Pract. 2024;15:393–400. https://doi.org/10.2147/AMEP.S457408.
12. DUÇEP DDK-EAK. National Core Curriculum for Undergraduate Dental Education (DUÇEP-2024) [in Turkish]. Ankara: Council of Higher Education of Türkiye. 2024. Available from: https://yok.gov.tr/.
13. TUKMOS. Dental Specialty Training Curriculum Framework (Tıpta Uzmanlık Kurulu Müfredat Oluşturma ve Standart Belirleme Sistemi – TUKMOS) [in Turkish]. Ankara: Ministry of Health of Türkiye. 2024. Available from: https://tuk.saglik.gov.tr/.
14. Republic of Türkiye Official Gazette. Law No. 1219 on the Practice of Medicine and Medical Specialties (Tababet ve Şuabatı San'atlarının Tarzı İcrasına Dair Kanun). 11 April 1928; Issue 863. Available from: https://www.mevzuat.gov.tr/.
15. Anderson LW, Krathwohl DR, editors. A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York: Longman; 2001.
16. Miller GE. The assessment of clinical skills/competence/performance. Acad Med. 1990;65(9 Suppl):S63–7. https://doi.org/10.1097/00001888-199009000-00045.
17. Monrad SU, Bibler Zaidi NL, Grob KL, et al. What faculty write versus what students see? Perspectives on multiple-choice questions using Bloom's taxonomy. Med Teach. 2021;43(5):575–82. https://doi.org/10.1080/0142159X.2020.1849752.
18. Çulhaoğlu AK, Kılıçarslan MA, Deniz KZ. Diş hekimliğinde uzmanlık sınavının farklı eğitim seviyelerindeki algı ve tercih durumlarının değerlendirilmesi. Atatürk Üniv Diş. Hek Fak Derg. 2021;31(3):1–7.
19. Yeo S. An analysis of verbs used in the course outcomes of outcome-based integrated courses at a medical school based on the taxonomy of educational objectives. Korean J Med Educ. 2019;31(3):261–9. https://doi.org/10.3946/kjme.2019.136.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.