



An Analysis of Final Exam Questions in Grade 12 Persian Language (Humanities Stream) Based on Bloom's Cognitive Taxonomy

Ghasem Mehravar Giglou*

Corresponding Author, Assistant Professor, Department of Persian Language and Literature Education, Farhangian University, Tehran, Iran.

Gh.mehravar@cfu.ac.ir

Abdolhossein Heydari

Department of English Language Teaching, Farhangian University, Tehran, Iran.

a.heidari@cfu.ac.ir

Ebrahim Samani

Department of English Language Teaching, Farhangian University, Tehran, Iran.

e.samani@cfu.ac.ir

Amirabbas Hajjami

Master's student in Persian Language and Literature, University of Guilan, Guilan, Iran.

amirabbashajjami@gmail.com

Abstract:

The purpose of the present study is to identify and analyze the final examination questions of the Grade 12 Persian language course in the humanities stream according to the different levels of Bloom's taxonomy. This study, using the content analysis method, examined the questions of the national Persian final exams administered between 2014 and 2024 (1393–1403). First, the questions were categorized based on the six levels of Bloom's taxonomy—knowledge, comprehension, application, analysis, evaluation, and creation—and their frequencies were determined. Then, the data were analyzed and examined statistically using the Chi-square test and Fisher's Exact Test. The findings indicate that in the early years, most of the questions were concentrated in the lower cognitive levels (knowledge and comprehension), with the main focus of the exams on basic concept learning and information reproduction. In contrast, the proportion of questions at higher levels such as synthesis and evaluation was very limited. This demonstrates that, during the examined period, the Persian final exams placed less emphasis on developing students' analytical skills, critical evaluation, and creativity. Furthermore, the results show that there have been significant changes over time in the distribution of Bloom's cognitive levels. In particular, higher levels such as application and analysis have increased in recent years, indicating a change in the assessment approach and a greater focus on higher cognitive levels. Compared with earlier years, lower-level questions have decreased, reflecting a shift in educational expectations and evaluation practices. Therefore, in recent years, notable changes have been observed in the structure and content of the Persian final exam questions, moving toward strengthening critical thinking and analytical skills.

***Cite this article:** Mehravar Giglou, Ghasem. Heydari, Abdolhossein. Samani, Ebrahim. Hajjami, Amirabbas. An Analysis of Final Exam Questions in Grade 12 Persian Language (Humanities Stream) Based on Bloom's Cognitive Taxonomy. Vol. 15, No. 1 (Tome 31), April 2026,59-84.

DOI: 10.30479/jtpsol.2026.22865.1732

Received on: 10/10/2025

Accepted on: 10/05/2026



© The Author(s).

Publisher: Imam Khomeini International University

Introduction

Examinations are essential instruments for evaluating students' progress and the effectiveness of educational systems. In Iran, the final national exams (“Azmoon-e Nahayi”) play a central role in assessing students' academic achievements at the end of high school. However, the cognitive depth of these exams — that is, how much they assess different levels of thinking — has often been questioned.

This study aims to evaluate the cognitive levels of the final exam questions in Persian language for 12th-grade humanities students over a ten-year period (2014–2024 / 1393–1403 in the Iranian calendar). The research specifically applies Bloom's revised taxonomy to classify and analyze the questions according to six cognitive levels: knowledge, comprehension, application, analysis, evaluation, and creation. By doing so, the study seeks to determine the extent to which these exams measure higher-order thinking skills, such as critical analysis, reasoning, and creativity, rather than simple memorization and recall.

Bloom's taxonomy (1956) and its revised version (Anderson & Krathwohl, 2001) provide a hierarchical model for categorizing cognitive processes in education. The taxonomy's six levels — from simple recall of facts to creative synthesis and evaluation — offer a structured approach to designing and analyzing educational objectives and assessments.

According to this framework, lower-order skills (knowledge and comprehension) emphasize memory and understanding, while higher-order skills (application, analysis, evaluation, creation) involve reasoning, synthesis, judgment, and innovation.

This research builds on the idea that balanced assessment across all levels is essential for nurturing critical thinking (Facione, 1990; Halpern, 1998) — a skill recognized as fundamental for academic success and informed citizenship. The study assumes that exams heavily weighted toward lower cognitive levels may promote rote learning and limit students' analytical and creative abilities.

The study addresses four key questions:

1. What cognitive levels do the national Persian exams for 12th-grade humanities students primarily assess?
2. Do these exams adequately cover all levels of Bloom's taxonomy?
3. Which levels are most and least represented in the exam questions?
4. What changes or improvements can be made to design exams that better promote higher-order thinking?

Methodology

This study employed a descriptive–analytical design using the method of content analysis to examine the final national examination questions in the Persian language for 12th-grade humanities students in Iran from 2014 to 2024 (1393–1403). The data set included all official exam papers issued by the Ministry of Education, collected through the national assessment archives. Each question was carefully coded and classified according to the six cognitive levels of Bloom's taxonomy — knowledge, comprehension, application, analysis, evaluation, and creation — in order to identify their frequency and distribution. To ensure

accuracy and reliability, two independent raters reviewed and cross-checked the classification results, and discrepancies were resolved through consensus. The categorized data were entered into SPSS version 25 for statistical analysis. Descriptive statistics were used to summarize the frequency and percentage of each cognitive level, while Chi-square and Fisher’s Exact tests were employed to examine the significance of changes across years. This systematic procedure provided both quantitative and qualitative insight into evolving trends in exam design and cognitive emphasis.

Results

Across the ten-year period, the results show a strong dominance of lower-level cognitive skills. Table 1 shows the results of 842 questions which were analyzed in this study.

Blooms’ Taxonomy levels	Frequency in Percentage
Knowledge	26.84%
Comprehension	63.18%
Application	1.54%
Analysis	7.84%
Evaluation	0.36%
Creation	0.24%

As Table 1 indicates, approximately 90% of all questions belong to the two lowest cognitive levels (knowledge and comprehension). In contrast, questions that require students to apply, analyze, evaluate, or create are extremely limited.

The Chi-square and Fisher’s Exact tests revealed a statistically significant relationship between exam year and cognitive level distribution ($p < 0.001$), suggesting some evolution over time. Specifically, questions at the “application” and “analysis” levels have modestly increased in recent years, particularly after 1400 (2021–2022). However, this growth remains limited, and the majority of questions still measure basic recall and understanding.

The findings clearly demonstrate a misalignment between educational objectives and assessment design. Despite the widespread advocacy for promoting critical thinking and analytical literacy within national curricula, the Persian final exams continue to prioritize memorization-based knowledge. The prevalence of low-level questions confirms the persistence of a memory-oriented examination culture, which reinforces surface learning rather than deep understanding.

This pattern echoes earlier studies (Philosofinejad et al., 2016; Shohamy, 2013), which found similar tendencies in Iranian and international contexts. The lack of higher-order questions means that students are rarely challenged to interpret, critique, or generate new ideas based on literary and linguistic content.

On the other hand, the gradual rise in “application” and “analysis” questions in later years suggests a slow shift in exam design philosophy, perhaps reflecting the influence of recent

educational reforms and professional training for teachers and exam writers. However, these shifts are not yet substantial enough to change the overall assessment landscape.

Conclusion

Over the decade studied, national Persian exams for Grade 12 humanities students have predominantly assessed knowledge and comprehension, with minimal attention to analysis, evaluation, or creation. Although minor improvements are evident in recent years, the general pattern indicates a sustained emphasis on surface learning.

To move toward a more dynamic and cognitively rich assessment system, Iran's educational policymakers and exam committees must consciously redesign their question frameworks to promote critical thinking, creativity, and problem-solving. Such transformation will help prepare students not just to recall information, but to apply, interpret, and evaluate it in meaningful contexts.

The predominance of lower-order cognitive questions in the national Persian exams has significant implications for the broader educational process. Such an assessment pattern reinforces rote learning, encouraging students to memorize rather than to think, analyze, and apply their knowledge critically. Consequently, learners are deprived of opportunities to engage in deeper forms of understanding, such as interpretation, synthesis, and creative expression. This imbalance not only limits the development of higher-order thinking skills but also diminishes motivation among students who might otherwise excel in analytical and evaluative tasks. Moreover, it narrows the educational outcomes by producing graduates who are adept at recalling information but insufficiently prepared for the intellectual demands of higher education and professional life.

To address these shortcomings, exam designers and curriculum planners need to adopt a more balanced and progressive approach to assessment design. Examinations should increasingly incorporate questions that target complex cognitive processes, including analysis, synthesis, and evaluation. Such reform would align testing practices more closely with the goals of the national curriculum, which aspires to nurture critical and creative thinkers. Furthermore, teacher training programs should emphasize strategies for constructing higher-order exam questions and for assessing reasoning and interpretive skills effectively. Periodic review and content analysis of national exams are also essential to ensure that all levels of Bloom's taxonomy are adequately represented. By embedding these changes into the structure of exam design, the education system can move beyond memorization-based evaluation and cultivate a learning culture centered on inquiry, reasoning, and intellectual growth.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Keywords: Bloom's taxonomy, Persian final exams, Content analysis of questions, Cognitive assessment, Chi-square test.



بررسی و تحلیل سؤالات آزمون‌های نهایی درس فارسی پایه دوازدهم علوم انسانی بر اساس طبقه‌بندی شناختی بلوم (پژوهشی)

قاسم مهرآور گیگلو*

نویسنده مسئول، استادیار گروه آموزش زبان و ادبیات فارسی، دانشگاه فرهنگیان، تهران، ایران.
Gh.mehrvavar@cfu.ac.ir

عبدالحسین حیدری

دانشیار گروه آموزش زبان انگلیسی، دانشگاه فرهنگیان، تهران، ایران.
a.heidari@cfu.ac.ir

ابراهیم سامانی

استادیار گروه آموزش زبان انگلیسی، دانشگاه فرهنگیان، تهران، ایران.
e.samani@cfu.ac.ir

امیرعباس حجامی

دانشجوی کارشناسی ارشد زبان و ادبیات فارسی، دانشگاه گیلان، گیلان، ایران.
amirabbashajami@gmail.com

چکیده

هدف از انجام پژوهش حاضر، شناسایی و تحلیل آزمون‌های درس فارسی پایه دوازدهم علوم انسانی طبق سطوح مختلف طبقه‌بندی بلوم است. این مطالعه با استفاده از روش تحلیل محتوا به بررسی سؤالات آزمون‌های نهایی فارسی در طول سال‌های ۱۳۹۳ تا ۱۴۰۳ پرداخته است. نخست، سؤالات بر اساس شش سطح طبقه‌بندی بلوم؛ شامل «دانش»، «درک»، «کاربرد»، «تحلیل»، «ارزیابی» و «ترکیب» دسته‌بندی شدند و فراوانی آن‌ها تعیین شد. سپس داده‌ها با استفاده از آزمون کای اسکوئر و تحلیل آماری Fisher's Exact Test مورد بحث و مذاقه قرار گرفت. یافته‌ها نشان می‌دهد که در سال‌های اولیه، بیشترین سهم سؤالات مربوط به سطوح پایین‌تر شناختی (دانش و درک) بوده و تمرکز عمده آزمون‌ها بر یادگیری مفاهیم پایه و بازتولید اطلاعات است. در مقابل، سهم سؤالات مرتبط با سطوح بالاتر؛ همچون «ترکیب» و «ارزیابی» بسیار محدود بوده است. این امر بیانگر آن است که آزمون‌های نهایی درس فارسی، در دوره مورد بررسی، کمتر بر تقویت مهارت‌های تحلیلی، ارزیابی انتقادی و خلاقیت دانش‌آموزان تأکید داشته‌اند. همچنین نتایج پژوهش نشان داد که تغییرات معناداری در توزیع سطوح مختلف طبقه‌بندی بلوم در طول زمان وجود دارد؛ به‌ویژه، سطوح بالاتر طبقه‌بندی بلوم مانند «کاربرد» و «تحلیل» در سال‌های اخیر افزایش یافته‌اند که نشان‌دهنده تغییر در رویکرد ارزیابی و توجه بیشتر به سطوح بالاتر شناختی است. در مقایسه با سال‌های پیش، سؤالات مربوط به سطوح پایین‌تر کاهش یافته است و این روند، حاکی از تغییر در انتظارات آموزشی و ارزیابی است. بنابراین، در سال‌های اخیر، تغییرات قابل توجهی در ساختار و محتوای سؤالات آزمون‌های نهایی فارسی مشاهده می‌شود که به سمت تقویت تفکر انتقادی و مهارت‌های تحلیلی حرکت کرده است.

کلیدواژه‌ها:

طبقه‌بندی بلوم، آزمون‌های نهایی فارسی، تحلیل محتوای سؤالات، ارزشیابی شناختی، تحلیل کای اسکوئر.

* استناد: مهرآور گیگلو، قاسم، حیدری، عبدالحسین، سامانی، ابراهیم، حجامی، امیرعباس. بررسی و تحلیل سؤالات آزمون‌های نهایی درس فارسی پایه دوازدهم علوم انسانی بر اساس طبقه‌بندی شناختی بلوم، سال پانزدهم، شماره اول (پیاپی ۳۱)، بهار و تابستان ۱۴۰۵، ۸۴-۵۹.

۱. مقدمه

امروزه آزمون، یک عنصر اساسی در نظام آموزش و پرورش به‌شمار می‌رود که به‌عنوان ابزاری برای نمایش پیشرفت دانش‌آموزان در حوزه‌های مختلف شناختی و مهارتی به‌کار گرفته می‌شود. برخلاف رفتارهای آشکار که به شکل مستقیم اندازه‌گیری می‌شوند، ویژگی‌هایی چون هوش، انگیزش، خلاقیت و ... به‌طور غیرمستقیم و با آزمون اندازه‌گیری می‌شوند. در تعریف آزمون می‌توان گفت؛ آزمون به مجموعه‌ای سؤال گفته می‌شود که اغلب به‌طور کتبی (مداد و کاغذی) برای پاسخ‌گویی در اختیار آزمون‌شوندگان قرار داده می‌شود (Seyf, 2019, p.34). آزمون، اطلاعات کافی در اختیار والدین، معلم و مدرسه قرار می‌دهد؛ تا از پیشرفت یادگیری دانش‌آموز آگاه شوند. آزمون همچنین محک مناسبی برای خود دانش‌آموز است؛ تا به نکات ضعف و قوت تحصیلی خود پی ببرد. آموزشگاه‌ها و معلمان از آزمون‌ها برای سنجش اثربخشی آموزش و منابع آموزشی بهره می‌برند و در این میان، آزمون نهایی برای این سنجش از اهمیت بیشتری برخوردار است. پژوهش‌های آموزشی نشان می‌دهند که آزمون‌های نهایی، علاوه بر سنجش عملکرد گذشته، به‌دلیل ماهیت جامع و فراگیر خود، می‌توانند فرایند یادگیری را نیز تقویت کرده و موجب تثبیت بهتر مطالب در حافظه بلندمدت شوند (گلس^۱ و همکاران، ۲۰۱۳).

ارزشیابی به‌طور کلی به دو نوع ارزشیابی تکوینی یا فرایندی^۲ و ارزشیابی تراکمی یا مجموعی^۳ تقسیم می‌شود. ارزشیابی تکوینی با هدف ارائه بازخورد به معلم و دانش‌آموز، برای ارتقای کیفیت آموزش و یادگیری در همان جریان تدریس انجام می‌شود؛ نه صرفاً برای داوری نهایی یادگیری. ارزشیابی تراکمی در پایان یک نیم‌سال یا یک برنامه آموزشی انجام می‌شود و هدف آن، تعیین نمره، ارزیابی میزان پیشرفت یا بررسی میزان اثربخشی یک برنامه درسی، دوره آموزشی یا طرح آموزشی است (بلوم^۴ و همکاران، ۱۹۷۱). زمانی یک آزمون به بهبود یادگیری کمک می‌کند که مثلاً دانش‌آموزان را به تفکر عمیق‌تر، مطالعه گسترده‌تر یا به‌کارگیری مهارت‌ها در موقعیت‌های واقعی سوق می‌دهد؛ با اثر مثبت آزمون^۵ همراه است و از سوی دیگر، وقتی آزمون‌ها فقط بر حفظیات و پاسخ‌های کوتاه تأکید کنند و باعث سطحی شدن آموزش شوند، اثر منفی آزمون^۶ را به‌همراه دارند (شوهامی^۷، ۲۰۱۳).

^۱ Glass

^۲ Formative Evaluation

^۳ Summative Evaluation

^۴ Bloom

^۵ positive washback

^۶ Negative washback

^۷ Shohamy

نظریه یادگیری طبقه‌بندی شناختی بلوم^۱ که چارچوبی ساختاری برای طراحی اهداف آموزشی، فعالیت‌های درسی و ارزیابی از آموخته‌ها ارائه می‌کند، فرایند شناختی را به شش طبقه دانش، درک، کاربرد، تحلیل، ترکیب و ارزیابی تفکیک می‌کند (نعمتی و همکاران، ۱۴۰۱). سطوح یادآوری و درک معمولاً ساده‌تر هستند و دانش‌آموز با یادآوری اطلاعات و توضیح آن‌ها می‌تواند پاسخ دهد. سطوح میانی مثل کاربرد و تحلیل، نیاز به تفکر فعال و ترکیب اطلاعات دارند و نسبتاً پیچیده‌تر هستند و در نهایت، سطوح بالاتر مثل ترکیب و ارزیابی که نیاز به قضاوت، نقد و تولید محتوای جدید دارند و پیچیده‌ترین سطح محسوب می‌شوند. در نسخه بازنگری، علاوه بر بازتعریف فرایندهای شناختی، بُعد «دانش» هم به چهار نوع دانش واقعی، دانش مفهومی، دانش مهارتی و دانش فرامعرفتی تقسیم شد. طبقه‌بندی بلوم علاوه بر مورد استفاده قرار گرفتن در سنجش محتوایی، اهداف آموزشی را شفاف می‌کند؛ به طراحی مؤثر تدریس و ارزشیابی کمک می‌کند و برنامه‌ریزی آموزشی را تسهیل می‌بخشد (اندرسون و کراثول^۲، ۲۰۰۱). یکی از اهداف مهم آموزش زبان و ادبیات فارسی، رشد تفکر و تقویت روحیه زیبایی‌دانی دانش‌آموزان است. ویگوتسکی^۳ (۱۹۶۲) بر اهمیت زبان در پیشرفت ادراکی کودکان و وابستگی زبان و تفکر به یکدیگر تأکید دارد. از نظر او رابطه تفکر و کلام چیزی جز یک روند نیست؛ بلکه یک حرکت ممتد رو به جلو و عقب از تفکر به زبان و از کلام به تفکر است. در واقع، رشد زبان به گستردگی تفکر منجر می‌شود و گسترش تفکر نیز بر توسعه زبان تأثیر دارد. واژه‌های بی‌شمار زبان که به صورت شبکه‌های گسترده و زنجیروار مرتبط به هم در ذهن ذخیره می‌شوند؛ با طرح‌واره‌های ذهنی متعدد درهم‌تنیده و پویا ارتباط دارند که مملو از مفاهیم انتزاعی است. بنابراین، با رشد مهارت‌های زبانی که همراه با افزایش دامنه واژگان در ذهن اتفاق می‌افتد، بر وسعت و گستردگی مفاهیم و تصورات ذهنی نیز افزوده می‌شود. همین روابط متقابل زبان و تفکر موجب شده است که همه نظام‌های آموزش و پرورش جهان، آموزش زبان را جزو اولویت‌های خود قرار دهند و بر آغاز آموزش زبان از دوره پیش‌دبستانی و دبستان تأکید داشته باشند. ارزشیابی از آموخته‌های دانش‌آموزان که گام نهایی در برنامه‌ریزی درسی آموزش زبان شمرده می‌شود؛ جزء جدایی‌ناپذیر از فرایند یاددهی - یادگیری است و مطالعه انواع مختلف آزمون‌ها، به‌ویژه آزمون‌های پایان سال تحصیلی کشوری، از اهمیت فوق‌العاده‌ای برخوردار است. در همین راستا، در پژوهش حاضر تلاش خواهد شد؛ تا سؤالات امتحان نهایی فارسی پایه دوازدهم از سال ۱۳۹۳ تا ۱۴۰۳ بررسی شود و هم‌خوانی آن‌ها با سطوح مختلف طبقه‌بندی بلوم ارزیابی و تحلیل شود و به سؤالات ذیل پاسخ داده شود:

۱- سطوح سؤالات امتحانی نهایی فارسی پایه دوازدهم بر اساس طبقه‌بندی بلوم چگونه است؟

^۱ Bloom's Taxonomy

^۲ Anderson & Krathwohl

^۳ Vygotsky

- ۲- آیا سؤالات آزمون نهایی پایه دوازدهم درس فارسی، تمامی سطوح شناختی بلوم را پوشش می‌دهند؟
- ۳- کدام یک از سطوح شناختی بلوم، بیشترین تکرار را در سؤالات آزمون نهایی فارسی پایه دوازدهم دارند؟
- ۴- چه تغییراتی در طراحی سؤالات آزمون نهایی فارسی پایه دوازدهم می‌تواند پوشش بهتری از سطوح بالاتر شناختی بلوم را ارائه دهد؟

۲. چارچوب نظری

تفکر انتقادی^۱ به‌عنوان مجموعه‌ای از مهارت‌ها تعریف می‌شود که شامل تحلیل استدلال‌ها، استنتاج با استفاده از منطق استقرایی یا قیاسی^۲، ارزیابی و تصمیم‌گیری است (الی^۳، ۲۰۰۱). دیوئی^۴ (۱۹۹۱) که اولین تعریف مربوط به تفکر انتقادی را به او نسبت داده‌اند، این مفهوم را این‌گونه تعریف می‌کند: «بازنگری فعال، مستمر و دقیق باورها یا اشکال مفروض دانش، در پرتو دلایلی که از آن‌ها حمایت می‌کنند و نتایج بعدی که به آن‌ها منتهی می‌شوند.» با تمام این تعاریف، پژوهشگرانی همچون ریدل^۵ (۲۰۰۷) معتقد هستند که تفکر انتقادی، بیش از آن که یک «تعریف صریح» باشد، نیازمند «توضیح» است؛ توضیحی درباره ویژگی‌های کلیدی و عناصر اصلی آن، نه تلاش برای خلاصه کردن این مفهوم پیچیده در یک جمله. تفکر انتقادی؛ شامل حوزه‌های شناختی و غیرشناختی می‌شود. حوزه شناختی معمولاً با ارزیابی محصولات فکری مختلف تعریف می‌شود (پشیچ^۶، ۲۰۱۱) و غیرشناختی شامل عادت‌های ذهنی و گرایش‌هاست (فاسیونه^۷، ۱۹۹۰).

بلوم و همکاران (۱۹۷۱) نیز با توسعه یک نظام طبقه‌بندی برای سه حوزه شناختی، عاطفی و روانی - حرکتی، طبقه‌بندی اهداف آموزشی را آغاز کردند. ایده اصلی این طبقه‌بندی، به این صورت است که آنچه معلمان می‌خواهند دانش‌آموزان بدانند، می‌تواند در یک سلسله‌مراتب از ساده به پیچیده مرتب شود. سطوح به‌صورت پی در پی در نظر گرفته می‌شوند؛ به طوری که تسلط بر یک سطح پیش‌نیاز، رسیدن به سطح بعدی است (هیت^۸، ۲۰۱۱). سطوح اصلی که بلوم و همکارانش (۱۹۵۶) مطرح کردند، به ترتیب عبارتند از:

^۱ Critical Thinking

^۲ Inductive & Deductive Logic

^۳ Eli

^۴ Dewey

^۵ Riddell

^۶ Pešić

^۷ Facione

^۸ Huitt

(۱) دانش^۱ (۲) درک^۲ (۳) کاربرد^۳ (۴) تحلیل^۴ (۵) ترکیب^۵ (۶) ارزیابی^۶.

بعدها اندرسون و کراوول^۷ (۲۰۰۱) طبقه‌بندی بلوم را بازنگاری و اسم سطوح را از اسم به فعل تغییر دادند: سطح «دانش» به «یادآوردن یا به خاطر سپردن»^۸، سطح «درک» به «فهمیدن»^۹، سطح «کاربرد» به «به‌کار بستن»^{۱۰}، «تحلیل» به «تحلیل کردن»^{۱۱}، «ترکیب» به «خلق کردن یا آفریدن»^{۱۲} و «ارزیابی» به «ارزشیابی کردن»^{۱۳}. چهار سطح نخست، همانند سلسله‌مراتب بلوم و همکاران (۱۹۷۱) باقی ماندند؛ اما جایگاه آفریدن و ارزشیابی کردن با یکدیگر تغییر کرد و نسخه نهایی به‌عنوان طبقه‌بندی اصلاح شده بلوم معرفی شد. در نهایت، طبقه‌بندی بلوم از فرایندهای شناختی سطح پایین‌تر (دانش، درک) شروع شده و به سمت فرایندهای شناختی سطح بالاتر (تحلیل، کاربرد، ترکیب، ارزیابی) منتهی می‌شود (علی‌مراد و رزمی یگانه، ۱۴۰۳). نکته مهم این است که این فرایندهای شناختی سطح بالا، تفکر انتقادی فرد را تشکیل می‌دهند (کندی و همکاران^{۱۴}، ۱۹۹۱). تفکر انتقادی علاوه بر فوایدی که در تدریس و آموزش دارد (کسب نمرات بالاتر، بهبود درک دانش‌آموز و دانشجو، افزایش توانایی دانش‌آموز و دانشجو در انتقال مهارت‌های آموخته‌شده به موقعیت‌های تازه و ...)، در امور روزمره و حیاتی نیز کمک‌کننده است (تصمیم‌گیری و حل مسائل به شکل ماهرانه، انتخاب‌های مدیرانه در روابط انسانی و عرصه‌های مدنی و شخصی، بهبود توجه و مشاهده و ...) (الوالی^{۱۵}، ۲۰۱۱).

تعداد زیادی از پژوهشگران تفکر انتقادی؛ از جمله کندی^{۱۶} و همکاران، ۱۹۹۱؛ هالپرن، ۱۹۹۸^{۱۷} تأیید کرده‌اند که مهارت‌ها و توانایی‌های تفکر انتقادی، قابل آموزش هستند. برای این آموزش، سه روش مستقیم (موضوع‌محور)، غیرمستقیم (عمومی) و ترکیبی پیشنهاد شده است که به نظر می‌رسد؛ روش ترکیبی، بیشترین

^۱ Knowledge

^۲ Comprehension

^۳ Application

^۴ Analysis

^۵ Synthesis

^۶ Evaluation

^{۱۵} Anderson & Krathwohl

^۸ Remembering

^۹ Understanding

^{۱۰} Applying

^{۱۱} Analyzing

^{۱۲} Create

^{۱۳} Evaluate

^{۱۴} Kennedy

^{۱۵} Alwali

^{۱۰} Kennedy

^{۱۱} Halpern

تأثیر را بر مهارت‌ها و گرایش‌های دانش‌آموزان دارد. معمولاً توافق وجود دارد که تنها با یادگیری برنامه درسی مبتنی بر محتوا، کودکان نمی‌توانند به تفکر کنندگان بهتر، قادر به ارائه دلایل برای نتیجه‌گیری‌های خود، انعطاف‌پذیر و خلاق، حل‌کننده مسئله و تصمیم‌گیرنده خوب تبدیل شوند (ماسا^۱، ۲۰۱۴).

بنابراین، در این میان، نقش معلم در آموزش تفکر انتقادی (با تدریس مناسب و طراحی آزمون‌هایی که در آن‌ها فرایندهای شناختی سطح بالا سنجیده می‌شود) برجسته می‌شود.

درباره میزان استفاده هر کدام از فرایندهای شناختی، نظرات متفاوتی بیان شده است. با این حال، تقریباً بیشتر پژوهشگران بر این نکته تأکید دارند که مهارت‌های شناختی سطح بالا باید در طراحی سؤالات آزمون‌ها، تعداد لازم را شامل شود. سؤالات سطح پایین، هم‌نیاز هستند؛ زیرا پایه‌های مهمی برای پیشبرد مغز برای رسیدن به سطوح بالاتر هستند.

البته نتایج کار ممکن است همیشه آن‌طور که معلمان می‌خواهند، ظاهر نشود. طراحان سؤال ممکن است چنین درک کنند که برخی پرسش‌ها نیازمند مهارت‌های تفکر سطوح بالاتر برای پردازش هستند؛ در حالی که پاسخ‌دهندگان تنها با بهره‌گیری از مهارت‌های تفکر سطوح پایین‌تر می‌توانند پاسخ صحیح ارائه دهند. به همین ترتیب، پرسش‌هایی که در ظاهر در سطح پایین‌تر قرار دارند، ممکن است در عمل، نیازمند مهارت‌های تفکر سطوح بالاتر برای پاسخ‌دهی صحیح باشند (زیدی^۲ و همکاران، ۲۰۱۸).

بررسی آزمون فارسی پایه دوازدهم، از جهت این‌که آخرین سال تحصیل در دوره متوسطه دوم است و پس از آن، دانش‌آموزان وارد دانشگاه و مؤسسات آموزش عالی یا بازار کار می‌شوند، از اهمیت ویژه‌ای برخوردار است. پژوهشگرانی همچون هک ورث^۳ (۲۰۱۰) بر این نکته تأکید کرده‌اند؛ برای آن‌که دانشجویان بتوانند در حرفه‌های آینده خود از مهارت‌های تفکر انتقادی بهره بگیرند، ابتدا باید بیاموزند که چگونه این مهارت‌ها را توسعه دهند. به‌تبع، این مهارت باید نحوه توسعه این مهارت در مدارس یاد داده شود. معلمان باید دانش‌آموزان را برای این فرایند آماده سازند و سؤالات را به نوعی طراحی کنند که از سطوح پایه تا سطوح بالای شناختی را شامل شود. باید دید که سؤالات درس مهم زبان و ادبیات فارسی، شامل سطوح شناختی از سطوح پایین به بالا می‌شود یا طراحی سؤالات به نحوی پیشرفته که یکی از سطوح، سهم بیشتری از دیگر سطوح دارد؟

^۱ Massa

^۲ Zaidi

^۳ Hackworth

۳. پیشینه پژوهش

ارزیابی کیفیت سؤالات آزمون‌های نهایی، به‌ویژه در دروسی مانند فارسی، از اهمیت بالایی در تعیین سطح یادگیری دانش‌آموزان و کارایی فرآیند آموزش برخوردار است. با توجه به تأثیر قابل توجه طبقه‌بندی شناختی بلوم در طراحی سؤالات با سطوح مختلف، پیچیدگی و ارزیابی یادگیری عمیق‌تر از صرف یادآوری مطالعات این پژوهش با هدف ارزیابی سؤالات امتحانات نهایی درس فارسی از منظر این طبقه‌بندی انجام شده است. پیشینه مطالعات نشان می‌دهد که پژوهش‌های متعددی به بررسی کیفیت سؤالات امتحانی در حوزه‌های مختلف پرداخته‌اند؛ اما تمرکز روی تطابق سؤالات امتحانات نهایی درس فارسی با سطوح مختلف طبقه‌بندی بلوم و تحلیل نقاط قوت و ضعف آن‌ها از نظر پوشش دامنه شناختی، تاکنون در مورد امتحانات نهایی درس فارسی مدارس کشور به‌طور جامع مورد بررسی قرار نگرفته است. بنابراین، این پژوهش با هدف پرکردن این خلأ و ارائه راهکارهایی برای بهبود کیفیت سؤالات امتحانات نهایی درس فارسی، طراحی و اجرا شده است. در ادامه، به برخی از پژوهش‌هایی که درباره آزمون‌های نهایی بعضی دروس انجام شده، اشاره می‌شود:

سیاسی (۱۳۸۵) در پژوهش خود با عنوان «بررسی و تحلیل سطوح حیطة شناختی و شاخص‌های روان‌سنجی امتحان نهایی دروس عربی، حسابان و زیست‌شناسی دانش‌آموزان دختر پایه سوم متوسطه، در سه منطقه متفاوت اقتصادی - اجتماعی استان خوزستان» به این نتیجه رسیده است که امتحانات، از دقت و ثبات کافی برخوردار هستند و بین اهداف و محتوای دروس، توازن مطلوبی برقرار است.

میرآقایی و همکاران (۱۳۹۴) در مقاله «بررسی و تحلیل شاخص‌های روان‌سنجی و سطوح حیطة شناختی سؤالات امتحانات نهایی دروس ریاضیات و علوم پایه سوم راهنمایی شهرستان خرم‌آباد» به این نتیجه رسیده‌اند که سؤالات، دارای پایایی و روایی محتوایی هستند و درس ریاضیات به سطوح بالای اندازه‌گیری؛ یعنی ترکیب و ارزشیابی اختصاص نداشتند.

فلسفی‌نژاد و همکاران (۱۳۹۵) در مقاله «ویژگی‌های روان‌سنجی امتحانات نهایی سال سوم متوسطه و قابلیت آن‌ها در گزینش داوطلبان ورود به دوره‌های کارشناسی» سؤالات آزمون نهایی درس زیست‌شناسی و فارسی را بر اساس CTT و IRT تحلیل کرده‌اند و به این نتیجه رسیده‌اند که آزمون ادبیات فارسی، فاقد سؤالات دشوار است و بیشتر سؤالات آسان آن، در نیمه اول آزمون قرار دارد. همچنین آزمون زیست‌شناسی نیز با سؤالات آسان آغاز می‌شود؛ هرچند که سؤالات آسان کمی دارد و اصل ترتیب سؤالات از ساده به مشکل در دو آزمون نامبرده، به نسبت رعایت شده است.

در دهه‌های جدید، طبقه‌بندی بلوم به‌عنوان یکی از ابزارهای اصلی در تحلیل و طراحی اهداف آموزشی و ارزیابی‌ها شناخته شده است. این طبقه‌بندی به معلمان و طراحان آزمون‌ها این امکان را می‌دهد که اهداف

آموزشی خود را به‌طور مؤثرتر تنظیم کرده و ارزیابی‌هایی طراحی کنند که نه‌تنها اطلاعات پایه را می‌سنجند؛ بلکه تفکر تحلیلی، انتقادی و خلاقانه را نیز ترویج می‌دهند.

با این حال، مطالعات مختلف نشان می‌دهند که در بسیاری از نظام‌های آموزشی، آزمون‌ها اغلب بر سطوح پایین‌تر شناختی متمرکز هستند و از این‌رو، فرصتی برای ارتقای مهارت‌های تفکر سطح بالاتر فراهم نمی‌کنند. در این راستا، بسیاری از پژوهشگران به تحلیل و بررسی توزیع سطوح مختلف طبقه‌بندی بلوم در آزمون‌های نهایی پرداخته‌اند و به این نتیجه رسیده‌اند که سؤالات آزمون‌ها عمدتاً در سطوح پایین‌تری مانند دانش و درک متمرکز هستند. مطالعات مختلفی که در این زمینه انجام شده، به‌طور مشترک بر این نکته تأکید دارند که طراحی آزمون‌ها بیشتر به ارزیابی سطوح پایین‌تر شناختی پرداخته و از توجه به سطوح بالاتر که شامل تحلیل، ارزیابی و خلاقیت می‌شود، غفلت کرده‌اند. برای مثال؛ شوهامی (۲۰۱۳) در پژوهش خود نشان داده است که آزمون‌های زبان، به‌طور عمده بر حفظ و یادآوری اطلاعات تمرکز دارند و هیچ تأکیدی بر تفکر انتقادی و حل مسائل پیچیده ندارند. او معتقد است که چنین آزمون‌هایی، نه‌تنها موجب محدود شدن ظرفیت‌های شناختی دانش‌آموزان می‌شود؛ بلکه آن‌ها را به سمت یادگیری سطحی سوق می‌دهند که توانایی‌های تحلیلی و نوآورانه را تضعیف می‌کند.

چنگ و کرتس^۱ (۲۰۰۴) نیز در تحقیق خود مفهوم «واش‌بک»^۲ را معرفی کرده‌اند که نشان می‌دهد؛ طراحی آزمون‌ها نه‌تنها بر یادگیری دانش‌آموزان اثر می‌گذارد؛ بلکه باعث شکل‌گیری روش‌های تدریس خاصی نیز می‌شود. آن‌ها می‌گویند که زمانی که آزمون‌ها بر سطوح پایین‌تر شناختی تأکید دارند، تدریس نیز به‌طور غیرمستقیم، به سمت تقویت حافظه‌سپاری و یادآوری اطلاعات سوق پیدا می‌کند و این امر، مانع از توسعه مهارت‌های تحلیلی و تفکر انتقادی در دانش‌آموزان می‌شود.

در همین راستا، کندی و همکارانش (۱۹۹۱) بر اهمیت تفکر انتقادی در فرآیند آموزشی تأکید کرده‌اند و معتقدند که این مهارت‌ها باید از طریق طراحی آزمون‌هایی که دانش‌آموزان را به تحلیل و ارزیابی اطلاعات و حتی خلق محتوای جدید وادار می‌کنند، پرورش یابند. آن‌ها اشاره می‌کنند که عدم توجه به تفکر انتقادی در طراحی آزمون‌ها می‌تواند منجر به محدودیت‌های جدی در فرآیند یادگیری شود؛ چرا که دانش‌آموزان تنها قادر به حفظ اطلاعات و پاسخ دادن به سؤالات ساده خواهند بود؛ بدون این‌که توانایی تحلیل و ارزیابی مسائل را کسب کنند.

به نظر می‌رسد که نیاز به یک بازنگری جدی در طراحی آزمون‌ها و شیوه‌های ارزیابی وجود دارد. باید تأکید بیشتری بر سطوح بالاتر طبقه‌بندی بلوم؛ مانند تحلیل، ارزیابی و ترکیب قرار گیرد. این تغییرات نه‌تنها می‌تواند

^۱ Cheng & Curtis

^۲ Washback

به ارتقای تفکر انتقادی و حل مسائل پیچیده کمک کند؛ بلکه موجب تقویت مهارت‌های شناختی و آماده‌سازی دانش‌آموزان برای مواجهه با چالش‌های دنیای واقعی خواهد شد.

۴. روش‌شناسی پژوهش

پژوهش حاضر با هدف بررسی سؤالات امتحانات نهایی کشوری درس فارسی پایه دوازدهم علوم انسانی در ایران، طی سال‌های ۱۳۹۳ تا ۱۴۰۳ انجام شده است. این پژوهش، از نوع توصیفی - تحلیلی است و با بهره‌گیری از روش تحلیل محتوا و تلفیق رویکردهای کمی و کیفی انجام گرفته است. در این راستا، داده‌های پژوهش؛ شامل مجموعه‌ای از سؤالات امتحانات نهایی درس فارسی است که از سوی وزارت آموزش و پرورش منتشر شده‌اند. این سؤالات از آنجا که دارای ساختار، محتوا و اهداف آموزشی هماهنگ و استاندارد هستند، مبنای مناسبی برای تحلیل علمی و مقایسه در طول سال‌های مورد بررسی محسوب می‌شوند.

ابزار گردآوری داده‌ها؛ شامل بانک سؤالات امتحانات کشوری و آرشیو رسمی مرکز ارزشیابی و تضمین کیفیت نظام آموزش و پرورش (<https://azmoon.medu.gov.ir/fa>) بود. تمامی سؤالات استخراج شده پس از گردآوری، به صورت نظام‌مند بر اساس سطوح شش‌گانه طبقه‌بندی شناختی بلوم؛ شامل «دانش»، «درک»، «کاربرد»، «تحلیل»، «ارزیابی» و «ترکیب» طبقه‌بندی شدند. در این تحقیق، دو کدگذار مستقل، مسئول طبقه‌بندی سؤالات بر اساس طبقه‌بندی بلوم بودند. برای اطمینان از دقت فرآیند کدگذاری، یک سوم از سؤالات، ابتدا توسط هر دو کدگذار کدگذاری شد. ضریب کاپا (Cohen's Kappa coefficient) برای توافق بین کدگذاران ۰/۷۸ به دست آمد که نشان‌دهنده سازگاری قابل توجه است. پس از رفع هرگونه اختلاف، کدگذاران به کدگذاری تمام سؤالات ادامه دادند. برای هر سال تحصیلی، فراوانی و درصد سؤالات در هر سطح شناختی محاسبه و سپس با سال‌های دیگر مقایسه شد؛ تا روند تغییرات در تأکید سطوح شناختی مشخص شود.

در گام بعدی، داده‌های گردآوری شده، به منظور تحلیل کمی، در نرم‌افزار SPSS نسخه ۲۵ وارد شد. نخست، تحلیل توصیفی برای نمایش توزیع سؤالات در هر سطح شناختی انجام شد؛ سپس، به منظور بررسی ارتباط آماری بین سال‌های تحصیلی و سطوح مختلف طبقه‌بندی بلوم، از آزمون کای اسکوائر (Chi-Square) استفاده شد. با توجه به این که در برخی از سلول‌های جدول‌های تقاطعی، فراوانی مورد انتظار کمتر از پنج بود، جهت افزایش دقت نتایج، از آزمون دقیق فیشر (Fisher's Exact Test) به عنوان آزمون مکمل استفاده شد. این آزمون برای تأیید معناداری روابط بین متغیرها در داده‌هایی با حجم نمونه نسبتاً کوچک یا توزیع نامتوازن مناسب‌تر است.

به منظور اطمینان از دقت تحلیل، دو سطح شناختی «ارزیابی» و «ترکیب» که در بسیاری از آزمون‌ها دارای فراوانی صفر یا بسیار پایین بودند، به صورت یک سطح ترکیبی در نظر گرفته شدند؛ تا از بروز خطای آماری در

نتایج جلوگیری شود. سپس مقادیر فراوانی، درصد و آزمون‌های استنباطی در قالب جدول‌ها و نمودارهای مقایسه‌ای تنظیم و تفسیر شدند. در این تحقیق، شامل چندین مرحله بود که به‌طور دقیق برای شناسایی روندها و تغییرات در سطح شناختی سؤالات آزمون‌ها انجام شد. همان‌طور که بیان شد؛ ابتدا تمامی سؤالات آزمون‌ها بر اساس شش سطح شناختی بلوم (۱۹۵۶) طبقه‌بندی شدند. این طبقه‌بندی به‌منظور شناسایی مهارت‌های شناختی مختلف مورد ارزیابی قرار گرفت. در سطح اول، سؤالات مربوط به «دانش» قرار گرفتند که به یادآوری یا بازیابی اطلاعاتی که قبلاً آموخته شده است، اشاره دارد. این سؤالات معمولاً شامل تعریف اصطلاحات کلیدی، فهرست کردن مراحل یک فرایند یا تکرار اطلاعات هستند. در سطح بعدی، «درک» قرار دارد که شامل پردازش اطلاعات جدید است. سؤالات در این دسته، نیاز به فهم و تفسیر اطلاعات دارند. پس از آن، در سطح «کاربرد»، یادگیرنده باید توانایی حل مسئله جدید با استفاده از اطلاعات آموخته‌شده را داشته باشد. در این سطح، سؤالات به‌گونه‌ای طراحی می‌شوند که دانش‌آموزان بتوانند مفاهیم آموخته‌شده را در موقعیت‌های جدید به‌کار ببرند. سطح «تحلیل» نیز به روابط میان اجزا و مفاهیم مختلف می‌پردازد و یادگیرندگان باید قادر باشند، تمایز قائل شوند و مقایسه‌ها و آزمایش‌ها را انجام دهند. در سطح «ترکیب»، یادگیرندگان باید به خلاقیت پرداخته و محصولات جدید و منحصربه‌فردی بسازند. اهداف یادگیری در این سطح، به‌طور معمول؛ شامل طراحی یک طرح یا تولید اطلاعات جدید است. در نهایت، سطح «ارزیابی» قرار دارد که در آن سؤالات نیازمند قضاوت و ارزیابی دانش بر اساس معیارهای مشخص هستند. این قضاوت‌ها معمولاً مبتنی بر محتوا و نقد نوشتاری هستند. پس از طبقه‌بندی سؤالات، داده‌ها جمع‌آوری و بر اساس سال‌های تحصیلی و سطوح شناختی دسته‌بندی شدند.

۵. ارائه و واکاوی داده‌ها

جدول (۱) نمونه‌ای از سؤالات آزمون‌های نهایی بر اساس طبقه‌بندی شناختی بلوم را نشان می‌دهد:

جدول (۱). نمونه‌ای از سؤالات آزمون‌های نهایی براساس طبقه‌بندی شناختی بلوم (۱۹۵۶)

سطح طبقه بندی بلوم	توضیحات	کلمات کلیدی	نمونه سؤالات مربوط در آزمون
دانش	دانش آموز پاسخ را به یاد می‌آورد و بدون تغییر و تفسیر جواب را بازیابی می‌کند.	نام بردن، مشخص کردن، بیان کردن، تعریف کردن و ...	دی ماه ۹۳ - سؤال ۳۵: نام اثر تحقیقی ۵ جلدی از «ایرج افشار یزدی» را بنویسید.
درک	در این بخش دانش‌آموز باید معنای ابیات را بفهمد و عبارات مشخص شده را تفسیر کرده و توضیح بدهد	توضیح دادن، خلاصه کردن، بازگو کردن، توصیف کردن و ...	دی ماه ۹۶ - سؤال ۲۷: آیه «و من یتوکل علی الله فهو حسبه» با کدام بیت زیر ارتباط معنایی دارد؟ الف) گل و گلزار خوش آید کسی را که گلزار و گلستانش تو باشی ب) چه باک آید ز کس آن را که او را

	یا گزینه مناسب را انتخاب کند.		نگهدار و نگهبانش تو باشی (ج) همه شادی و عشرت است ای دوست در آن خانه که مهمانش تو باشی (د) خوشا چشمی که رخسار تو ببیند خوشا ملکی که سلطانش تو باشی
کاربرد	علاوه بر بازیابی و فهم، دانش‌آموز باید دانش خاصی را به موقعیت جدیدی منتقل و استفاده کند.	استفاده کردن، ساختن، شناسایی کردن، شبیه-سازی کردن، اعمال کردن و ...	دی ماه ۱۴۰۳ - سؤال ۲۲: آرایه درست مربوط به هر بیت را از ستون مقابل آن انتخاب کنید (یک آرایه در ستون «ب» اضافی است). الف) خیس خون داغ سهراب و سیاوش‌ها/ روکش تابوت تختی هاست ب) شانه می‌آید به کار زلف در آشفنگی/ آشنایان را در ایام پریشانی بپرس ج) رخ شاه کاووس پر از شرم دید/ سخن گفتنش با پسر نرم دید د) تا درد و ورم فرو نشیند/ کافور بر آن ضماد کردند ۱- حس آمیزی ۲- حسن تعلیل ۳- ایهام ۴- اسلوب معادله ۵- پارادوکس
تحلیل	دانش‌آموز باید ساختار منطقی و رابطه علت و معلول متن را شناسایی کند، نه صرفاً معنی یا اطلاعات را یادآوری کند.	ساده‌سازی کردن، مقایسه کردن، بررسی کردن، تفسیر کردن و ...	دی ماه ۹۷- سؤال ۳۱: در بیت زیر از داستان «طوطی و بازرگان»، طوطی چه پیامی به همتای خویش فرستاد؟ «طوطی زان طوطیان لرزید بس اوفتاد و مرد و بگسستش نفس»
ترکیب	دانش‌آموز، دانش به دست آمده را تلفیق و چیز منحصر به فردی تولید می‌کند.	ایجاد کردن، ترکیب کردن، پیشنهاد کردن، خلق کردن، توسعه دادن و ...	خردادماه ۱۴۰۰- سؤال ۱۱: در یک گروه اسمی، واژه اسب را با یک وابسته پیشین و یک وابسته پسین بکار ببرید.
ارزیابی	دانش‌آموز درباره ارزش و اعتبار موضوع‌های مختلف به ارزیابی و قضاوت می‌پردازد.	کشف کردن، اثبات کردن، ارزیابی کردن، قضاوت کردن، نقد کردن و ...	دی ماه ۹۹- سؤال ۳۴: سروده زیر، بیانگر کدام ویژگی‌های عشق است؟ «چه حرف تازه‌ای برای گفتن مانده است، یا چه چیز تازه‌ای برای نوشتن/ که بتواند عشق مرا یا سجایای ارزشمند تو را بازگو کند؟»

این دسته‌بندی به منظور شناسایی روندها و تغییرات در توزیع سؤالات و ارزیابی‌ها در طول سال‌های مختلف انجام شد. نتایج به دست آمده از تحلیل داده‌ها، به شرح ذیل و در دو بخش (۱. تحلیل آمار توصیفی ۲. تحلیل آمار استنباطی) ارائه می‌شود.

۵.۱. نتایج تحلیل آمار توصیفی

هدف از این تحلیل، بررسی توزیع سطوح طبقه‌بندی بلوم در طول سال‌ها و نیم‌سال‌ها و شناسایی روندها و الگوها در توسعه مهارت‌های شناختی دانش‌آموزان است. توزیع فراوانی هر سطح، که به تفکیک سال‌ها دسته‌بندی شده است، دیدگاه دقیقی از چگونگی تغییرات مهارت‌های شناختی دانش‌آموزان در طول مطالعه ارائه می‌دهد. علاوه بر این، تعداد کل هر سطح از طبقه‌بندی بلوم در تمام سال‌ها، دید کلی‌تری از روند کلی توسعه مهارت‌های شناختی در طول دوره مطالعاتی فراهم می‌آورد.

جدول (۲): توزیع فراوانی سطوح طبقه‌بندی بلوم بر اساس سال

مرحله سال	دانش	درک	کاربرد	تحلیل	ترکیب	ارزیابی	تعداد سؤالات هر سال	
۱۳۹۳	۲۱	۲۱	۰	۳	۰	۰	۴۵	دی ماه
۱۳۹۴	۱۳	۲۶	۰	۱	۰	۰	۴۰	
۱۳۹۵	۱۷	۲۲	۰	۲	۰	۰	۴۱	
۱۳۹۶	۱۷	۲۵	۰	۲	۰	۰	۴۴	
۱۳۹۷	۶	۲۵	۲	۲	۰	۰	۳۵	
۱۳۹۸	۳	۲۶	۰	۴	۰	۰	۳۳	
۱۳۹۹	۴	۲۴	۰	۵	۰	۱	۳۴	
۱۴۰۰	۵	۱۸	۰	۶	۰	۰	۲۹	
۱۴۰۱	۱۱	۲۴	۱	۲	۰	۰	۳۸	
۱۴۰۲	۹	۲۱	۰	۲	۰	۰	۳۲	
۱۴۰۳	۹	۲۳	۱	۳	۰	۰	۳۶	
۱۳۹۳	۱۶	۳۱	۰	۱	۰	۰	۴۸	خرداد ماه
۱۳۹۴	۱۷	۳۲	۰	۰	۰	۰	۴۹	
۱۳۹۵	۸	۲۳	۰	۲	۰	۰	۳۳	
۱۳۹۶	۱۳	۳۱	۰	۰	۰	۰	۴۴	
۱۳۹۷	۱۳	۲۷	۲	۱	۰	۰	۴۳	
۱۳۹۸	۷	۲۸	۳	۳	۰	۰	۴۱	
۱۳۹۹	۸	۲۹	۰	۳	۱	۱	۴۲	
۱۴۰۰	۴	۱۷	۱	۶	۱	۰	۲۹	
۱۴۰۱	۷	۱۳	۱	۵	۰	۱	۲۷	
۱۴۰۲	۸	۲۵	۱	۴	۰	۰	۳۸	
۱۴۰۳	۱۰	۲۱	۱	۹	۰	۰	۴۱	
جمع	۲۲۶	۵۳۲	۱۳	۶۶	۲	۳	۸۴۲	
درصد	۲۶.۸۴%	۶۳.۱۸%	۱.۵۴%	۷.۸۴%	۰.۲۴%	۰.۳۶%	۱۰۰%	

جدول (۲) نمونه‌ای از سؤالات آزمون‌های نهایی، بر اساس طبقه‌بندی شناختی بلوم (۱۹۵۶) را نشان می‌دهد. این تقسیم‌بندی، امکان درک دقیق‌تری از توزیع مهارت‌های شناختی را فراهم می‌آورد. همان‌طور که در جدول (۲) مشاهده می‌شود، داده‌های هر سال، به‌طور دقیق بر اساس سطح طبقه‌بندی بلوم، از ابتدایی‌ترین سطح (دانش) تا سطوح پیچیده‌تر (ارزیابی و ترکیب)، دسته‌بندی شده‌اند. این روش، به شناسایی نوسانات یا تغییرات قابل توجه در تمرکز ارزیابی‌ها در طول زمان کمک می‌کند.

برای ارائه دیدگاه جامع‌تری، جدول (۳) تعداد کل هر سطح از طبقه‌بندی بلوم را در طول تمام سال‌ها به‌طور خلاصه نشان می‌دهد. این داده‌های تجمعی امکان بررسی روندهای بلندمدت در توسعه مهارت‌های شناختی را فراهم می‌آورد و الگوهای قابل توجه یا نواحی که نیاز به توجه بیشتری دارند را نشان می‌دهد. با مقایسه تعداد کل‌ها، می‌توان ارزیابی کرد که آیا تأکید مداوم بر مهارت‌های خاصی بوده است؛ یا این‌که در طول زمان، تمرکز روی مهارت‌های خاصی تغییر کرده است.

جدول (۳): تعداد کل سطوح طبقه‌بندی بلوم در طول تمام سال‌ها

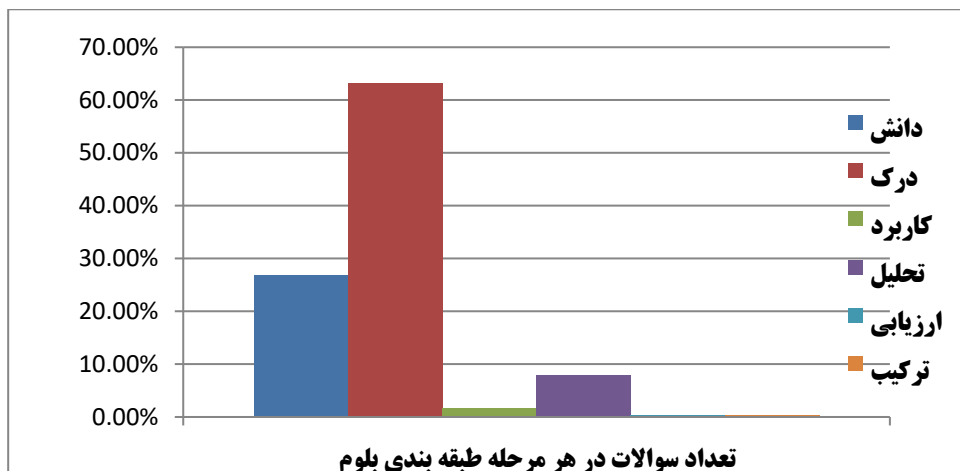
سطوح طبقه بندی بلوم	تعداد سؤالات در این مرحله
دانش	۲۲۶
درک	۵۳۲
کاربرد	۱۳
تحلیل	۶۶
ارزیابی	۳
ترکیب	۲

این جداول، زمانی که به‌طور همزمان تحلیل شوند، بینش ارزشمندی از روند ارزیابی‌های دانش‌آموزان ارائه می‌دهند و می‌توانند به بهبودهای آینده در برنامه درسی کمک کنند؛ تا توسعه مهارت‌های شناختی در تمام سطوح طبقه‌بندی بلوم به‌طور متوازن صورت گیرد.

نتایج این بررسی نشان می‌دهد که توزیع سؤالات آزمون‌ها به‌شدت نامتوازن است و بیشتر تمرکز بر سطوح ابتدایی قرار دارد. سطح دانش که ابتدایی‌ترین سطح است، ۲۶٫۸۴٪ سؤالات را به خود اختصاص داده است. سطح درک که کمی پیچیده‌تر از سطح دانش است، بیشترین سؤالات را به خود اختصاص داده است؛ یعنی ۶۳٫۱۸٪ سؤالات. در این میان، سطح کاربرد ۱٫۵۴٪، سطح تحلیل ۷٫۸۴٪ و دو سطح ترکیب و ارزیابی، تنها ۰٫۲۴٪ و ۰٫۳۶٪ درصد را به خود اختصاص داده‌اند.

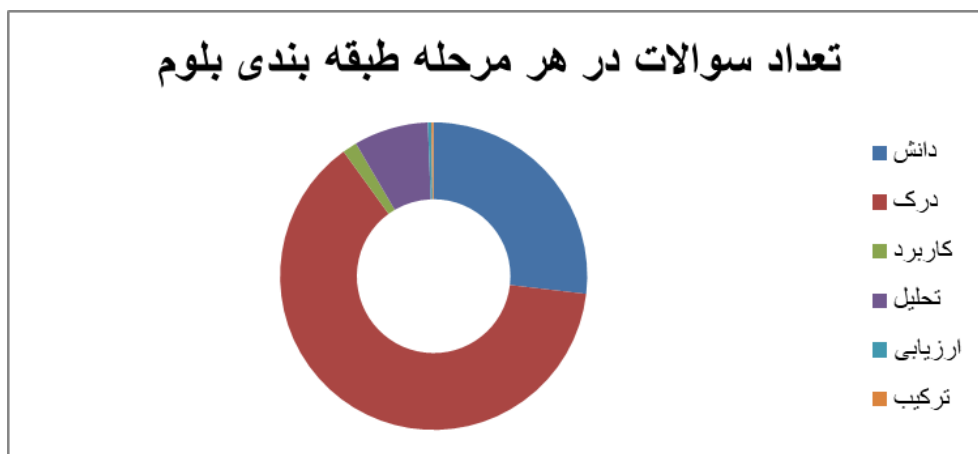
آمار نشان می‌دهد که بیشتر پرسش‌ها به حافظه‌محور بودن و بازشناسی اطلاعات متکی هستند. با توجه به جدول، مشخص می‌شود که توانایی‌های تفکر انتقادی، استدلال و خلاقیت، کمتر مورد توجه بوده‌اند. مقایسه سال‌ها (۱۳۹۳ - ۱۴۰۳) نشان می‌دهد که تغییر معناداری در طراحی سؤالات رخ نداده است و تقریباً همیشه

تأکید روی دانش و درک بوده است. نمودارهای ۱ و ۲ توزیع سطوح طبقه‌بندی بلوم در آزمون‌های نهایی را نشان می‌دهند. نمودار اول؛ نسبت سؤالات بر اساس سطوح مختلف طبقه‌بندی بلوم را مقایسه می‌کند. نمودار دوم؛ تعداد سؤالات هر سطح از طبقه‌بندی بلوم را نمایش می‌دهد. این نمودارها روند تمرکز بر مهارت‌های شناختی سطح پایین را نشان می‌دهند و بینش‌هایی را درباره الگوهای مشاهده‌شده در ارزیابی‌ها ارائه می‌دهند.



نمودار ۱: نسبت سؤالات بر اساس سطوح طبقه‌بندی بلوم

نمودار ۱ نشان‌دهنده نسبت سؤالات مربوط به هر سطح از طبقه‌بندی بلوم در آزمون‌های نهایی است. همان‌طور که مشاهده می‌شود، درصد زیادی از سؤالات در سطوح دانش و درک قرار دارند؛ در حالی که سطوح کاربرد، تحلیل، ارزیابی و خلق درصد کمتری از سؤالات را تشکیل می‌دهند. این امر نشان‌دهنده تمرکز بیشتر آزمون‌ها بر روی مهارت‌های شناختی سطح پایین است و تأکید کمتری بر تفکر تحلیلی و حل مسئله دارد.



نمودار ۲: تعداد سؤالات در هر سطح از طبقه‌بندی بلوم

نمودار دایره‌ای (۲) تعداد سؤالات هر سطح از طبقه‌بندی بلوم را نشان می‌دهد. همان‌طور که در این نمودار مشاهده می‌شود، بخش عمده‌ای از سؤالات، به سطوح دانش و درک اختصاص دارد؛ در حالی که سطوح به‌کارگیری و تحلیل تعداد کمتری از سؤالات را شامل می‌شوند. این نمودار به‌طور واضح‌تر، روند تأکید بیشتر بر سطوح پایین‌تر شناختی و کمبود توجه به مهارت‌های تفکر انتقادی و تحلیلی را نشان می‌دهد.

۲.۵. نتایج تحلیل آمار استنباطی

پس از بررسی آمار توصیفی و توزیع سطوح طبقه‌بندی بلوم در آزمون‌های نهایی، برای تحلیل روابط معنادار بین متغیرها و بررسی تفاوت‌های مشاهده‌شده در داده‌ها، از آمار استنباطی استفاده شد؛ به‌ویژه، آزمون کای اسکوئر برای بررسی ارتباطات بین سال‌های تحصیلی و سطوح مختلف طبقه‌بندی بلوم به کار گرفته شد. جدول (۴) خلاصه پردازش داده‌ها را نشان می‌دهد.

جدول ۴. خلاصه پردازش داده‌های آزمون کای اسکوئر

تعداد موارد معتبر	درصد	تعداد موارد ناقص	درصد	تعداد کل	درصد
۱۱۰	۱۳,۱٪	۷۳۲	۸۶,۹٪	۸۴۲	۱۰۰

همان‌طور که جدول (۴) نشان می‌دهد؛ این تحلیل شامل ۱۱۰ مورد معتبر (۱۳,۱٪ از کل داده‌ها) بود و ۷۳۲ مورد ناقص (۸۶,۹٪ از کل) وجود داشت. با توجه به وجود فراوانی‌های مورد انتظار کوچک در جدول، جدول تقاطعی (۱۱۰ سلول دارای فراوانی مورد انتظار کمتر از ۵ بود)، از آزمون دقیق فیشر برای تأیید معناداری آماری نتایج استفاده شد. جدول نتایج آزمون کای اسکوئر برای رابطه بین سطوح طبقه‌بندی بلوم و سال‌ها/نیم‌سال‌های تحصیلی را نشان می‌دهد.

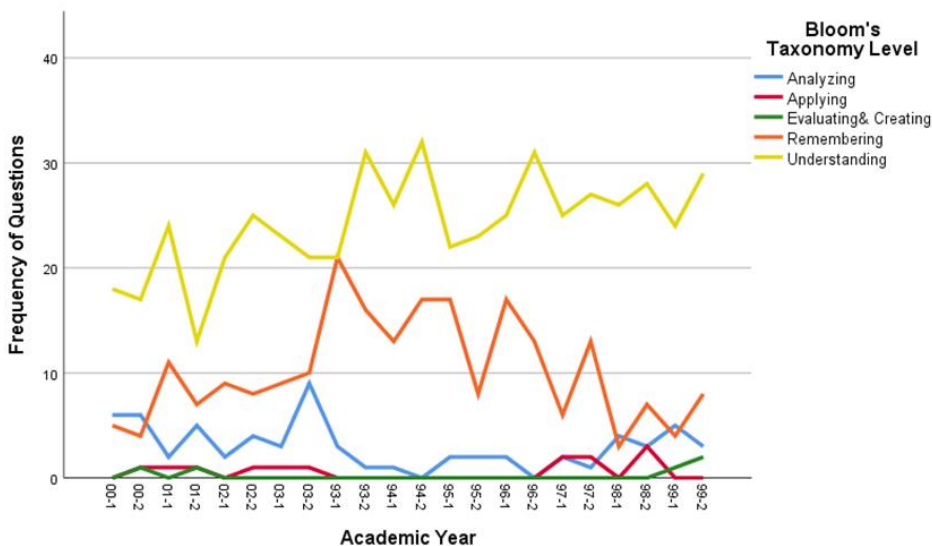
جدول ۵. نتایج آزمون کای اسکوئر برای رابطه بین سطوح طبقه‌بندی بلوم و سال/نیم‌سال تحصیلی

مقدار Exact Sig (1sided)	مقدار Exact Sig (2-Sided)	مقدار Asymptotic Sig (2-sided)	مقدار df	value	
	. ^b	۱.000	۸۴	.000 ^a	کای اسکوئر پیرسون
	۱.000	۱.000	۸۴	.000	نسبت احتمال
	۱.000			۱۷,۰۶۵	آزمون دقیق فیشر
. ^b	. ^b	۱.000	۱	.000	آزمون ارتباط خطی به خطی
				۱۱۰	داد موارد معتبر

^a ۱۱۰ سلول (۱۰۰٪) دارای شمارش مورد انتظار کمتر از ۵ بودند که نشان‌دهنده لزوم استفاده از آزمون دقیق فیشر به جای آزمون کای اسکوئر است.

^b محاسبه انجام نشد؛ زیرا حافظه کافی موجود نیست.

نتایج آزمون دقیق فیشر نشان داد که بین سطوح طبقه‌بندی بلوم و سال‌ها/نیم‌سال‌ها یک ارتباط آماری معنادار وجود ندارد ($\chi^2(84) = 17.065, p > 0.05$). نمودار (۳) روند تغییرات سطوح طبقه‌بندی بلوم در سال‌های تحصیلی مختلف را نشان می‌دهد.



نمودار (۳). روند تغییرات سطوح طبقه‌بندی بلوم در سال‌های تحصیلی مختلف

نمودار (۳) روند تغییرات تعداد سطوح مختلف طبقه‌بندی بلوم (دانش، درک، کاربرد و غیره) را در طول سال‌های تحصیلی مختلف نشان می‌دهد. در این نمودار، هر خط نمایانگر یکی از سطوح طبقه‌بندی بلوم است و نشان‌دهنده تعداد پرسش‌هایی است که به هر سطح اختصاص داده شده است. تغییرات موجود در این نمودار می‌تواند نشان‌دهنده تغییرات در رویکرد ارزیابی و تمرکز بر سطوح مختلف شناختی در آزمون‌ها در طول سال‌ها باشد. به‌طور خاص، سطوح بالاتر طبقه‌بندی بلوم؛ مانند کاربرد و تحلیل در سال‌های بعدی افزایش پیدا کرده‌اند که نشان‌دهنده تغییر در تمرکز ارزیابی به سمت تفکر انتقادی و حل مسئله است؛ اما نتایج نشان می‌دهد که این تغییرات، رابطه معناداری با سال مورد بررسی ندارد.

۶. تفسیر داده‌ها

نتایج این مطالعه که با هدف تحلیل توزیع سطوح طبقه‌بندی بلوم در آزمون‌های نهایی درس فارسی پایه دوازدهم در سال‌های ۱۳۹۳ تا ۱۴۰۳ انجام شد، دیدگاه‌های ارزشمندی را در خصوص روند مهارت‌های شناختی مورد تأکید در ارزیابی‌های ملی ارائه می‌دهد. این مطالعه با استفاده از ترکیب آمار توصیفی و آمار استنباطی،

شامل آزمون کای اسکوتر و آزمون دقیق فیشر، رابطه بین سال‌های تحصیلی و توزیع سطوح طبقه‌بندی بلوم در آزمون‌ها را بررسی کرد.

نتایج این مطالعه به وضوح نشان می‌دهد که مهارت‌های شناختی سطح پایین‌تر (دانش و درک) در آزمون‌ها غالب است؛ در حالی که مهارت‌های شناختی سطح بالاتر (کاربرد، تحلیل، ارزیابی و خلق) توجه کمتری را به خود جلب کرده‌اند. این الگو نشان می‌دهد که با وجود تأکید فزاینده بر تفکر انتقادی و مهارت‌های شناختی بالاتر در سیاست‌ها و طراحی برنامه درسی (همان‌طور که بلوم، اندرسون و کراوول، ۲۰۰۱ اشاره کرده‌اند)، آزمون‌های نهایی همچنان بیشتر بر حافظه‌سپاری اطلاعات و درک سطحی تأکید دارند. این یافته‌ها با تحقیقات فلسفی نژاد و همکاران (۲۰۱۶) همخوانی دارد که به کمبود سؤالات با سطوح شناختی بالا در آزمون‌های فارسی اشاره کرده‌اند و همچنین شوهمی (۲۰۱۳) که تأکید کرده است که آزمون‌ها باید علاوه بر یادآوری، مهارت‌های تفکر و تحلیل را نیز ارزیابی کنند.

افزایش توجه به مهارت‌های شناختی سطح بالاتر در سال‌های اخیر، همان‌طور که در تحلیل‌ها مشاهده شد، نشان‌دهنده تغییر در طراحی آزمون‌ها است. اگرچه نتایج آزمون کای اسکوتر تغییرات معناداری را نشان داد، این تغییرات همچنان نسبت به سطوح ابتدایی همچون دانش و درک محدود باقی مانده است. این تغییرات، روندی مشابه با یافته‌های کندی و همکاران (۱۹۹۱) را تأیید می‌کند که بر لزوم وجود تعادل در ارزیابی‌ها و تشویق به تفکر انتقادی تأکید دارند؛ درحالی‌که در این مطالعه همچنان کمبود تأکید بر مهارت‌های تفکر تحلیلی و حل مسئله مشاهده می‌شود.

در این راستا، توجه کم به سطوح ارزیابی و - خلق مهارت‌های ضروری برای تفکر انتقادی و حل مسائل خلاقانه (فاسیونه، ۱۹۹۰) - حاکی از آن است که هنوز هم نیاز به بهبود طراحی آزمون‌ها برای تقویت این مهارت‌ها احساس می‌شود. همان‌طور که ویگوتسکی (۱۹۶۲) اشاره کرده است، زبان و تفکر به شدت با هم مرتبط هستند و آموزش تفکر انتقادی، از طریق ارزیابی‌هایی که بر تحلیل، ارزیابی و خلق تأکید دارند، می‌تواند به‌طور قابل توجهی به پیشرفت تفکر دانش‌آموزان کمک کند.

این تحقیق همچنین بر لزوم اتخاذ رویکردی متوازن در طراحی آزمون‌ها تأکید دارد؛ در حالی که مهارت‌های شناختی سطح پایین مانند حافظه‌سپاری و درک برای پایه‌گذاری دانش ضروری هستند، نباید از تأکید بر مهارت‌های تفکر انتقادی مانند تحلیل و ارزیابی غفلت کرد؛ چرا که این مهارت‌ها برای رشد فکری دانش‌آموزان و آمادگی آن‌ها برای تحصیلات عالی و بازار کار ضروری هستند (هک ورت، ۲۰۱۰). مطابق با تحقیقات چنگ و کرتس (۲۰۰۴)، این مطالعه نشان می‌دهد که اگر در طراحی آزمون‌ها به مهارت‌های شناختی سطح بالا توجه نشود، دانش‌آموزان ممکن است برای انجام کارهایی که نیاز به تفکر و تحلیل دارند، به اندازه کافی آماده نباشند.

در نهایت، نتایج این مطالعه تأکید بر لزوم اصلاح و بازنگری در برنامه‌ی درسی و آزمون‌های نهایی دارند. تأکید باید بر توسعه‌ی تفکر انتقادی و مهارت‌های شناختی سطح بالا قرار گیرد، به‌ویژه در زمینه‌های ارزیابی و خلق؛ تا دانش‌آموزان برای یادگیری مادام‌العمر و حل مسائل در دنیای پیچیده‌ی امروزی آماده شوند. این نتایج می‌توانند به معلمان، سیاست‌گذاران و طراحان آزمون‌ها در ایجاد ارزیابی‌هایی که نه تنها دانش پایه را می‌سنجند؛ بلکه تفکر انتقادی را نیز ترویج می‌دهند، کمک کنند.

۷. نتیجه‌گیری و پیشنهادهای آموزشی و پژوهشی

تحلیل داده‌های این پژوهش نشان داد که آزمون‌های نهایی در طول سال‌های ۱۳۹۳ تا ۱۴۰۳ به‌طور غالب بر مهارت‌های شناختی سطح پایین‌تر؛ مانند دانش و درک تأکید دارند؛ در حالی که مهارت‌های شناختی سطح بالاتر؛ همچون تحلیل، ارزیابی و خلق، در مقایسه با سطوح اولیه، به میزان کمتری در سؤالات گنجانده شده‌اند. این یافته‌ها به وضوح نشان می‌دهند که در طراحی سؤالات آزمون‌های نهایی فارسی، بیشتر بر ارزیابی حافظه‌سپاری و درک مفاهیم تأکید شده است؛ تا بر مهارت‌های تفکر تحلیلی، ارزیابی انتقادی و حل مسائل پیچیده. هرچند دو سطح دانش و درک، سطوح پایه و اساسی در حوزه‌ی شناخت محسوب می‌شوند و لازم است که دانش‌آموزان در این دو سطح، به حد کافی آموزش دیده و فعالیت نمایند؛ اما امکان ظهور و تجلی خلاقیت‌ها و ابتکارات دانش‌آموزان از طریق پرداختن به سطوح بالاتر حوزه‌ی شناخت میسر است. بنابراین، بخش‌های مختلف نظام تعلیم و تربیت باید در تولید و سازمان‌دهی محتوا، اجرای تدریس و انجام ارزشیابی به تمام سطوح حوزه‌ی شناختی، نگاه متوازن داشته باشند. این مطالعه همچنین تأکید می‌کند که توجه بیشتر به تفکر انتقادی و حل مسئله در طراحی آزمون‌ها، می‌تواند به ارتقای مهارت‌های شناختی دانش‌آموزان کمک کند و آن‌ها را برای مواجهه با چالش‌های پیچیده‌تر آماده سازد. همان‌طور که در نتایج این پژوهش مشاهده شد؛ سؤالات آزمون‌ها به‌ویژه در سطوح بالا همچون تحلیل و ارزیابی، توجه کمی داشته‌اند، در حالی که این سطوح به‌عنوان مهارت‌های ضروری برای تفکر انتقادی و حل مسائل پیچیده شناخته می‌شوند. این مسئله به‌ویژه در راستای آموزش زبان و ادبیات فارسی که تأثیر زیادی در شکل‌دهی تفکر انتقادی و استدلال منطقی دانش‌آموزان دارد، نگران‌کننده است.

توزیع فعلی سؤالات، توانایی پاسخ به تمامی اهداف آموزشی طبقه‌بندی بلوم (به‌عنوان یکی از معتبرترین مرجع‌های سنجش) را ندارد. سطوح دانش و درک، به‌تنهایی ۹۰٪ سؤالات را به خود اختصاص داده‌اند و این حاکی از آن است که طراحان سؤالات امتحانی به دلایل مختلف (که یکی از آن‌ها سنجش آسان‌تر دانش‌آموزان و عدم دخالت نظر شخصی نمره‌دهندگان در نمره‌دهی است) سؤالات را در حد سطوح شناختی پایین طراحی

می‌کنند و توجه اندکی به سطوح بالا؛ از جمله ترکیب و ارزیابی دارند. از آنجا که هرکدام از این سطوح، در پرورش تفکر انتقادی نقش بسزایی دارند، نیاز است به میزان مساوی سؤالات را پوشش دهند.

همچنین، از آنجا که مهارت‌های تفکر سطح بالا از اهمیت ویژه‌ای برخوردارند، این مطالعه بر لزوم بازنگری در طراحی سؤالات آزمون‌های نهایی فارسی تأکید دارد. باید توجه بیشتری به سطوح شناختی بالاتر؛ مانند تحلیل، ارزیابی و خلق در طراحی سؤالات معطوف شود. این تغییرات، نه تنها باعث تقویت تفکر انتقادی در دانش‌آموزان می‌شود؛ بلکه آن‌ها را قادر می‌سازد تا مهارت‌های تحلیلی و استدلالی خود را به‌طور مؤثرتری به‌کار بگیرند.

یافته‌های این تحقیق پیامدهای عملی قابل توجهی برای طراحان سؤالات آزمون‌های نهایی دارند. آن‌ها می‌توانند از این نتایج برای بازنگری در روش‌های طراحی سؤالات استفاده کنند و تلاش کنند تا در کنار ارزیابی حافظه‌سپاری و درک مفاهیم، تفکر تحلیلی و حل مسئله نیز در سؤالات گنجانده شود؛ به‌ویژه در درسی مانند فارسی که هدف آن علاوه بر تقویت زبان‌آموزی، پرورش تفکر انتقادی و خلاقیت است، استفاده از سؤالاتی که به تحلیل و ارزیابی محتوای متنی می‌پردازند، می‌تواند به بهبود مهارت‌های شناختی دانش‌آموزان کمک کند. طراحی سؤالاتی که از دانش‌آموزان بخواهند؛ تا اطلاعات را تجزیه و تحلیل کرده و آن‌ها را در زمینه‌های مختلف مقایسه کنند، می‌تواند تأثیر زیادی در تقویت تفکر انتقادی داشته باشد.

علاوه بر این، این مطالعه شکاف‌های تحقیقاتی قابل توجهی را در زمینه طراحی سؤالات آزمون‌های نهایی درس فارسی نشان می‌دهد. بیشتر پژوهش‌های موجود به آزمون‌های سطح پایین‌تر پرداخته‌اند و توجه کمتری به سطوح بالاتر طبقه‌بندی بلوم دارند. به همین دلیل، پژوهش‌های آینده باید به‌طور خاص به تحلیل سؤالات آزمون‌های نهایی در دروس عمومی مانند فارسی بپردازند و عواملی را که باعث کاهش تمرکز بر مهارت‌های تحلیلی و ارزیابی می‌شوند، شناسایی کنند. در این راستا، تحقیقاتی که به بررسی چگونگی تأثیر طراحی سؤالات با سطوح شناختی بالا بر تفکر دانش‌آموزان و نتایج تحصیلی آن‌ها می‌پردازند، می‌توانند به غنای علمی این حوزه کمک کنند.

همچنین، تحقیقات آینده می‌توانند به بررسی ارتباط میان هدف‌گذاری آموزشی و طراحی آزمون‌ها بپردازند. به‌عبارت دیگر، باید بررسی شود که چگونه هدف‌های آموزشی در سیستم‌های آموزشی ایران با ارزیابی‌های استاندارد در هم‌خوانی هستند و چگونه این هم‌خوانی می‌تواند باعث افزایش کارآمدی آموزش و ارتقای مهارت‌های شناختی دانش‌آموزان شود. علاوه بر این، پژوهش‌های آینده می‌توانند به تحلیل و ارزیابی روش‌های نوین طراحی آزمون‌های نهایی بپردازند که نه تنها دانش پایه را ارزیابی کنند؛ بلکه تفکر انتقادی، حل مسئله و خلاقیت را نیز به چالش بکشند.

در نهایت، این پژوهش تأکید می‌کند که تغییرات در طراحی سؤالات آزمون‌های نهایی درس فارسی باید به سمت گنجاندن مهارت‌های شناختی سطح بالا سوق داده شود؛ تا به دانش‌آموزان کمک کند، تفکر انتقادی و

مهارت‌های تحلیلی خود را توسعه دهند. به طور کلی، این مطالعه با هدف بهبود طراحی و تحلیل سؤالات آزمون‌ها و افزایش توانمندی‌های شناختی دانش‌آموزان در سطح ملی، پیشنهاد می‌کند که طراحان سؤالات و سیاست‌گذاران آموزشی به طور جدی‌تری به تفکر تحلیلی و تفکر انتقادی توجه کنند و در این مسیر، ابزارهای ارزیابی کارآمدتر و هدفمندتری ایجاد کنند.

فهرست منابع:

- سپاسی، حسین. (۱۳۸۵). بررسی و تحلیل سطوح حیطة شناختی و شاخص‌های روان‌سنجی امتحان نهایی دروس عربی، حسابان و زیست‌شناسی دانش‌آموزان دختر پایه سوم متوسطه در سه منطقه متفاوت اقتصادی اجتماعی استان خوزستان، *مجله علوم تربیتی و روان‌شناسی*، دوره ۳، ش ۴، صص ۷۸-۵۷. <https://doi.org/10.22055/edus.2007.15981>
- سیف، علی‌اکبر. (۱۳۹۹). *اندازه‌گیری، سنجش و ارزشیابی آموزشی* (ویرایش هفتم). تهران: دوران.
- علی‌مراد، زهرا و رزمی، یگانه، رویا. (۱۴۰۳). ارزیابی محتوایی مجموعه «فارسی بیاموزیم» براساس چارچوب طبقه‌بندی تجدیدنظرشده بلوم، *پژوهش‌نامه آموزش زبان فارسی به غیرفارسی‌زبانان*، دوره ۱۳، ش ۱، صص ۱۳۰-۱۰۱. <https://doi.org/10.30479/jtpsol.2024.20294.1666>
- فلسفی‌نژاد، محمدرضا، فرخی، نورعلی و بهرامی، لیلیا. (۱۳۹۵). ویژگی‌های روان‌سنجی امتحانات نهایی سال سوم متوسطه و قابلیت آن‌ها در گزینش داوطلبان ورود به دوره‌های کارشناسی، *فصلنامه اندازه‌گیری تربیتی*، دوره ۷، ش ۳، صص ۷۶-۴۵. <https://doi.org/10.22054/jem.10.22054>
- میرآقایی، علی عباس، سپاسی، حسین، مهاجران، بهناز و قلعه‌ای، علیرضا. (۱۳۹۴). بررسی و تحلیل شاخص‌های روان‌سنجی و سطوح حیطة شناختی سؤالات امتحانات نهایی دروس ریاضیات و علوم پایه سوم راهنمایی شهرستان خرم‌آباد، *مجله روان‌شناسی مدرسه*، دوره ۴، ش ۳، صص ۱۱۸-۱۰۲. <https://doi.org/10.22098/jsp.10.22098>
- نعمتی سرخی، سعیدی، زری و صحرائی، رضامراد. (۱۴۰۱). تحلیل محتوای نرم‌افزارگردشگری فرهنگی بر مبنای نظریه یادگیری تجدیدنظرشده بلوم، *پژوهش‌نامه آموزش زبان فارسی به غیرفارسی‌زبانان*، دوره ۱۱، ش ۱، صص ۳۰۲-۲۷۵. <https://doi.org/10.30479/jtpsol.10.30479>

References:

- Alimorad, Z., & Razmi, Y. R. (2024). Content evaluation of the “Learn Persian” series based on the revised Bloom’s taxonomy framework. *Persian Language Teaching Research Journal for Non-Persian Speakers*, 13(1), 101–130. [In Persian]
- Alwali, A. K. (2011). Benefits of using critical thinking in high education. In L. Gómez Chova, I. Candel Torres, & A. López Martínez (Eds.), *Proceedings of the 5th International Technology, Education and Development Conference (INTED2011)* (pp. 2527–2532). IATED. Available at: <https://library.iated.org/view/ALWALI2011BEN>. [In Persian]
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. McGraw-Hill.

- Cheng, L., & Curtis, A.** (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in Language Testing: Research Contexts and Methods* (pp. 3–40). Routledge. <https://doi.org/10.4324/9781410609731-9>
- Eli, A. R.** (2011). Critical thinking: A literature review. Defense Acquisition University. <https://www.dau.edu/sites/default/files/2023-12/CriticalThinkingReview.pdf>
- Facione, P. A.** (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. The California Academic Press. Retrieved from <https://www.qcc.cuny.edu/socialSciences/ppecorino/CT-Expert-Report.pdf>
- Falsafi-Nejad, M. R., Farrokhi, N., & Bahrami, L.** (2016). Psychometric characteristics of final exams of third-grade high school and their applicability in selecting candidates for undergraduate programs. *Educational Measurement Quarterly*, 6(23), 45–76. [In Persian]
- Forehand, M.** (2010). Bloom's taxonomy: Original and revised. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology*. Retrieved from <https://cdn.vanderbilt.edu/vu-wp0/wp-content/uploads/sites/59/2010/06/12092513/BloomsTaxonomy-mary-forehand.pdf>
- Glass, A. L., Ingate, M., & Sinha, N.** (2013). The effect of a final exam on long-term retention. *Journal of General Psychology*, 140(3), 224–241. <https://doi.org/10.1080/00221309.2013.797379>
- Hackworth, R. M.** (2010). Radiation science educators' perception of obstacles in the use of critical thinking (Master's thesis, Ohio State University). OhioLINK Electronic Theses and Dissertations Center. Retrieved from http://rave.ohiolink.edu/etdc/view?acc_num=osu1262120623
- Halpern, D. F.** (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53(4), 449–455. <https://doi.org/10.1037/0003-066X.53.4.449>
- Huitt, W.** (2011). Bloom et al.'s taxonomy of the cognitive domain. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved August 31, 2025, from <http://www.edpsycinteractive.org/topics/cognition/bloom.html>
- Kennedy, M., Fisher, M. B., & Ennis, R. H.** (1991). Critical Thinking: Literature Review and Needed Research. In L. Idol & B. F. Jones (Eds.), *Educational Values and Cognitive Instruction* (pp. 11–40). Routledge. <https://doi.org/10.4324/9781315044392-2>
- Massa, S.** (2014). The development of critical thinking in primary school: The role of teachers' beliefs. *Procedia - Social and Behavioral Sciences*, 141, 387–392. <https://doi.org/10.1016/j.sbspro.2014.05.068>
- MirAghaei, A. A., Sepasi, H., Mohajeran, B., & Qalaei, A.** (2015). Analysis of psychometric indices and cognitive domain levels of final exam questions in mathematics and science of third-grade middle school students in Khorramabad. *School Psychology Journal*, 4(3), 102–118. [In Persian]
- Nemati-Sorkhi, M., Saedi, Z., & Sahraei, R. M.** (2022). Content analysis of cultural tourism software based on the revised Bloom's learning theory. *Persian Language Teaching Research Journal for Non-Persian Speakers*, 11(1), 275–302. [In Persian]

- Pešić, J.** (2011). Sličnosti i razlike u konceptualizovanju kritičkog mišljenja [Similarities and differences in conceptualizing critical thinking]. *Psihološka istraživanja*, 14(1), 5–23. <https://scindeks-clanci.ceon.rs/data/pdf/0352-7379/2011/0352-73791101005P.pdf>
- Riddell, T.** (2007). Critical assumptions: Thinking critically about critical thinking. *Journal of Nursing Education*, 46(3), 121–126. <https://doi.org/10.3928/01484834-20070301-06>.
- Seif, A. A.** (2020). *Measurement, assessment, and educational evaluation* (7th ed.). Tehran: Doran Publishing. [In Persian]
- Sepasi, H.** (2006). Analysis of cognitive domain levels and psychometric indices of final exams in Arabic, calculus, and biology of third-grade female high school students in three socio-economic regions of Khuzestan Province. *Journal of Educational Sciences and Psychology*, 3(4), 57–78. [In Persian]
- Shohamy, E.** (2013). The power of tests: A critical perspective on the uses of language tests (Reprinted). Routledge. <https://doi.org/10.4324/9781315837970>
- Sievertsen, H. H.** (2023). Assessments in education. In Oxford Research Encyclopedia of Economics and Finance. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190625979.013.846>
- Vygotsky, L. S.** (1962). *Thought and language*. Cambridge, MA: MIT Press.
- Zaidi, N. L. B., Grob, K. L., Monrad, S. M., Kurtz, J. B., Tai, A., Ahmed, A. Z., Gruppen, L. D., & Santen, S. A.** (2018). Pushing critical thinking skills with multiple-choice questions: Does Bloom's taxonomy work? *Academic Medicine*, 93(6), 856–859. <https://doi.org/10.1097/ACM.0000000000002087> . [In Persian]