



Review and analysis of final exam questions for the 12th grade literary sciences  
and arts course based on Bloom's taxonomy

GH. Mehravar<sup>\*,1</sup>, A. Heidari<sup>2</sup>, E. Samani<sup>3</sup>, A. Hajjami

1. Department of Persian Language and Literature Education, Farhangian University, Tehran, Iran  
(Corresponding Author).

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

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

4. Master's student in Persian Language and Literature, Department of Literary Criticism and Theory,  
University of Guilan.

ABSTRACT

**Keywords:**

Bloom's taxonomy,  
Cognitive levels,  
Critical thinking,  
Final examinations,  
Fisher's exact test,  
Question analysis.

**Background and Objectives:** The present study aimed to analyze the final examination questions of the Persian Literature and Rhetoric course (Grade 12) based on Bloom's cognitive taxonomy during the period 2018–2025 (1397–1404 in the Iranian calendar). **Methods:** The study employed a descriptive-analytical design using a mixed-method approach combining both quantitative and qualitative analyses. The dataset consisted of all national final exam questions administered by the Iranian Ministry of Education during the specified years. In the first phase, questions were classified into Bloom's six cognitive levels—Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. To ensure validity and reliability, dual independent coders reviewed the classifications, and inter-rater agreement exceeded 0.85. Descriptive statistics were used to determine the frequency and percentage of each cognitive level, and Fisher's Exact Test was applied to examine the association between exam year and cognitive level distribution. **Findings:** The results revealed that over 80% of the questions belonged to the three lower levels (Knowledge, Comprehension, and Application), while no items were found at the Synthesis or Evaluation levels. Although Fisher's Exact Test indicated a statistically significant relationship between year and cognitive level, the pattern of change lacked a consistent upward trend toward higher-order thinking. **Conclusion:** Overall, the findings demonstrate the dominance of rote-learning tendencies in national assessments and a relative neglect of higher-order cognitive skills, such as critical thinking and creativity. The study's implications highlight the need for systematic revision of exam design and curriculum planning to promote advanced cognitive engagement in students.


1. Corresponding author  
✉ gh.mehravar@cfu.ac.ir

Received: 1404/08/07

Reviewed: 1404/10/12

Accepted: 1404/10/15

PP: 17

Citation (APA): Mehravar, GH & et al. (2026). Review and analysis of final exam questions for the 12th grade literary sciences and arts course based on Bloom's taxonomy. *The Journal of Theory and Practice in Teachers Education*, 7 (1), 35-51.  <https://doi.org/10.48310/rpllp.2026.21329.1286>



## بررسی و تحلیل سؤالات آزمون‌های نهایی درس علوم و فنون ادبی پایه دوازدهم بر اساس طبقه‌بندی شناختی بلوم

### مقاله پژوهشی

قاسم مهرآور\*<sup>۱</sup>، عبدالحسین حیدری<sup>۲</sup>، ابراهیم سامانی<sup>۳</sup>، امیرعباس حجاجی<sup>۴</sup>

۱. استادیار گروه آموزش زبان و ادبیات فارسی، دانشگاه فرهنگیان، تهران، ایران. (نویسنده مسئول)

۲. گروه آموزش زبان انگلیسی، دانشگاه فرهنگیان، تهران، ایران.

۳. گروه آموزش زبان انگلیسی، دانشگاه فرهنگیان، تهران، ایران.

۴. دانشجوی کارشناسی ارشد زبان و ادبیات فارسی گرایش نقد و نظریه ادبی دانشگاه گیلان.

### چکیده

**پیشینه و اهداف:** هدف پژوهش حاضر، تحلیل سؤالات آزمون‌های نهایی درس علوم و فنون ادبی پایه دوازدهم بر اساس سطوح طبقه‌بندی شناختی بلوم در بازه زمانی سال‌های ۱۳۹۷ تا ۱۴۰۴ است. **روش‌ها:** پژوهش از نوع توصیفی - تحلیلی و با رویکرد آمیخته (کمی و کیفی) انجام شده است. داده‌های پژوهش شامل کلیه سؤالات آزمون‌های نهایی کشوری این درس در سال‌های یادشده است. در مرحله نخست، سؤالات بر اساس شش سطح شناختی بلوم (دانش، درک، کاربرد، تحلیل، ترکیب و ارزیابی) طبقه‌بندی شدند و برای اطمینان از روایی و پایایی، فرایند کدگذاری به صورت دو مرحله‌ای و با نظر متخصصان انجام شد. سپس فراوانی و درصد هر سطح محاسبه و توزیع سؤالات در سال‌های مختلف با استفاده از آمار توصیفی و آزمون دقیق فیشر بررسی شد. **یافته‌ها:** نتایج پژوهش نشان داد که بیش از ۸۰ درصد از سؤالات در سه سطح پایین‌تر طبقه‌بندی بلوم (دانش، درک و کاربرد) قرار دارند و هیچ سؤالی در سطوح ترکیب و ارزیابی طراحی نشده است. آزمون دقیق فیشر وجود رابطه آماری معنادار میان سال برگزاری آزمون و سطح شناختی سؤالات را تأیید کرد؛ هرچند تغییرات فاقد الگوی منظم و تصاعدی به سمت سطوح بالاتر شناختی بودند. **نتیجه‌گیری:** این یافته‌ها بیانگر غلبه رویکرد حافظه‌محور در طراحی سؤالات و کم‌توجهی به مهارت‌های تفکر انتقادی و خلاقیت در نظام ارزشیابی رسمی است. نتایج این پژوهش می‌تواند در بازنگری برنامه‌های درسی و طراحی آزمون‌های آینده با هدف تقویت سطوح بالای شناختی مورد استفاده قرار گیرد.

<https://doi.org/10.48310/rpllp.2026.21329.1286>

### واژه‌های کلیدی:

آزمون دقیق فیشر،  
آزمون‌های نهایی،  
تحلیل سؤالات،  
تفکر انتقادی،  
سطوح شناختی،  
طبقه‌بندی بلوم.

۱. نویسنده مسئول

gh.mehravar@cfu.ac.ir

تاریخ دریافت: ۱۴۰۴/۰۸/۰۷

تاریخ بازنگری: ۱۴۰۴/۱۰/۱۲

تاریخ پذیرش: ۱۴۰۴/۱۰/۱۵

شماره صفحات: ۱۷

## مقدمه

یکی از رایج‌ترین ابزارهای ارزیابی در آموزش، برگزاری آزمون‌ها است. آزمون‌ها نه تنها به‌عنوان یک ابزار تلقی می‌شوند، بلکه فراتر از آن، به‌عنوان روش‌های استاندارد در نظر گرفته می‌شوند که برای اندازه‌گیری سیستماتیک نمونه‌ای از رفتار افراد از طریق ارائه مجموعه‌ای از سؤالات به کار می‌روند. آزمون‌ها به منظور سنجش کیفیت، توانایی، مهارت یا دانش یک نمونه در برابر استاندارد مشخص طراحی می‌شوند، که معمولاً می‌تواند قابل قبول یا غیرقابل قبول باشد. در عمل آموزشی، آزمون‌ها روش‌هایی هستند که برای تعیین توانایی دانش‌آموزان در انجام وظایف خاص یا نشان دادن تسلط بر یک مهارت یا دانش استفاده می‌شوند. آزمون‌ها برای اندازه‌گیری سطح یادگیری دانش‌آموزان نیز به کار می‌روند (آدم<sup>۱</sup> و همکاران، ۲۰۲۰). در این میان امتحانات و آزمون‌های نهایی اهمیت ویژه‌ای دارند. پژوهش‌های (گلس<sup>۲</sup> و همکاران، ۲۰۱۳) نشان داد که آزمون‌های نهایی علاوه بر سنجش عملکرد گذشته، به دلیل ماهیت جامع و فراگیر خود، می‌توانند فرایند یادگیری را نیز تقویت کرده و موجب تثبیت بهتر مطالب در حافظه بلندمدت شوند. همچنین باید توجه داشت که آزمون باید طوری طراحی شود که پیشرفت را برای دانش‌آموز به ارمغان آورد. یافته‌های جنکینز و شولتز<sup>۳</sup> (۲۰۱۳) نمایانگر این نکته بود که دانشجویانی که در امتحانات پایانی شرکت می‌کنند، تغییر قابل توجهی در یادگیری و پیشرفتشان ایجاد نمی‌شود؛ همچنین یافته‌های پژوهش احمدی صفا و جعفری (۲۰۱۳) بیانگر آن است که امتحانات نهایی موجب بازخورد منفی<sup>۴</sup> بر روش تدریس معلمان شده و پدیده تدریس برای آزمون<sup>۵</sup> را تقویت می‌کند. طبق نظر شوهامی<sup>۶</sup> (۲۰۱۳/۱/۲۰۱۳)، آزمون‌های سنتی<sup>۷</sup> به کاربرد آزمون و اثرات منفی آن توجهی ندارند و در مقابل در آزمون‌دهی کاربردمحور<sup>۸</sup> این پیامدها (چه مثبت و چه منفی) در نظر گرفته می‌شوند. یکی از ابزارهای کاربردی برای بررسی سؤالات آزمون‌ها، طبقه‌بندی بلوم<sup>۹</sup> است. طبقه‌بندی بلوم یک مدل چندلایه از طبقه‌بندی بر اساس شش سطح شناختی از ساده به پیچیده است. سه سطح پایین یا ساده این طبقه‌بندی به ترتیب عبارتند از: «دانش<sup>۱۰</sup>»، «درک<sup>۱۱</sup>» و «کاربرد<sup>۱۲</sup>». این سطوح در آزمون‌ها معمولاً بیشتر به کار می‌روند. سه سطح بالاتر عبارتند از: «تحلیل<sup>۱۳</sup>»، «ترکیب<sup>۱۴</sup>» و «ارزیابی<sup>۱۵</sup>». این طبقه‌بندی سلسله مراتبی است؛ یعنی برای مثال اگر دانش آموزی در سطح «کاربرد» عمل کند، سطوح «دانش» و «درک» را نیز فراگرفته است (فورهند<sup>۱۶</sup>، ۲۰۱۰). طبقه‌بندی بلوم 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. سطوح

1. Adom
2. Glass
3. Jenkins & Schultz
4. Negative washback
5. Teaching to the test
6. Shohamy
7. Traditional testing
8. Use-oriented' testing
9. Bloom's Taxonomy
10. Knowledge
11. Comprehension
12. Application
13. Analysis
14. Synthesis
15. Evaluation
16. Forehand

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 می‌کند. اندرسون و کراسول<sup>۱</sup> (۲۰۰۱) یکی از اهداف مهم آموزش زبان و ادبیات را رشد تفکر و تقویت روحیه زیبایی‌شناختی دانش‌آموزان تلقی می‌کنند. ویگوتسکی<sup>۲</sup> (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، 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 زنجیر وار 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 توأم 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 علوم و فنون ادبی پایه دوازدهم از سال ۱۳۹۷ تا ۱۴۰۴ 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 علوم و فنون ادبی پایه دوازدهم بر اساس طبقه‌بندی بلوم چگونه است؟
۲. آیا سؤالات امتحان نهایی پایه دوازدهم درس علوم و فنون ادبی تمامی سطوح شناختی بلوم را پوشش می‌دهند؟
۳. آیا بین سال برگزاری امتحان و توزیع سطوح شناختی سؤالات امتحان نهایی (بر اساس طبقه‌بندی بلوم) رابطه‌ی معناداری وجود دارد؟
۴. نسبت سطوح شناختی سؤالات امتحان نهایی بر اساس طبقه‌بندی بلوم در سال‌های ۱۳۹۷ تا ۱۴۰۴ چه تغییری داشته است؟

### پیشینه پژوهش

ارزیابی کیفیت سؤالات آزمون‌های نهایی، به‌ویژه در دروسی مانند علوم و فنون ادبی، از اهمیت بالایی در تعیین سطح یادگیری دانش‌آموزان و کارایی فرآیند آموزش برخوردار است. مطالعات متعدد نشان داده‌اند که نحوه طراحی سؤالات، نه تنها میزان یادگیری دانش‌آموزان را بازتاب می‌دهد؛ بلکه جهت‌گیری تدریس معلمان را نیز تحت تأثیر قرار می‌دهد

1. Anderson & Krathwohl  
2. Vygotsky

(گولیکرز<sup>۱</sup> و همکاران، ۲۰۰۸). 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 علوم و فنون ادبی از منظر این طبقه‌بندی انجام شده است. در پژوهش‌های اخیر، ارزیابی کیفی آزمون‌ها با استفاده از طبقه‌بندی‌های شناختی مانند بلوم به عنوان ابزاری برای سنجش عمق یادگیری به کار رفته است (اندرسون و کارسول، ۲۰۰۱). پیشینه مطالعات نشان می‌دهد که پژوهش‌های متعددی به بررسی کیفیت سؤالات امتحانی در حوزه‌های مختلف پرداخته‌اند؛ اما تمرکز بر روی تطابق سؤالات امتحانات نهایی درس علوم و فنون ادبی با سطوح مختلف طبقه‌بندی بلوم و تحلیل نقاط قوت و ضعف آنها از نظر پوشش دامنه شناختی، تا کنون در مورد امتحانات نهایی درس علوم و فنون ادبی مدارس کشور به‌طور جامع مورد بررسی قرار نگرفته است. تحلیل سؤالات امتحانی بر اساس چارچوب‌های شناختی، یکی از معتبرترین روش‌های سنجش کیفیت آموزشی است (هالادینا<sup>۲</sup>، ۲۰۰۴)؛ بنابراین، این پژوهش با هدف پرکردن این خلأ و ارائه راهکارهایی برای بهبود کیفیت سؤالات امتحانات نهایی درس علوم و فنون ادبی، طراحی و اجرا شده است. در ادامه به برخی از پژوهش‌هایی که در مورد آزمون‌های نهایی بعضی دروس انجام شده اشاره می‌شود:

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 مذکور به نسبت رعایت شده است. سپاسی (۱۳۸۵) در پژوهش خود با عنوان «بررسی و تحلیل سطوح حیطه 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 حیطه 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. از این رو، پژوهش حاضر خلأ مهمی را پر می‌کند و نخستین تلاش نظام‌مند برای تحلیل روند تغییرات شناختی سؤالات امتحانات علوم و فنون ادبی در بازه ۱۳۹۷ تا ۱۴۰۴ است.

همچنین در سطح بین‌المللی نیز، مطالعات متعددی، کیفیت سؤالات امتحانی را با استفاده از طبقه‌بندی بلوم بررسی کرده‌اند. به عنوان نمونه، عمر و ایتان<sup>۳</sup> (۲۰۱۵) در مالزی نشان دادند که بیش از ۷۰٪ سؤالات آزمون‌های نهایی در سطوح «دانش» و «درک» متمرکز هستند و کمتر از ۱۰٪ سؤالات در سطوح «تحلیل» و بالاتر طراحی شده‌اند. در بررسی دیگری که اسکالی<sup>۴</sup> (۲۰۱۷) در ایرلند انجام داد، گزارش کرد که سؤالات دانشگاهی نیز گرایش شدیدی به سطوح پایین شناختی دارند و این امر موجب کاهش توانایی حل مسئله در دانشجویان می‌شود. در مطالعه دیگری که

- 
1. Gulikers
  2. Haladyna
  3. Omar & Atan
  4. Scully

در کشور هند با هدف بررسی سؤالات در چارچوب سطوح شناختی بلوم انجام شد، کاور و سینگ<sup>۱</sup> (۲۰۱۹) نشان دادند که نبود توازن میان سطوح بلوم، مانع از تقویت مهارت‌های فکری سطح بالا در دانش‌آموزان می‌شود. در نتیجه استفاده آگاهانه از سطوح بالای بلوم می‌تواند موجب ارتقای توانایی تحلیل و استدلال در دانش‌آموزان شود (ایراسیان و میراندا<sup>۲</sup>، ۲۰۰۲).

درحالی‌که پژوهش‌های متعددی در حوزه‌های علوم، ریاضی و زبان انگلیسی انجام شده، درس علوم و فنون ادبی به‌ویژه در پایه دوازدهم، کمتر مورد توجه قرار گرفته است. نبود مطالعات طولی در این حوزه موجب شده که تغییرات احتمالی در تمرکز شناختی سؤالات در طول سال‌های مختلف ناشناخته باقی بماند. از این‌رو، پژوهش حاضر با استفاده از آزمون دقیق فیشر، برای نخستین‌بار رابطه آماری بین سال برگزاری امتحان و توزیع سطوح شناختی بلوم را در بازه زمانی ۱۳۹۷ تا ۱۴۰۴ تحلیل کرده است.

### چارچوب نظری پژوهش

پیش‌تر از کاربرد آزمون و اهمیت آزمون‌های نهایی بحث شد، اما یکی از مهم‌ترین مهارت‌هایی که آزمون باید در دانش‌آموزان تقویت کند، مهارت تفکر انتقادی است. درباره تعریف تفکر انتقادی اتفاق نظری وجود ندارد. تفکر انتقادی را به صورت: «ذهن باز، بی‌طرفی، گرایش به جستجوی دلیل، پرسشگری، میل به آگاهی بیشتر، انعطاف‌پذیری و احترام به دیدگاه‌های دیگران» تعریف می‌کنند (لای<sup>۳</sup>، ۲۰۱۱). تفکر انتقادی «قضاوتی هدفمند و خودتنظیم است که منجر به تفسیر، تحلیل، ارزیابی و استنتاج می‌شود، و همچنین شامل تبیین ملاحظات شواهدی، مفهومی، روش‌شناختی، معیاری یا زمینه‌ای است که این قضاوت بر آن‌ها و «ارزیابی درست گزاره‌ها» (انیس<sup>۴</sup>، ۱۹۶۲) استوار است (فاسیونه<sup>۵</sup>، ۱۹۹۰). همچنین تفکر انتقادی را «اندیشیدن معقول و تامل بر تصمیم‌گیری درباره این‌که چه چیزی را باید باور کرد یا چه کاری را باید انجام داد» (انیس، ۱۹۹۳) تعریف کرده‌اند.

نکته قابل توجهی که انیس درباره تعاریف تفکر انتقادی به آن اشاره می‌کند، این مورد است که سه سطح بالای طبقه‌بندی اهداف آموزشی بلوم یعنی تحلیل، ترکیب و ارزیابی اغلب به عنوان تعریف تفکر انتقادی ارائه شده است. حتی میان پژوهشگرانی که اعتقادی به تعریف صریح برای تفکر انتقادی نداشته و این مهارت را نیازمند توضیح درباره عناصر اصلی این مفهوم و نه خلاصه سازی آن می‌دانند (ریدل<sup>۶</sup>، ۲۰۰۷) هم می‌توان اثراتی از سطوح شناختی بلوم یافت. این گروه از پژوهشگران تفکر انتقادی را شامل حوزه‌های شناختی و غیرشناختی می‌دانند. حوزه شناختی معمولاً با ارزیابی محصولات فکری مختلف (پشیچ<sup>۷</sup>، ۲۰۱۱) و تفسیر، تحلیل، استنباط و توضیح گره خورده است. حوزه غیرشناختی هم شامل عادت‌های ذهنی و گرایش‌ها و... است (فاسیونه، ۱۹۹۰).

نظریه تفکر انتقادی در آموزش و پرورش بر پایه اندیشه‌های جان دیویی<sup>۸</sup> (با نام تفکر تأملی<sup>۹</sup>) استوار است. او تفکر تأملی را چنین تعریف می‌کند: «تأمل فعال، مداوم و دقیق درباره‌ی هر باور یا شکل ادعایی از دانش در پرتو شواهدی

1. Kaur & Singh
2. Airasian & Miranda
3. Lai
4. Ennis
5. Facione
6. Riddell
7. Pešić
8. John Dewey
9. Reflective Thinking

که از آن حمایت می‌کند و نتایجی که به آن می‌انجامد؛ که شامل تلاشی آگاهانه و ارادی برای استوارساختن باور بر مبنای شواهد و عقلانیت است» (یلدریم و اوزکاهرامان<sup>۱</sup>، ۲۰۱۱).

مسئله مورد بحث پیرامون تفکر انتقادی، این است که آیا مهارت‌های تفکر انتقادی عمومی‌اند یا مختص رشته‌ای خاص؟ در این باره ۳ دیدگاه «عمومی‌گرا»<sup>۲</sup>، «رشته‌محور»<sup>۳</sup> و «ترکیبی»<sup>۴</sup> وجود دارد. برای مثال در دیدگاه عمومی‌گرا (پیترز و سودن<sup>۵</sup>، ۲۰۰۰) می‌گویند: «سخت است که نوع گسترده‌ای از تفکر را تصور کنیم که هیچ کاربرد قابل توجهی خارج از یک رشته خاص نداشته باشد». در مقابل مک‌پک<sup>۶</sup> (۲۰۱۷) بر ضرورت دانش محتوایی و رشته‌محور در فرآیند تفکر انتقادی تأکید دارد. مک‌پک سه دلیل برای رد دیدگاه عمومی‌گرا ارائه می‌دهد:

۱. تمام تفکر در مورد چیزی انجام می‌شود؛
۲. توانایی تفکر انتقادی عمومی ممکن نیست، زیرا دانش یک موضوع برای تفکر انتقادی لازم است؛
۳. تفکر انتقادی به شدت بین رشته‌ها متفاوت است. برای مثال تفکر انتقادی در علوم تجربی ممکن است ممکن است با تفکر انتقادی در ادبیات متفاوت باشد.

دیدگاه ترکیبی بین پژوهشگران مقبولیت بیشتری دارد و ترکیبی از دو دیدگاه قبلی است. فاسیونه به عنوان یکی از طرفداران این دیدگاه با وجود اینکه معتقد است تفکر انتقادی از دانش محتوایی متفاوت است، اما می‌گوید: «یکی از بهترین روش‌ها برای یادگیری تفکر انتقادی، آموزش آن در یک حوزه موضوعی خاص است» (فاسیونه، ۱۹۹۰). همچنین کورفیس<sup>۷</sup> (۱۹۸۸) و لای (۲۰۱۱) معتقد بودند از آنجا که ملاک‌های شواهد در رشته‌های مختلف متفاوت است، مهارت‌های عمومی تفکر انتقادی لازم اما به تنهایی کافی نیستند تا تفکر انتقادی مؤثر در یک رشته خاص شکل بگیرد.

با این تعاریف ارتباط بین سطوح شناختی مدنظر در طبقه‌بندی بلوم و تفکر انتقادی مشخص می‌شود. طبقه‌بندی بلوم تعاریف دقیقی برای هر یک از شش طبقه اصلی در حیطه شناختی ارائه داده است. این طبقات عبارتند از: «دانش»، «درک یا فهم»، «کاربرد»، «تحلیل»، «ترکیب» و «ارزیابی». این طبقات به ترتیب از ساده به پیچیده و از عینی به انتزاعی مرتب شده‌اند. افزون بر این، فرض بر این بود که طبقه‌بندی اصلی یک سلسله‌مراتب جمعی را نشان می‌دهد؛ بدین معنا که تسلط بر هر مقوله ساده‌تر، پیش‌نیاز تسلط بر مقوله پیچیده‌تر بعدی است. یکی از پرکاربردترین استفاده‌ها از طبقه‌بندی اصلی، طبقه‌بندی اهداف برنامه‌ی درسی و سؤالات آزمون‌ها به‌منظور نشان‌دادن گستره (یا نبود گستره) اهداف و سؤالات در طیف مقوله‌ها بوده است. تقریباً همیشه، این تحلیل‌ها نشان داده‌اند که تأکید زیادی بر اهدافی وجود دارد که تنها به شناخت یا یادآوری اطلاعات نیاز دارند؛ اهدافی که در مقوله‌ی «دانش» قرار می‌گیرند، اما اهدافی که شامل درک و استفاده از دانش هستند یعنی آنهایی که در مقوله‌های «فهم» تا «ترکیب» قرار می‌گیرند معمولاً به‌عنوان مهم‌ترین اهداف آموزش در نظر گرفته می‌شوند (کراثول، ۲۰۰۲).

در این میان نقش معلمان در آموزش تفکر انتقادی بسیار اهمیت دارد. برای مثال پاجارس<sup>۸</sup> (۱۹۹۲) با استناد به چندین منبع، ارتباطی قوی میان باورهای آموزشی معلمان و برنامه‌ریزی، تصمیم‌های آموزشی و شیوه‌های تدریس آنان

- 
1. Yildirim and Özkahraman
  2. Generalist Position
  3. Specifist Position
  4. Blended Position
  5. Pithers & Soden
  6. McPeck
  7. Kurfiss
  8. Pajares

را نشان داد یا ماسا<sup>۱</sup> (۲۰۱۴) معتقد بود باورهای معلمان درباره یادگیری و هوش می‌تواند تفاوت در میزان استفاده از فعالیت‌های تفکر انتقادی در کلاس درس را توضیح دهد. لیپمن<sup>۲</sup> (۱۹۸۸) معتقد بود برای کمک به دانشجویان در توسعه مهارت‌های تفکر انتقادی، معلمان باید فرایندهایی را که تفکر انتقادی را شکل می‌دهند درک کرده و فعالیت‌های آموزشی‌ای را به کار گیرند که پرورش‌دهنده این فرایندها باشد. لیپمن توصیه می‌کند معلمان اطلاعات را تعریف و روشن کنند، پرسش‌های مناسب مطرح نمایند، گزاره‌ها یا باورها را شفاف‌سازی یا به چالش بکشند، اعتبار منابع را قضاوت کنند و مسائل را با پیش‌بینی پیامدهای احتمالی به شیوه‌ای منطقی یا استنتاجی حل نمایند.

با توجه به مطالب بالا نقش معلمان در آموزش و نهادینه کردن تفکر انتقادی مشخص شد و نیاز به ذکر نیست که یکی از موقعیت‌هایی که باور معلمان در آن نمایان می‌شود، در طراحی آزمون‌ها است. بررسی آزمون‌های درس علوم و فنون ادبی پایه دوازدهم بر اساس طبقه‌بندی بلوم از چندین جهت حائز اهمیت است. از آنجایی که تا به حال بررسی در زمینه علوم و فنون بر اساس طبقه‌بندی بلوم نشده، پژوهش فوق می‌تواند شروعی برای بررسی عمیق‌تر این درس باشد. همچنین همانطور که نظرات پژوهشگران درباره عمومی یا تخصصی بودن تفکر انتقادی بیان شد، مشخص شد که دیدگاه ترکیبی مقبولیت بیشتری بین پژوهشگران دارد یعنی بهترین راه یادگیری تفکر انتقادی آموزش آن در یک رشته خاص است. این پژوهش با بررسی سؤالات آزمون‌های علوم و فنون ادبی علاوه بر طبقه‌بندی سؤالات در سطوح شناختی بلوم، مشخص می‌کند که سؤالات به شیوه یکسان بین سطوح پخش شده‌اند یا نه یا به بیان دیگر سؤالات صرفاً پایه‌های رسیدن به تفکر انتقادی را تقویت کرده (مرتبط با سطوح پایین‌تر طبقه‌بندی بلوم هستند) یا تفکر انتقادی دانش‌آموزان را به درستی تقویت و سنجش می‌کنند؟ پایه دوازدهم از این جهت انتخاب شد که آخرین سال تحصیل در دوره متوسطه دوم است و پس از آن دانش‌آموزان یا وارد دانشگاه و مؤسسات آموزش عالی یا بازار کار می‌شود و نیاز است تا مهارت تفکر انتقادی را داشته باشند و بتوانند در موقعیت‌های مختلف از آن استفاده کنند.

### روش پژوهش

این پژوهش به بررسی سؤالات امتحانات نهایی درس علوم و فنون ادبی پایه دوازدهم در ایران طی سال‌های ۱۳۹۷ تا ۱۴۰۴ پرداخته است. داده‌های مورد استفاده شامل سؤالات امتحانی منتشرشده توسط وزارت آموزش و پرورش بوده که از نظر ساختار و اهداف آموزشی استاندارد و هماهنگ طراحی شده‌اند. ابزار اصلی گردآوری داده‌ها، بانک سؤالات امتحانی کشوری و آرشيو سؤالات کشوری علوم و فنون ادبی منتشرشده توسط مرکز سنجش وزارت آموزش و پرورش بود. این پژوهش که از نظر هدف، ماهیتی توصیفی - تحلیلی و از نظر نوع داده‌ها، ترکیبی (آمیخته) دارد؛ بدین معنا است که از هر دو رویکرد کمی و کیفی برای بررسی سؤالات آزمون‌های نهایی درس علوم و فنون ادبی پایه دوازدهم استفاده شده است. این ترکیب روش‌ها به پژوهشگر اجازه می‌دهد تا ضمن تحلیل آماری دقیق، به بررسی محتوایی و تفسیر معناشناختی سؤالات نیز بپردازد. استفاده از دو رویکرد مکمل، امکان تبیین عمیق‌تر نتایج و استنتاج‌های معتبرتر را فراهم ساخته است. این سوالات پس از گردآوری، بر اساس ساختار پرسش‌ها، موضوعات درسی و اهداف آموزشی متناظر با محتوای کتاب درسی، سامان‌دهی و کدگذاری شدند. در مرحله‌ی نخست، هر سؤال بر اساس نوع مهارت شناختی موردنیاز برای پاسخ‌دهی، در یکی از شش سطح طبقه‌بندی شناختی بلوم یعنی دانش، درک، کاربرد، تحلیل، ترکیب و ارزیابی قرار گرفت. برای اطمینان از صحت کدگذاری، از دو متخصص در حوزه برنامه‌ریزی درسی و آموزش زبان و ادبیات فارسی درخواست شد تا به صورت مستقل سؤالات را در سطوح بلوم طبقه‌بندی کنند. سپس میزان توافق بین داوران محاسبه شد و در موارد اختلاف، با بحث و اجماع نهایی دسته‌بندی اصلاح گردید. این اقدام به

1. Massa  
2. Lipman

افزایش روایی محتوایی و پایایی طبقه‌بندی انجامید. سوالات در مرحله اول بر اساس شش سطح مختلف شناختی بلوم طبقه‌بندی شدند (بلوم و همکاران، ۱۹۵۶).

دانش: دانش به‌عنوان یادآوری یا بازیابی مطالبی که قبلاً آموخته شده تعریف می‌شود. اهداف یادگیری در این سطح معمولاً شامل تعریف اصطلاحات کلیدی، فهرست کردن مراحل یک فرایند یا تکرار چیزی که شنیده یا دیده شده است می‌باشد.

درک: بزرگ‌ترین دسته از مهارت‌ها و توانایی‌های شناختی است. مهارت کلیدی در این سطح پردازش اطلاعات جدید است. سوالات این حیطه نیازمند فهم اطلاعات هستند.

کاربرد: در این سطح، یادگیرنده باید بتواند یک مسئله جدید را با به‌کارگیری اطلاعات حل کند، بدون اینکه نیاز به یادآوری یا راهنمایی مداوم داشته باشد. اهداف در این سطح می‌توانند شامل تفسیر اطلاعات، نشان دادن تسلط بر یک مفهوم، یا به‌کارگیری یک مهارت آموخته‌شده باشند.

تحلیل: یادگیرنده باید بتواند روابط میان اجزاء را تشخیص دهد. اهداف در این سطح اغلب شامل افعالی چون تمایز قائل‌شدن، مقایسه و مقابله‌کردن، نقدکردن یا آزمایش کردن هستند.

ترکیب: این سطح نیازمند رفتار خلاقانه است، زیرا یادگیرندگان محصولات تازه‌ای می‌سازند که اغلب منحصربه‌فرد نیز هستند. اهداف یادگیری در این سطح ممکن است از یادگیرندگان بخواهد یک طرح ایجاد کنند، پیشنهادی ارائه دهند، محصولی طراحی کنند یا اطلاعات را سازمان‌دهی نمایند.

ارزیابی: سوالات در این بخش نیازمند قضاوت و ارزیابی دانش بر اساس معیارهای معین هستند. قضاوت این بخش مبتنی بر محتوا و نقد نوشتار است.

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

جدول ۱: نمونه‌ای از سوالات آزمون‌های نهایی بر اساس طبقه‌بندی شناختی بلوم (۱۹۵۶)

سطح طبقه‌بندی بلوم	توضیحات	کلمات کلیدی	نمونه سوالات مربوطه در آزمون
دانش	دانش‌آموز پاسخ را به یاد می‌آورد و بدون تغییر و تفسیر جواب را بازیابی می‌کند.	نام‌بردن، مشخص کردن، بیان کردن، تعریف کردن و ...	دی ماه ۱۴۰۰ - سؤال ۳: کتاب «دستور زبان عشق» اثر کیست؟ خرداد ماه ۱۴۰۲ سؤال ۸: مفهوم کدام بیت با مفهوم «آزادی» در دوره بیداری، متفاوت است؟
درک	در این بخش دانش‌آموز باید معنای ابیات را بفهمد و عبارات مشخص شده را تفسیر کرده و توضیح بدهد یا گزینه مناسب را انتخاب کند.	توضیح‌دادن، خلاصه‌کردن، بازگوکردن، توصیف کردن و ...	الف) در محیط طوفان‌زا ماهرانه در جنگ است/ ناخدای استبداد با خدای آزادی ب) اسیری هیچ آزادی نجوید/ چو دل در بند گیسوی تو دارد
کاربرد	علاوه بر بازیابی و فهم، دانش‌آموز باید دانش خاصی را به موقعیت جدیدی منتقل و استفاده کند.	استفاده‌کردن، ساختن، شناسایی‌کردن، شبیه‌سازی‌کردن، اعمال کردن و ...	دی ماه ۹۸ - سؤال ۱۹: در هر یک از بیت‌های گروه «الف» کدام آرایه ادبی گروه «ب» به کار رفته است؟ «قسمت» «ب» یک مورد اضافه است. الف) شب تاریک و بیم موج و گردابی چنین هایل/ کجا دانند حال ما سبکباران ساحل‌ها

ب) برب کسوه جنون خنده شیرین بهار/  
نقش زخمی است که از تیشه فرهاد شکفت  
ج) اگرچه نقش دیوارم به ظاهر از گران  
خوابی/ اگر رنگ از رخ گل می‌پرد، بیدار  
می‌گردم  
۱- اغراق  
۲- حسن تعلیل  
۳- تلمیح  
۴- مراعات نظیر

با توجه به متن زیر از «چرند و پرند  
دهخدا» به سؤالات پاسخ دهید:  
«گفتم: بابا! والله من مُرده، شماها زنده؛ شما  
از وکیل خیر نخواهید دید، مگر همان  
مشروطه خالی چطور است؟  
گفتند: برو پی کارت! سواد نداری، حرف  
نزن. باری، حالا بعد از دو سال سر حرف من  
افتاده‌اند. حالا تازه می‌فهمند که شأن مُقنن  
(= قانون‌گذار) بالاتر از آن است که به قانون  
عمل کند. اینها را مردم تازه می‌فهمند؛ من  
از قدیم می‌فهمیدم، چون گریه‌های مادرم را  
دیده بودم، برای این که چشم‌های مثل آرزق  
شامی بایام هنوز یادم بود».  
الف) با توجه به نثر فوق، بنیادی‌ترین تفکر  
و خواست مشروطه‌خواهان چه بوده است؟  
ب) بارزترین ویژگی زبانی نثر دوره بیداری  
را با توجه به متن بنویسید.

ساده‌سازی کردن، دانش‌آموز باید ساختار منطقی و رابطه علت  
مقایسه کردن، بررسی کردن، و معلول متن را شناسایی کند، نه صرفاً تحلیل  
تفسیر کردن و... معنی یا اطلاعات را یادآوری کند.

ایجاد کردن، دانش‌آموز، دانش به‌دست‌آمده را تلفیق و ترکیب  
ترکیب کردن، پیشنهاد کردن، چیز منحصر به فردی تولید می‌کند  
خلق کردن، توسعه دادن و...  
کشف کردن، اثبات کردن، دانش‌آموز درباره ارزش و اعتبار موضوع‌های  
ارزش‌یابی کردن، ارزش‌یابی کردن، مختلف به ارزیابی و قضاوت می‌پردازد.  
ارزیابی  
قضاوت کردن، نقد کردن و ...



جدول ۱ 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. برای تحلیل آماری، از دو سطح آمار توصیفی و استنباطی استفاده شد. در سطح توصیفی، فراوانی‌ها، درصدها و نمودارهای مقایسه‌ای برای هر سال ترسیم شد. در سطح استنباطی، به‌منظور بررسی رابطه میان سال برگزاری آزمون و سطح شناختی سؤالات، ابتدا آزمون کای اسکوئر<sup>۱</sup> انجام شد. اما با توجه به آن که در بخش عمده‌ای از جدول توافقی<sup>۲</sup> تعداد مورد انتظار هر سلول کمتر از ۵ بود، پیش‌فرض‌های آزمون کای اسکوئر نقض شد. بر این اساس، برای حفظ دقت و اعتبار تحلیل، از آزمون دقیق فیشر استفاده شد که مناسب‌ترین روش برای جداول توافقی با حجم نمونه کوچک و فراوانی‌های پایین محسوب می‌شود. آزمون دقیق فیشر امکان سنجش وجود یا عدم وجود رابطه آماری معنادار میان متغیر سال و سطوح شناختی بلوم را

1. Chi-square
2. Contingency Table

بدون اتکا به تقریب توزیع کای اسکوتر فراهم می‌کند. از آنجا که تمام سلول‌های جدول در این پژوهش دارای فراوانی کمتر از ۵ بودند، استفاده از این آزمون ضروری بود. تحلیل داده‌ها با نرم‌افزار SPSS نسخه ۲۶ انجام گرفت و سطح معناداری آماری در تمامی آزمون‌ها ۰/۰۵ ( $p < 0.05$ ) در نظر گرفته شد. در کنار آزمون فیشر، شاخص‌های تکمیلی مانند درصد پوشش سطوح شناختی، روند تغییرات فراوانی سالانه، و نمودارهای ستونی تطبیقی برای تبیین الگوهای تغییر طراحی شدند.

برای اطمینان از روایی محتوایی، از روش قضاوت خبرگان بهره گرفته شد. سه نفر از اساتید حوزه آموزش زبان فارسی و ارزشیابی تحصیلی، طبقه‌بندی سؤالات را بازبینی کردند و مطابقت آن را با سطوح طبقه‌بندی بلوم تأیید نمودند. همچنین برای بررسی پایایی بین‌کدگذار، ضریب توافق کاپا میان دو کدگذار مستقل محاسبه شد که عدد به‌دست‌آمده بیش از ۰.۸۵ بود و این مقدار بیانگر توافق بالا و پایایی قابل قبول طبقه‌بندی است.

### یافته‌ها

در تحلیل نتایج، تمرکز اصلی بر دو پرسش کلیدی بود: نخست، این‌که آیا سطوح مختلف شناختی در آزمون‌های سال‌های مورد بررسی به شکل متوازن به‌کار رفته‌اند یا خیر؛ و دوم، آیا رابطه‌ای معنادار میان سال برگزاری آزمون و سطح شناختی سؤالات وجود دارد. 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.

جدول ۲: توزیع فراوانی سطوح طبقه‌بندی بلوم بر اساس سال

تعداد	مرحله سال	دانش	درک	کاربرد	تحلیل	ترکیب	ارزیابی	سؤالات هر سال
۲۸	۱۳۹۷	۹	۹	۵	۵	۰	۰	۲۸
۲۵	۱۳۹۸	۹	۴	۹	۳	۰	۰	۲۵
۳۳	۱۳۹۹	۱۰	۱۰	۸	۵	۰	۰	۳۳
۳۱	۱۴۰۰	۱۰	۱۲	۵	۴	۰	۰	۳۱
۳۴	۱۴۰۱	۱۰	۱۰	۹	۵	۰	۰	۳۴
۲۸	۱۴۰۲	۱۰	۱۱	۱	۶	۰	۰	۲۸
۳۴	۱۴۰۳	۸	۱۵	۴	۷	۰	۰	۳۴
۲۷	۱۳۹۸	۷	۸	۶	۶	۰	۰	۲۷
۳۶	۱۳۹۹	۱۱	۱۰	۸	۷	۰	۰	۳۶
۳۵	۱۴۰۰	۸	۱۱	۸	۸	۰	۰	۳۵
۳۰	۱۴۰۱	۸	۱۱	۶	۵	۰	۰	۳۰
۳۶	۱۴۰۲	۷	۱۵	۸	۶	۰	۰	۳۶
۳۸	۱۴۰۳	۱۴	۱۱	۶	۷	۰	۰	۳۸
۴۰	۱۴۰۴	۸	۱۸	۱۱	۳	۰	۰	۴۰
۴۵۵	جمع	۱۲۹	۱۵۵	۹۴	۷۷	۰	۰	۴۵۵
%۱۰۰	درصد	%۲۸.۴	%۳۴.۱	%۲۰.۷	%۱۶.۹	%۰	%۰	%۱۰۰

نتایج جدول ۲ نشان می‌دهد که توزیع سؤالات آزمون‌ها به شدت نامتوازن است. سطوح ابتدایی یعنی دانش، درک و کاربرد بیش از ۸۰٪ سؤالات را به خود اختصاص داده‌اند و از سطوح بالا تنها سطح تحلیل ۱۶.۹٪ سؤالات را به خود اختصاص داده و هیچ سوالی در دو سطح ترکیب و ارزیابی یافت طرح نشده است. جدول ۳ پراکندگی سؤالات بر اساس سطوح طبقه بندی بلوم در طول بازه تحقیق را نشان می‌دهد.

**جدول ۳: سطح پوشش و تعداد سؤالات در هر مرحله از طبقه بندی بلوم**

سطوح طبقه بندی بلوم	تعداد سؤالات در این مرحله
دانش	۱۲۹
درک	۱۵۵
کاربرد	۹۴
تحلیل	۷۷
ترکیب	.
ارزیابی	.

جدول‌های ۲ و ۳ نشان دادند که توزیع سؤالات بر اساس سطوح شناختی کاملاً نامتوازن است. سه سطح پایین‌تر بلوم؛ یعنی دانش، درک و کاربرد بیش از ۸۰ درصد از کل سؤالات را تشکیل می‌دهند. این یافته به‌وضوح بیانگر تمرکز طراحان بر سطوح اولیه و بازتولیدی یادگیری است؛ جایی که دانش‌آموز صرفاً موظف به یادآوری یا تفسیر مطالب است و کمتر به تحلیل یا قضاوت خلاقانه می‌پردازد. در مقابل، دو سطح بالای «ترکیب» و «ارزیابی» به‌طور کامل در سؤالات امتحانات نهایی غایب بوده‌اند؛ این غیبت از دیدگاه آموزشی، نشانه‌ای از کم‌توجهی به رشد مهارت‌های تفکر انتقادی، خلاقیت و توانایی حل مسئله در دانش‌آموزان است. در ادامه برای پاسخ به سوال سوم و چهارم مطرح شده در این تحقیق، از آمار استنباطی آزمون دقیق فیشر استفاده شد. جدول ۴ نتایج این بررسی آماری را نشان می‌دهد.

**جدول ۴. نتایج آزمون کای مربع برای رابطه بین سطوح طبقه بندی بلوم و سال/نیم سال تحصیلی**

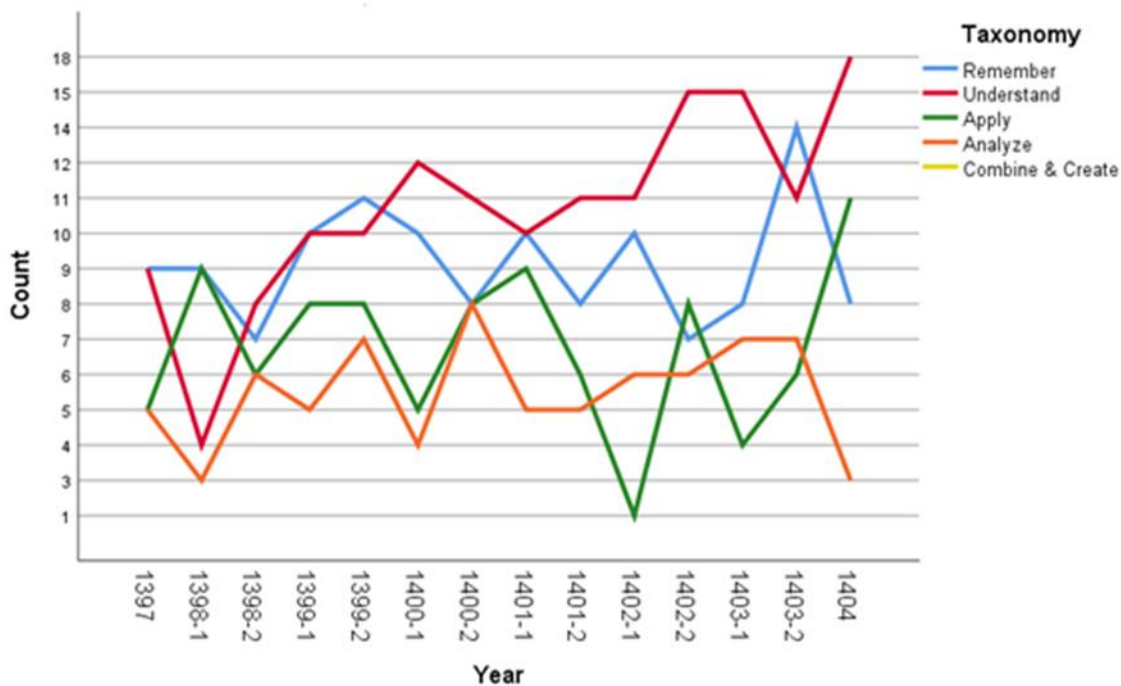
value	مقدار df	مقدار مقدار	مقدار Exact Sig (1-sided)	مقدار Exact Sig (2-Sided)	مقدار Asymptotic Sig (2-sided)
۱۷۶.۹۳۳	۱۶۹	۰.۳۲۲	۰.۰۰۰	۰.۰۰۰	۰.۰۰۰
۱۳۶.۶۷۸	۱۶۹	۰.۹۶۸	۰.۰۰۰	۰.۰۰۰	۰.۰۰۰
۰.۰۰۰					
۳.۰۶۷	۱	۰.۰۸۰	۰.۰۰۰	۰.۰۰۰	۰.۰۰۰
۵۶					

یادداشت: ۱۹۶ سلول (۱۰۰٪) دارای شمارش مورد انتظار کمتر از ۵ بودند، که نشان‌دهنده لزوم استفاده از آزمون دقیق فیشر به جای آزمون کای مربع است

همانطور که نتایج جدول ۴ نشان می‌دهد، از منظر آماری، اجرای آزمون دقیق فیشر که به‌دلیل کم‌بودن فراوانی مورد انتظار در بسیاری از سلول‌های جدول توافقی ضروری بود، نشان داد که رابطه‌ای آماری میان سال برگزاری آزمون و توزیع سطوح شناختی وجود دارد. این بدان معناست که تغییراتی در نحوه طراحی سؤالات در طول زمان رخ داده است، هرچند این تغییرات فاقد روند مشخص و خطی هستند. به‌عبارت دیگر، نمی‌توان ادعا کرد که آزمون‌ها در گذر زمان به سمت سطوح شناختی بالاتر حرکت کرده‌اند.

بررسی جزئی‌تر داده‌ها نشان می‌دهد که در برخی سال‌ها، به‌ویژه در آزمون‌های دی‌ماه ۱۴۰۳ و خردادماه ۱۴۰۴، تمرکز سؤالات از سطح «دانش» به سطح «درک» منتقل شده است؛ یعنی طراحان تلاش کرده‌اند از صرف حافظه‌محوری فاصله بگیرند و مهارت فهم و تفسیر را بیشتر بسنجند. با این حال، سهم سؤالات سطح «تحلیل» در کل

بازه‌ی زمانی کمتر از ۱۷ درصد بوده و دو سطح «ترکیب» و «ارزیابی» هیچ حضوری نداشته‌اند. این الگو به روشنی نشان می‌دهد که نظام ارزشیابی رسمی هنوز به مرحله‌ی سنجش سطوح بالای شناختی نرسیده است.



با توجه به نمودار ۱ نتایج آماری و محتوایی به‌طور هم‌زمان بر یک نکته کلیدی تأکید دارند: نظام طراحی آزمون‌های نهایی درس علوم و فنون ادبی هنوز در مرحله‌ی «یادگیری سطح پایین» قرار دارد و نیازمند بازنگری در رویکردهای طراحی سؤال است. سؤالات باید نه تنها حافظه‌ی اطلاعاتی را بسنجند، بلکه مهارت درک، تفسیر، استدلال، ترکیب ایده‌ها و ارزیابی انتقادی را نیز دربر گیرند تا فرآیند یادگیری به‌سوی عمق و خلاقیت سوق داده شود.

### بحث و نتیجه‌گیری

از جنبه‌ی نظری، این یافته‌ها هم‌راستا با مطالعات پیشین داخلی (فلسفی‌نژاد، ۱۳۹۵؛ سپاسی، ۱۳۸۵؛ میرآقایی و همکاران، ۱۳۹۴) است که همگی تأکید داشتند سؤالات امتحانات نهایی در ایران عمدتاً در سطوح پایین‌تر بلوم طراحی می‌شوند. از سوی دیگر، نتایج پژوهش حاضر با یافته‌های بین‌المللی نیز همخوانی دارد. پژوهش‌های عمر و عطان (۲۰۱۵)، اسکالی (۲۰۱۷) و کاور و سینگ (۲۰۱۹) نیز نشان دادند که گرایش به طرح سؤالات در سطوح پایین‌تر شناختی یک پدیده جهانی است، به‌ویژه در آزمون‌های استاندارد و سراسری که تصحیح عینی و یکسان اهمیت زیادی دارد. با وجود این، از نظر آموزشی، چنین رویکردی پیامدهای قابل توجهی دارد. سؤالاتی که صرفاً سطوح دانش و درک را می‌سنجند، موجب تثبیت یادگیری سطحی می‌شوند و فرصت تمرین تفکر تحلیلی یا خلاق را از دانش‌آموز می‌گیرند. این امر در بلندمدت باعث می‌شود دانش‌آموزان در مواجهه با متون ادبی و پرسش‌های باز، صرفاً به حفظ مطالب بسنده کنند و توانایی قضاوت ادبی یا تحلیل انتقادی در آنان پرورش نیابد. نمودار شماره ۱ نیز به روشنی نشان می‌دهد که گرچه در برخی سال‌ها سهم سؤالات تحلیلی اندکی افزایش یافته است، اما الگوی کلی تغییرات نوسانی بوده و فاقد جهت‌گیری نظام‌مند است.

بررسی سؤالات امتحان نهایی در هر رشته و درسی بر اساس طبقه‌بندی بلوم می‌تواند مشخص کند که طراحان تا چه اندازه‌ای دانش‌آموزان را برای یادگیری و تقویت مهارت تفکر انتقادی آماده می‌کنند یا یادگیری این فرآیند را در حد پایه حفظ می‌کنند. بررسی سؤالات علوم و فنون ادبی نشان داد که طراحان، سطوح پایین شناختی را بیشتر مدنظر قرار

داده و هر چه به سمت سطوح بالاتر رفتیم، از تعداد سؤالات کم شد و در نهایت سؤالات در سطح تحلیل متوقف شدند. از آنجا که سطوح دانش، درک و کاربرد پایه‌های تقویت مهارت تفکر انتقادی هستند، باید متوازن وجود داشته باشند، اما اینکه بیش از ۸۰٪ سؤالات را به خود اختصاص بدهند، نشان از عدم توجه به سطوح بالاتر شناختی است. سؤالات در سال‌های مورد اشاره دارای نوسان بوده است. با توجه به نمودار شماره ۱ مشخص شد که در سؤالات دی ماه ۱۴۰۳ و خرداد ماه ۱۴۰۴ توجه طراحان سؤالات امتحانی از سطح دانش به سمت سطح درک معطوف شده است. شاید یکی از دلایل عدم وجود سؤالی در سطوح ترکیب و ارزیابی، نبودن جواب مشخص و در نهایت امکان تصحیح سلیقه‌ای توسط مصحح امتحانات باشد؛ اما این دلیل قانع کننده و علمی نمی‌تواند باشد که طراحان از این دو سطح که بالاترین سطوح شناختی هستند غافل شوند.

این تحقیق با توجه به نتایج خود برخی پیشنهادات را برای طراحان آزمون علوم و فنون ادبی و دانش آموزان به شرح زیر ارائه می‌دهد:

نتایج این تحقیق نشان داد که توجه طراحان تنها به ۴ سطوح شناختی بلوم بوده و طراحان از سطوح ترکیب و ارزیابی غافل مانده‌اند. توزیع سؤالات در سطوح دانش، درک، کاربرد و تحلیل نشان داد که طراحان می‌توانند به شکل متوازن سؤالات را به نحوی طراحی کنند که نیازهای دانش‌آموزان برای یادگیری و انتظار معلمان برآورده شود. اگر همین توازن در دو سطح دیگر نیز برقرار می‌شد آنگاه می‌شد ادعا کرد که این آزمون جهت سنجش و تقویت مهارت تفکر انتقادی مناسب است.

از دانش‌آموزان نیز انتظار می‌رود که با مهارت تفکر انتقادی و کارکرد برنامه درسی و آزمون در پرورش این مهارت مهم آشنا باشند. مطالب این پژوهش هم بر اهمیت و تفکر انتقادی و کارکرد بهتر این مهارت در ارتباط با رشته‌های خاص تاکید داشته و ارتباط بین این مهارت و سطوح شناختی بلوم را نشان داد. دانش‌آموزان می‌توانند با مطالعه پژوهش حاضر به درک درستی از تفکر انتقادی و ارتباط خواندن یک درس و رشته خاص (در این پژوهش درس علوم و فنون ادبی) با این مهارت را درک کنند. این فهم، هم در درونی‌سازی محتوای درسی و هم در ارائه پاسخ‌های دقیق‌تر به پرسش‌ها، به دانش‌آموزان یاری می‌رساند. آگاهی از سطوح شناختی و طبقه‌بندی بلوم به یادگیرندگان نشان می‌دهد که معلم چه سطحی از عملکرد یا انتظار را از آنها مدنظر دارد. در چنین شرایطی، دانش‌آموزان نیز قادر خواهند بود خواسته‌های معلم را بهتر برآورده کنند. البته پیش‌نیاز این امر آن است که خود معلم شناخت کافی از این فرایندها داشته باشد و بتواند آن را به شاگردان منتقل کند. این تعامل دوسویه سبب می‌شود که معلم برنامه درسی و آزمون‌های کلاسی را به گونه‌ای سازمان دهد که همه سطوح طبقه‌بندی بلوم را در برگیرد؛ در نتیجه، هم مسیر پرورش تفکر انتقادی هموار می‌شود و هم دانش‌آموز با درکی روشن‌تر در این مسیر گام برمی‌دارد.

### مشارکت نویسندگان

میزان مشارکت نویسندگان در نگارش مقاله به طور یکسان است.

### تشکر و قدردانی

از داوران مقاله که با ارائه نظرات ارزشمند خود موجب ارتقا و کیفیت مقاله گردیدند، بی‌نهایت سپاسگزاریم.

### تعارض منافع

هیچ‌گونه تعارض منافع وجود ندارد.

#### COPYRIGHTS



©2021 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

## 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 ۵۷-۷۸.
- سیف، علی‌اکبر (۱۳۹۹). *اندازه‌گیری، سنجش و ارزشیابی آموزشی* (ویرایش هفتم). تهران: دوران.
- فلسفی‌نژاد، محمدرضا و همکاران (۱۳۹۵). ویژگی‌های روان‌سنجی امتحانات نهایی سال سوم متوسطه و قابلیت آنها در گزینش داوطلبان ورود به دوره‌های کارشناسی، *فصلنامه اندازه‌گیری تربیتی*، دوره ۶، ش ۲۳، صص ۴۵-۷۶.
- میرآقایی، علی‌عباس و همکاران (۱۳۹۴). بررسی و تحلیل شاخص‌های روان‌سنجی و سطوح حیطة شناختی سوالات امتحانات نهایی دروس ریاضیات و علوم پایه سوم راهنمایی شهرستان خرم‌آباد، *مجله روان‌شناسی مدرسه*، دوره ۴، ش ۳، صص ۱۰۲-۱۱۸.

## Reference

- Adom, D., Mensah, J. A., & Dake, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education*, 9(1), 109–119. <https://doi.org/10.11591/ijere.v9i1.20457>
- Ahmadi Safa, M., & Jafari, F. (2016). The washback effects of high school third grade exam on EFL teachers' methodology, evaluation and attitude. *Applied Research on English Language*, 5(1), 77–98. [https://are.ui.ac.ir/article\\_20397\\_0.html](https://are.ui.ac.ir/article_20397_0.html)
- Airasian, P. W., & Miranda, H. (2002). The role of assessment in the revised taxonomy. *Theory Into Practice*, 41(4), 249–254.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: complete edition*. Addison Wesley Longman, Inc..
- Ennis, R.H. (1962). A concept of critical thinking. *Harvard Educational Review*, 29, 128-136.
- Ennis, R. H. (1993). Critical thinking assessment. *Theory Into Practice*, 32(3), 179–186. <https://doi.org/10.1080/00405849309543594>
- 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>
- 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>

- Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2008). Authentic assessment, student and teacher perceptions: The practical value of the five-dimensional framework. *Journal of Vocational Education and Training*, 60(4), 401–423.
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items*. Lawrence Erlbaum.
- Jenkins, H., & Schultz, S. (2013). The Final Exam — To Have or Have Not. In *Proceedings of the 2013 ASEE Southeastern Section Conference*. American Society for Engineering Education. <http://sites.asee.org/se/wpcontent/uploads/sites/56/2021/04/2013ASEESE152.pdf>
- Kaur, H., & Singh, G. (2019). A study on the cognitive levels of English examination questions in Indian secondary schools. *Journal of Education and Practice*, 10(12), 12–19.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, 41(4), 212–218. [https://doi.org/10.1207/s15430421tip4104\\_2](https://doi.org/10.1207/s15430421tip4104_2)
- Kurfiss, J. G. (1988). *Critical thinking: Theory, research, practice, and possibilities*. Washington, DC: Association for the Study of Higher Education. Retrieved from <http://eric.ed.gov/?id=ED304041>
- Lai, E. (2011). *Critical thinking: A literature review*. Pearson. <http://www.pearsonassessments.com/research>
- Lipman, M. (1988). Critical thinking: What can it be? *Educational Leadership*, 38-43
- 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>
- McPeck, J. E. (2017). *Critical thinking and education*. Abingdon, UK: Routledge.
- Omar, N., & Atan, N. A. (2015). Analysis of examination questions using Bloom's taxonomy. *Procedia – Social and Behavioral Sciences*, 172, 613–620.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307–332. <https://doi.org/10.3102/00346543062003307>
- 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>
- Pithers, R. T., & Soden, R. (2000). Critical thinking in education: A review. *Educational Research*, 42(3), 237–249. <https://doi.org/10.1080/001318800440579>
- 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>

- Scully, D. (2017). Constructing multiple-choice items to measure higher-order thinking: Applying Bloom's taxonomy. *Practical Assessment, Research & Evaluation*, 22(4), 1–13.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Pearson Education Limited. (Republished 2013 by Routledge)  
<https://www.routledge.com/The-Power-of-Tests-Shohamy/p/book/9780582423367>
- Vygotsky, L. S. (1962). *Language and thought* (pp. 11193-000). Cambridge, MA: MIT Press.
- Yildirim, B., & Özkahraman, Ş. (2011). Critical thinking theory and nursing education. *International Journal of Humanities and Social Science*, 1(17, Special Issue), 176–185.