

تحلیل سؤال‌های آزمون ورودی دانشگاه و عملکرد داوطلبان در پاسخگویی به آنها بر اساس سطوح حیطة شناختی بلوم^۱.

مرضیه کرامتی نوجه ده سادات^۲
اعظم غلامی^۴

چکیده:

هدف پژوهش حاضر، تحلیل سؤال‌های آزمون ورودی دانشگاه و عملکرد داوطلبان در پاسخگویی به آنها بر اساس سطوح حیطة شناختی بلوم. پژوهش از نوع کاربردی و تحلیلی - توصیفی است. جامعه آماری، داوطلبان آزمون ورودی در سال‌های ۱۳۹۵ تا ۱۳۹۷ بودند که بر مبنای جدول کرجسی - مورگان، ۳۸۴ نفر از شهر تهران انتخاب شدند. ابزار گردآوری داده‌ها، کارنامه آزمون داوطلبان ورودی ۱۳۹۵ تا ۱۳۹۷ بود. پس از گردآوری داده‌ها، تجزیه و تحلیل آنها بر اساس کاپای کوهن و آمار توصیفی بود. ابتدا از طریق تحلیل محتوای سؤال‌های آزمون سطح آنها بر اساس سطوح حیطة شناختی بلوم مشخص شد، سپس، عملکرد داوطلبان در سؤال‌های هر یک از سطوح مورد تحلیل قرار گرفت. بدین منظور از آزمون‌های ۱۳۹۵ تا ۱۳۹۷ استفاده شد. یافته‌ها نشان داد سؤال‌ها بالای ۴۰ درصد در سطح فهمیدن است و عملکرد داوطلبان در هر یک از سطوح نامطلوب بود (کمتر از ۵۰ درصد). از یافته‌ها نتیجه‌گیری شد غالب سؤال‌های در حد پائین بود. پیشنهاد می‌شود سؤال‌ها در سطوح بالای حیطة شناختی بلوم تدوین شود تا بتواند

۱. تاریخ وصول: ۱۴۰۰/۱۰/۲۴

تأیید نهایی: ۱۴۰۲/۰۱/۱۶

۲. مقاله حاضر اقتباس شده از طرح پژوهشی به نام «میزان پاسخگویی داوطلبان آزمون ورودی سال‌های ۹۵ تا ۹۷ به سوال‌های درس زیست شناسی در حیطة شناختی بلوم» دانشگاه فرهنگیان ۱۳۹۷ است.

۳. استادیار گروه علوم پایه دانشگاه فرهنگیان، تهران. ایران. نویسنده مسؤول، mk.sadat@gmail.com.

۴. استادیار گروه علوم پایه دانشگاه فرهنگیان، تهران. ایران.

گزینش دانشجو در رشته‌های مختلف از جمله رشته‌های دانشگاه فرهنگیان به صورت مطلوب باشد.

واژه‌های کلیدی: آزمون ورودی علوم تجربی؛ درس زیست‌شناسی؛ حیطه شناختی بلوم؛ میزان پاسخگویی داوطلبان.

۱. مقدمه

یکی از نیازهای جوامع بشری، تعلیم و تربیت است که هر کشوری بر اساس فرهنگ جامعه خود، در طول رشد و تکامل انسان‌ها، برنامه‌ریزی لازم کرده است. ساختار پیچیده فرهنگ و اجتماعات بشری و الزام تربیت آنان و آماده کردن آنها برای یک زندگی متحول و دگرگون، دست‌اندرکاران تعلیم و تربیت و همچنین روان‌شناسان را بر آن داشته است که در پی برنامه‌ریزی و سازماندهی صحیح برای آموزش جوامع بشری باشند.

تعلیم و تربیت موجب رشد و تعالی جنبه‌های مختلف فکری و شکوفایی توانایی‌ها و استعدادها می‌شود، طوری که آموزش و پرورش، مخصوصاً مدارس نقش بسیار مهمی در شکل‌گیری تعلیم و تربیت دارند. هر نظام آموزشی دارای عناصر آموزشی درون‌داد، برون‌داد، فرایند و بازخورد است که طراحی آموزشی به وسیله طراحان آموزشی باعث شناخت درستی از آموزش در زمینه‌های تحلیل، هدف، محتوا، توالی، راهبردهای آموزشی، مواد آموزشی و در نهایت ارزشیابی را فراهم می‌کند که برای تسهیل در آموزش و یادگیری بکار می‌رود (نوروزی، ۱۳۹۸).

ارزشیابی فرایندی است که در طول آموزش و به صورت مداوم و مستمر صورت می‌گیرد (موریسون^۱، ۲۰۰۴). یکی از اهداف آموزش و پرورش، ارزیابی دانش‌آموزان از آموزش و یادگیری آنان، به منظور سنجش میزان پیشرفت تحصیلی، آماده کردن آنها برای شرکت در آزمون‌های ورودی و در نهایت ادامه تحصیل و کسب دانش مهارت و نگرش‌های لازم برای زندگی آتی است. ارزشیابی شامل داوری ارزشی در باره کیفیت موضوع مورد ارزشیابی است که مبتنی بر تحلیل و تفسیر اطلاعات گردآوری شده است که به کمک اندازه‌گیری و سنجش است و هرچه اطلاعات گردآوری شده جامع‌تر باشد ارزشیابی درست‌تر صورت می‌گیرد (سیف، ۱۳۹۸).

درس زیست‌شناسی، یکی از مهم‌ترین و کلیدی‌ترین درس رشته علوم تجربی در دوره متوسطه است. منتهی تدریس آن بیشتر به صورت سخنرانی و از نظر اهداف آموزشی فقط سطوح پائین حیطه شناختی را پوشش می‌دهد. حجم بالای کتب درسی و ساختار محتوای نه‌چندان غنی، کمبود وقت، ارزشیابی‌های ناکارآمد، روش سنتی، عدم استفاده از آزمایشگاه یا کامل نبودن آن و کمبود فرصت و مشکل بودجه برای بازدیدها و گردش‌های

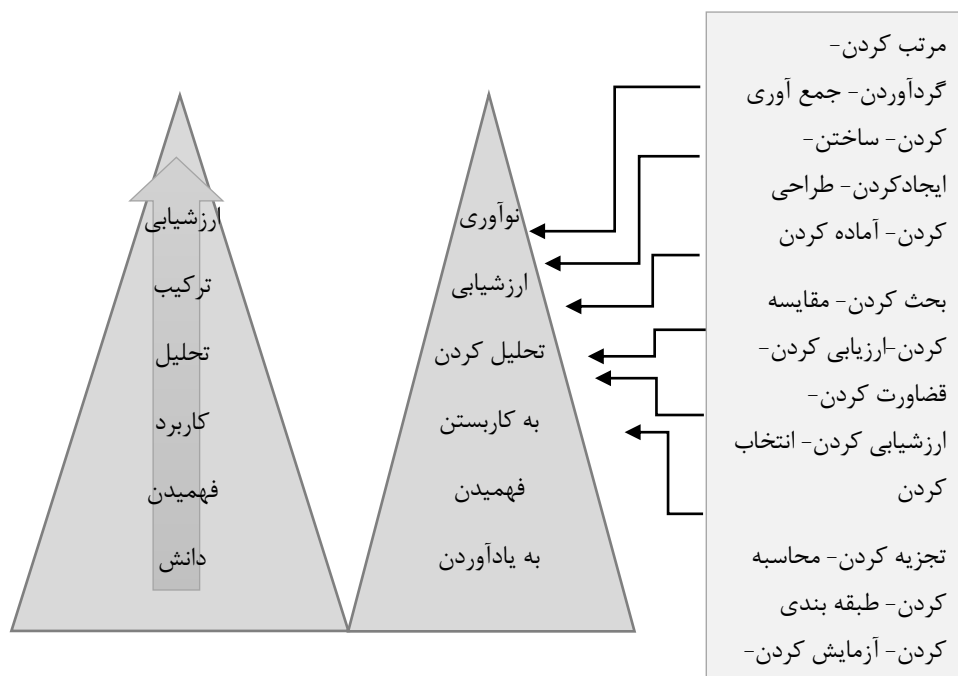
علمی، تراکم بسیار بالای دانش‌آموزان در کلاس درس و نامناسب بودن وضعیت فیزیکی کلاس‌ها و آزمایشگاه‌ها، مشکلات شخصی برخی دانش‌آموزان، نداشتن انگیزه کافی برای درس خواندن و ترک تحصیل و عدم رغبت برخی معلمان در به‌کارگیری روش‌های مؤثر و فعال تدریس از مهم‌ترین چالش‌های موجود در آموزش زیست‌شناسی است (کارگری، ۱۳۹۵: ۳۰/۱۶-۱۴).

آموزش زیست‌شناسی بخشی از فرایند آموزشی دانشگاه فرهنگیان است که دانشجویان برای کسب دانش، مهارت و نگرش معلمی مبادرت می‌ورزند. بر اساس سند تحول بنیادین جمهوری اسلامی ایران در افق ۱۴۰۴ شایسته است ایران کشوری توسعه‌یافته و دارای جایگاه اول اقتصادی علمی و فناوری در سطح منطقه و دارای اشتغال کامل باشد و برای رسیدن به این اهداف می‌بایست «مهارت» جایگزین «مدرک»، «نوآوری و خلاقیت» به جای ایستایی، «کارآفرینی» جایگزین «کارگذاری»، «روش مدرن به تولید» به جای «روش‌های سنتی تولید» قرار گیرد (میرکی قرایی، ۱۳۹۳: ۱۰-۴۷-۴۴)؛ بنابراین در هر رشته دانشگاهی کسب مباحث نظری و عملی از اهمیت خاصی برخوردار است تا بتواند نیروهای کارآمد تربیت کند. برای رسیدن به این اهداف، آموزش از دوره ابتدایی تا دبیرستان بر مبنای اصول روش‌های تعلیم و تربیت صورت می‌گیرد؛ بنابراین مهم است که برای گزینش دانشجو در دانشگاه از چه شیوه‌های ارزشیابی استفاده شود.

برای گزینش دانشجو از ملاک و معیارهای مناسبی استفاده می‌کنند. آزمون‌های ورودی دانشگاه‌ها، با در نظر گرفتن ملاک‌های معین و استاندارد، به اندازه‌گیری سطح توانایی داوطلبان ورود به این مراکز جهت گزینش آنان اقدام می‌کنند. بدیهی است ضرورت و اهمیت دقت چنین اندازه‌گیری‌هایی، با توجه به تأثیرات مستقیمی که در آینده افراد و به تبع آن جامعه خواهد داشت، غیرقابل انکار است؛ لذا پیوسته سعی بر این خواهد بود که اولاً با به‌کارگیری ملاک‌های صحیح و معتبر و رعایت اصول لازم در طراحی و ساخت این آزمون‌ها، روایی و پایایی هرچه بیشتر سؤال‌های آزمون تأمین گردد و ثانیاً با استفاده از تئوری‌های اندازه‌گیری و مدل‌های مناسب جهت تحلیل پاسخ‌ها بتوان دقت برآورد سطح توانایی آزمودنی‌ها را افزایش داد (معلمی اوره، ۱۳۸۷).

ابزار زیست‌شناسی بلوم^۱، یک ابزار ارزیابی مبتنی بر طبقه‌بندی بلوم است که به دانش‌آموزان کمک می‌کند تا در فعالیتهای آموزشی، مهارت‌های تحصیلی و فراشناخت خود را ارتقا بخشند (کرو^۲، دیرکس^۳، وندروت^۴، ۲۰۰۸: ۷/ ۳۶۸-۳۸۱). بلوم و همکارانش هدفهای حوزه شناختی را به شش طبقه دانش^۵، فهمیدن^۶، کار بستن^۷، تحلیل^۸، ترکیب^۹ و ارزشیابی^{۱۰} طبقه‌بندی کردند. در سال ۲۰۰۱ تغییراتی در طبقه‌بندی بلوم به نام طبقه‌بندی شناختی تجدیدنظر شده^{۱۱} صورت گرفت که یک طبقه‌بندی برای یادگیری، آموزش و سنجش است (سیف، ۱۳۹۸). سطح اول هر دوی اینها مبتنی بر دانش است که به‌موجب آن صرفاً اطلاعات منتقل می‌شود به‌طوری‌که فراگیران دانش پدیده را می‌آموزند؛ سطح بعدی در مورد ایجاد درک آن پدیده است؛ مرحله سوم کاربرد، در جایی است که دانش در متن زندگی واقعی به کار می‌رود. سپس تحلیل آن پدیده و چگونگی ایجاد ارتباط با دیگر مجموعه اطلاعات فراهم می‌شود. پس از اتمام این مرحله، مرحله ارزشیابی فرامی‌رسد تا فراگیران بتوانند اطلاعات کسب شده را با توجه به کاربرد آن و به‌طور انتقادی ارزیابی کنند که چگونه می‌توان متناسب با نیاز آن را اصلاح کرد. سرانجام، مرحله ترکیب یا نوآوری است که بالاترین مرحله از سطح دستاورد محسوب می‌شود و به‌موجب آن اطلاعات یا ایده جدیدی بر اساس تجربیات یادگیری از سطوح گذشته ایجاد می‌شود (پل^{۱۲}، نایک^{۱۳}، رانه^{۱۴}، پاور^{۱۵}، ۲۰۱۲: ۱۴۴-۱۴۸). در طبقه‌بندی تجدیدنظر مرحله ششم با مرحله پنجم جابه‌جا شده است. شکل ۱ بیان‌کننده هر دو حیطة شناختی به‌طور خلاصه است.

-
4. Bloom Biology Tool
 5. crowe
 3. Dirks
 4. Wenderoth
 6. Knowledge
 6. Comprehension
 7. Application
 - Analyze 8.
 9. Synthesis
 10. Evaluation
 11. Revised Cognitive Classification
 12. Paule
 13. Nike
 14. Rane
 15. Pawar



شناختی بلوم و حیطه شناختی تجدیدنظرشده (یانچیندا^۱ و همکاران ۲۰۱۶؛ حیدر^۲ و بامنی^۳، ۲۰۱۶).

در آموزش علمی، به ویژه در آموزش زیست‌شناسی، معلمان از انواع ابزار اندازه‌گیری، آزمون‌های چندگزینه‌ای، آزمون‌های صحیح و غلط، تطبیق آزمون‌ها، آزمون‌های مقدماتی، آزمون‌های پاسخ کوتاه، امتحانات شفاهی، آزمون‌های عملکرد، و یا نوشته‌های مداد - کاغذی استفاده می‌کنند. امتحان مداد - کاغذی طبق نظر دیویس، امتحان یا تست کتبی، ابزار قدرتمند آموزشی در آموزش علمی است (دیویس^۴، ۱۹۹۳). سؤال‌های پرکردنی، صحیح و غلط و چندگزینه‌ای می‌توانند سطح تجزیه و تحلیل دانش‌آموزان را بسنجد. سؤال‌هایی که تجزیه و تحلیل هستند و در کلاس درس مطرح می‌شوند اگر به‌عنوان سؤال در امتحان آیند در سطح یادآوری قرار می‌گیرند. سؤال‌های چندگزینه‌ای می‌توانند برای سنجش سؤال‌های ارزشیابی نیز استفاده گردند به شرط آنکه ارزش نسبی و

1. Yanchinda
2. Hyder
3. Bhamani
4. Davis

یا شایستگی را تعیین کنند. از آنجائی که همه پاسخ‌ها در سؤال‌های چند گزینه‌ای ارائه می‌شوند، در نتیجه نمی‌توانند سطح ترکیب را بیان کنند و این مسئله نیاز و ضرورت نوآوری و خلاقیت را حذف می‌کند (ادوویچ^۱، ۱۹۹۶). معلمان زیست‌شناسی معمولاً از امتحان‌های کتبی برای تعیین اینکه آیا دانش‌آموزان روش‌های مطالعه، حقایق علمی، مطالب درسی، قوانین، نظریه‌ها و نوآوری‌های زیست‌شناسی را یاد می‌گیرند یا خیر، اعمال می‌کنند. برای ارزیابی تدریس زیست‌شناسی در تمام دوره‌های دبیرستان، تعیین کیفیت پرسش‌هایی که در آزمون‌های مدرسه ارائه می‌شود، مهم است. زیرا سؤال‌های سطح بالای شناختی می‌توانند به‌عنوان سؤال‌هایی تعریف شوند که نیاز به دانش‌آموزان برای تفکر عمیق و استفاده از مهارت‌های استدلال دارند. به‌منظور بهبود کیفیت آموزش، اعتقاد بر آن است که باید بتوان به‌خوبی پرسش کنند. درخواست سؤال‌های خوب، تعامل بین معلم و دانش‌آموزان را افزایش می‌دهد (آلن^۲ و تانر^۳، ۲۰۰۵: ۴/۲۶۲-۲۶۸).

سؤال‌هایی که برای آزمون استفاده می‌شود، شایسته است به ضریب دشواری^۴ و ضریب تمیز^۵ نیز توجه شود. زیرا یکی از راه‌های تحلیل انواع آزمون‌هاست. ضریب دشواری به درصد کل آزمون‌شوندگانی که به یک سؤال پاسخ صحیح می‌دهند و میزان آسان بودن یا دشوار بودن یک سؤال را برای آزمون‌شوندگان بیان می‌کند. ضریب تمیز به بررسی قدرت تمایزگذاری یا تشخیص بین گروه قوی و ضعیف را مشخص می‌کند. یعنی سؤال تا چه اندازه می‌تواند گروه قوی را از گروه ضعیف تفکیک کند (سیف، ۱۳۹۸). این مسئله نیز در طراحی سؤالات آزمون باید موردتوجه قرار گیرد که چگونه سؤالات آزمون می‌تواند دو گروه را از هم تفکیک دهد.

پژوهشگران داخلی و خارجی برای ارزشیابی دانش‌آموزان و تعیین میزان یادگیری آنها در دوره‌های تحصیلی و متعاقب آن آزمون سراسری، مطالعات مختلفی انجام دادند که به تعدادی از آنها می‌پردازیم. کاوه، خدایی، موسوی، مقدم‌زاده و یونسسی (۱۳۹۸) با ارائه مدلی برای پیوند زدن نمره‌های سوابق تحصیلی با نمره‌های آزمون ورودی به دانشگاه بیان می‌دارد که روش همصدک سازی کمتر به گروه وابسته است، به صورتی که با به‌کارگیری

1. Udovic
2. Allen
3. Tanner
4. Difficulty index
5. Discriminative index

این روش، نمره‌های آزمون برای تصمیم‌گیری با یکدیگر قابل مقایسه هستند. فتح‌آبادی، شلانی و صادقی (۱۳۹۶) تأثیر آزمون سراسری بر تدریس و ارزشیابی از دیدگاه معلمان مورد بررسی قرار داد. آنها معتقد بودند، آزمون سراسری بر ترجیحات، آموزش و تدریس معلمان، انگیزه فراگیران برای یادگیری، هدف‌گذاری و نگرش معلمان، نحوه ارزشیابی آنها و نگرش به کلاس‌های کنکوری تأثیرات معنی‌داری دارد. درویش (۱۳۹۶)، تحقیقی در باره ارزیابی سؤال‌های شیمی آزمون سراسری سال ۱۳۹۵ با استفاده از نظریه سؤال - پاسخ^۱ به عمل آورد. نتایج حاصل بیانگر ضریب دشواری^۲، ضریب تمیز^۳ و ضریب پایایی^۴ بود. علاوه بر این با استفاده از نظریه کلاسیک، ۳۴ سؤال از ۳۵ سؤال بخش شیمی دشوار بود. همچنین از نظر جداسازی آزمودنی‌های قوی و ضعیف از یکدیگر (ضریب تمیز) ۹ سؤال کارآمدتر محسوب شد. میرزایی تبار (۱۳۹۴) در پژوهش سؤال‌های آزمون ورودی در درک مطلب زبان انگلیسی، معتقد است بررسی سؤال‌های آزمون بر اساس طبقه‌بندی بلوم ضروری است. تلاش معلمان در جهت تشویق دانش‌آموزان به سه سطح بالای تفکر است. هرچند سؤال‌های آزمون ورودی در سطوح پایین تفکر، تمرکز داشته است. میری (۱۳۹۴) در پژوهش مربوط به ویژگی‌های روان‌سنجی بخش فیزیک آزمون سراسری بیان می‌دارد که کلیه آزمون‌های درس فیزیک آزمون ورودی ۱۳۹۰-۱۳۸۸ تک‌بعدی بوده و استقلال موضعی این آزمون برقرار است. نتایج حاصل از تحقیق سلیمان‌زاده (۱۳۹۳) در باره سؤال‌های شیمی امتحان نهایی و آزمون سال‌های ۱۳۹۲-۱۳۹۰ اعلام می‌دارد که سؤال‌های آزمون در سطح پایین سطوح حیطة شناختی بلوم است و بیشترین درصد به سطوح درک و فهم تعلق گرفته است. سؤال‌های امتحان داخلی و نهایی بیشتر به سطح دانش توجه کرده است. باین‌حال سؤال‌های آزمون ورودی از این نظر نسبت به امتحان داخلی وضع بهتری داشته است و تقریباً امتحان نهایی و هماهنگ از نظر اصول آزمون‌سازی نسبت به امتحان داخلی برتر است. آبکار (۱۳۹۱) درباره ویژگی‌های روان‌سنجی سؤال‌های آزمون ورودی ۱۳۸۹ از نظر تئوری سؤال - پاسخ بیان می‌دارد که اکثر سؤال‌های دارای دشواری بسیار زیاد و ضریب تشخیص بالا و پایایی مطلوبی است. همچنین نتایج در درس زیست‌شناسی، از لحاظ دشواری در مدل یک، دو و سه پارامتری

1. Item-Response-Theory (IRT)

2. Difficulty index

3. Discriminative index

4. Reliability index

در حد متوسط و ضریب تشخیص در مدل یک پارامتری دارای یک مقدار ثابت و در مدل دو و سه پارامتری در حد متوسط را نشان می‌دهد و میزان پارامتر حدس در مدل سه پارامتری در حد بسیار کم است؛ بنابراین می‌توان گفت، نتایج پژوهشگران داخلی بر روی امتحانات نهایی و آزمون در دروس مختلف، حاکی از سطوح پایین حیطه شناختی است که سؤال‌های آزمون از جنبه حیطه شناختی بلوم و نظریه سؤال - پاسخ مورد بررسی قرار گرفته بود.

پژوهش‌های خارجی مانند پژوهش هیرشی^۱ (۲۰۱۶) بیان می‌کند که آزمون‌های سرنوشت‌ساز بر روش‌شناسی معلمان در کلاس تأثیر زیادی می‌گذارد. هانگ و لی^۲ (۲۰۱۵) در باره رابطه نحوه ارزشیابی و برنامه درسی بیان می‌دارد که چگونه برنامه درسی می‌تواند بر ارزشیابی و همچنین ارزشیابی بر برنامه درسی تأثیر متقابل داشته باشند. لورد^۳ و باویسکار^۴ (۲۰۰۷) در باره انتقال دانش‌آموزان از یادآوری اطلاعات به فهمیدن یا سطوح بالاتر از طریق ایجاد سؤال‌های علمی تحقیق کرد. از دیدگاه وی، تمام آموخته‌های دانش‌آموزان به صورت جزئی است و مفاهیم مبتنی بر واقعیت است؛ لذا ارزیابی از آنها بر اساس توانایی‌های آنان در یادآوری و خلاصه‌کردن اطلاعات است. چنین فراگیرانی، تمرکز خود را بر روی اصطلاحات و تعاریف معطوف می‌دارند و توجه کمی به تجزیه و تحلیل مطالب علمی دارند. براین اساس معلمان به‌تراست، توجه زیادی به طرح سؤال‌هایی در سطوح بالای طبقه‌بندی بلوم داشته باشند. منت^۵ (۲۰۰۶) در پژوهش تحت عنوان «باز هم بیشتر از حقایق» به بررسی استانداردهای مرجع آموزش در کالیفرنیا پرداخته است و از طبقه‌بندی بلوم به‌عنوان یک استاندارد برای ارزیابی استفاده کرده است. به اعتقاد وی، رسیدن به سطوح یادآوری و فهمیدن به‌عنوان معیارهای اصلی آموزش در مدارس است. همچنین سطوح دانش و فهم فراگیران را از رسیدن به آفرینندگی باز می‌دارد و خلاقیت در سطح بالاتر از ترکیب قرار دارد و معلمان را به پرورش خلاقیت در دانش‌آموزان فرامی‌خواند. در پژوهش‌های خارجی نیز برای ارزشیابی سؤال‌ها از طبقه‌بندی بلوم به‌عنوان یک استاندارد برای ارزشیابی استفاده شده است.

1. Hirsch
2. Huang & Lee
3. Loard
4. Baviskar
5. Ment

فیزیک، شیمی و زبان از جمله دروسی هستند که ضریب پایین تری نسبت به درس زیست‌شناسی در آزمون علوم تجربی دارند. در نتیجه می‌توان گفت، زیست‌شناسی با ضریب بالا نقش مهمی در انتخاب رشته دانش‌آموزان دارد (سازمان سنجش، ۱۳۹۸). باتوجه‌به پژوهش‌های اشاره شده، تحقیقات مبسوطی در باره سؤال‌های درس زیست‌شناسی به دلیل اهمیت این درس در سرنوشت دانش‌آموزان به عمل نیامده است. علاوه بر این، اهمیت اهداف آموزشی بر اساس سطوح طبقه‌بندی بلوم و ضرورت ارزیابی سؤال‌ها از دیدگاه صاحب‌نظران در این مورد و میزان پاسخگویی داوطلبان به سؤال‌های زیست‌شناسی منجر به انجام پژوهش در این زمینه گردید؛ بنابراین پژوهش حاضر درصدد پاسخگویی به سؤال‌های ذیل است.

سؤال ۱: سؤال‌های زیست‌شناسی آزمون سال‌های ۹۵ تا ۹۷ بر اساس سطوح مختلف حیطة شناختی چگونه است؟

سؤال ۲: میزان عملکرد داوطلبان به پاسخگویی سؤال‌های درس زیست‌شناسی آزمون ورودی رشته علوم تجربی بر حسب سطوح حیطة شناختی بلوم در سال‌های ۹۷-۱۳۹۵ چگونه است؟

۲. بحث و بررسی

۲-۱. روش‌شناسی پژوهش

روش پژوهش حاضر، تحلیلی - توصیفی و از نظر دسته‌بندی تحقیقات برحسب هدف، کاربردی است. جامعه آماری، داوطلبان آزمون ورودی در سال‌های ۱۳۹۵ تا ۱۳۹۷ بودند که بر مبنای جدول کرجسی - مورگان، ۳۸۴ نفر از شهر تهران و به‌صورت تصادفی به‌عنوان نمونه آماری انتخاب شدند. ابزار گردآوری داده‌ها، کارنامه آزمون داوطلبان ورودی ۱۳۹۵ تا ۱۳۹۷ بود. پس از گردآوری داده‌ها، تجزیه و تحلیل آنها بر اساس آمار توصیفی بود. ابتدا از طریق تحلیل محتوای سؤال‌های آزمون در سال‌های ۱۳۹۵ تا ۱۳۹۷ بر اساس سطوح حیطة شناختی بلوم مشخص شد، سپس، عملکرد داوطلبان در پاسخ داده شده به سؤالات آزمون ورودی در هر سه سال، مورد تجزیه و تحلیل قرار گرفت.

به‌منظور بررسی تک‌تک سؤال‌های آزمون ورودی درس زیست‌شناسی، بر اساس سطوح مختلف حیطة شناختی، سه نفر از متخصصان رشته زیست‌شناسی و مدرس در دوره دبیرستان و دانشگاه فرهنگیان و مؤلفین کتاب زیست‌شناسی دوره متوسطه دوم دعوت به

عمل آمد که اشراف کامل به سطوح حیطة شناختی بلوم و سؤال‌های آزمون درس زیست‌شناسی رشته علوم تجربی داشتند. آنها به‌عنوان ارزشیاب ۱ و ۲ و ۳ به بررسی و توصیف عینی، واقعی و منظم خصوصیات هر ۱۵۰ سؤال بر اساس ویژگی‌های سطوح مختلف حیطة شناختی بلوم پرداختند و نتایج را در جدولی از پیش تعیین شده علامت گذاشتند. کدگذاری بدین صورت بود که به دانش=۱، فهمیدن=۲، کاربرد=۳، تحلیل=۴، ترکیب=۵ و ارزشیابی=۶ داده شد. جامعه آماری، شامل سؤال‌های زیست‌شناسی گروه C در سه سال ۹۵، ۹۶ و ۹۷ آزمون ورودی ایران بود که از طریق سازمان سنجش، سؤال‌ها و پاسخ‌های صحیح دریافت شد. سؤال‌های هر سال شامل ۵۰ سؤال زیست‌شناسی است که در دفترچه اختصاصی نوشته شده بود؛ بنابراین ۱۵۰ سؤال مورد بررسی و تجزیه و تحلیل توسط سه ارزشیاب صورت گرفت. در دفترچه آزمون ورودی هر سال، سؤال‌های زیست‌شناسی از شماره ۱۵۶ شروع می‌شود و به سؤال ۲۰۵ ختم می‌یابد. برای تعیین میزان پاسخگویی داوطلبان، جامعه آماری، تمام داوطلبان شرکت‌کننده دختر و پسر در آزمون ورودی سال‌های ۹۵ به تعداد ۵۱۹۴۸۸، در سال ۹۶ به تعداد ۵۸۰۳۰۱ نفر و ۹۷ به تعداد ۶۴۲۲۸۸ نفر گزارش شد. از آنجائی که تمام سؤال‌های زیست‌شناسی هر سه سال مورد بررسی قرار گرفت، لذا نیاز به نمونه‌گیری نبود و از استراتژی تمام شماری استفاده شد. در مورد داوطلبان آزمون نیز به‌صورت تصادفی از داوطلبان پسر و دختر مدارس دولتی شهرستان‌های مختلف در سهمیه یک استفاده شد که طبق جدول مورگان به ازاء ۱۰۰۰۰۰ نفر به بالا در جامعه آماری ۳۸۴ نفر به‌عنوان نمونه آماری کافی بود؛ ولی به دلیل داشتن آمار بیشتر در هر سال، ۵۰۰ نفر در هر ورودی و هر جنس انتخاب شدند. سؤال‌های آزمون سراسری رشته علوم تجربی سه سال (در هر سال ۵۰ سؤال) به‌عنوان ابزار اندازه‌گیری است که برای انجام داده‌های آماری استفاده شد. این آزمون‌ها از نظر محتوای آموزشی یعنی چهار سال کتب درسی دوره متوسطه دوم همان سال‌ها تهیه شده است و روایی و پایایی آنها از طریق طراحان سازمان سنجش و آموزش کشور مورد تأیید بود. روش تحلیل سؤال‌ها از آمار توصیفی برای تعیین جداول، توزیع فراوانی، درصد و نمودار استفاده شد. نحوه عمل بدین صورت بود، هر کدام از ارزشیابان به بررسی سؤال‌های آزمون سراسری بر مبنای ویژگی‌های حیطة شناختی بلوم پرداخته و نظرات خود را در جداول مشابه ثبت کردند. لازم به ذکر است مشخصات سطوح مختلف حیطة

شناختی بلوم از کتاب سیف چاپ ۱۳۹۸ در اختیار همکاران قرار داده شد. سپس در هر آزمون، هر کدام از طبقات مشخص شده (فراوانی) و بعد از آن درصد فراوانی هر طبقه محاسبه گردید و در جدولی ثبت شد. بعد از جمع‌آوری و ثبت داده‌های ارزشیابان، به صورت دوجه دو طبقه آماره کاپای کوهن مقایسه گردید. همه موارد ذکر شده در نرم‌افزار SPSS^{۲۵} صورت گرفت. برای تحلیل میزان پاسخگویی داوطلبان از آمار توصیفی (فراوانی و درصد) در مورد تک‌تک سؤال‌ها پرداخته شد و کمترین و بیشترین درصد و میانگین کل پاسخ به سؤال‌های بیان گردید.

۲-۲. یافته‌های پژوهش

بعد از بررسی سؤال‌های آزمون زیست‌شناسی بر اساس سطوح مختلف حیطه شناختی و ثبت در جداول، تعیین درصد، فراوانی به تفکیک هر سه سال و مقایسه آنها، می‌بایست به سؤال‌های پژوهش پرداخت. سؤال‌ها و نتایج حاصل از آنها به قرار زیر است.

۲-۲-۱. سؤال ۱: سؤال‌های زیست‌شناسی آزمون سال‌های ۹۷-۱۳۹۵ بر اساس سطوح مختلف حیطه شناختی چگونه است.

الف) نتایج آماری فراوانی و درصد مربوط به سؤال‌های درس زیست‌شناسی

آزمون ورودی ۹۷-۱۳۹۵ توسط ارزشیابان

نتایج آماری سؤال‌های زیست‌شناسی سه سال به صورت خلاصه در جدول زیر به تفکیک سال مشخص شده است.

جدول ۱: فراوانی و درصد مربوط به سؤال‌های آزمون ورودی درس زیست‌شناسی بر حسب نظر

ارزشیابان

ارزشیاب ۳		ارزشیاب ۲		ارزشیاب ۱		۹۵
درصد	فراوانی	درصد	فراوانی	درصد	فراوانی	
۰	۰	۰	۰	۰	۰	دانش
۹۰	۴۵	۴۴	۲۲	۵۶	۲۸	فهمیدن
۸	۴	۰	۰	۸	۴	کاربرد
۰	۰	۳۲	۱۶	۳۰	۱۵	تحلیل
۰	۰	۲۲	۱۱	۴	۲	ترکیب
۰	۰	۰	۰	۰	۰	ارزشیابی
۱۰۰	۵۰	۱۰۰	۵۰	۱۰۰	۵۰	کل

ارزشیاب ۳		ارزشیاب ۲		ارزشیاب ۱		۹۶
درصد	فراوانی	درصد	فراوانی	درصد	فراوانی	
۶	۳	۶	۳	۱۰	۵	دانش
۸۸	۴۴	۳۴	۱۷	۵۲	۲۶	فهمیدن
۶	۳	۰	۰	۸	۴	کاربرد
۰	۰	۳۴	۱۷	۲۰	۱۰	تحلیل
۰	۰	۲۶	۱۳	۱۰	۵	ترکیب
۰	۰	۰	۰	۰	۰	ارزشیابی
۱۰۰	۵۰	۱۰۰	۵۰	۱۰۰	۵۰	کل
ارزشیاب ۳		ارزشیاب ۲		ارزشیاب ۱		۹۷
درصد	فراوانی	درصد	فراوانی	درصد	فراوانی	
۶	۳	۶	۳	۸	۴	دانش
۸۰	۴۰	۴۲	۲۱	۵۶	۲۸	فهمیدن
۸	۴	۰	۰	۲	۱	کاربرد
۶	۳	۳۶	۱۸	۲۶	۱۳	تحلیل
۰	۰	۱۶	۸	۸	۴	ترکیب
۰	۰	۰	۰	۰	۰	ارزشیابی
۱۰۰	۵۰	۱۰۰	۵۰	۱۰۰	۵۰	کل

نتایج جدول فوق نشان می‌دهد که درصد بالای سؤال‌های زیست‌شناسی، بیشتر در سطوح پائین‌تر مخصوصاً فهمیدن است و سطوح بالاتر یعنی ترکیب و ارزشیابی از درصد بسیار پائین و در برخی موارد صفر است.

برای مقایسه دو به دوی ارزشیابان و تعیین میزان توافق بین آنها در زمینه کدگذاری‌های انجام شده از شاخص کاپای کوهن استفاده می‌شود. شاخص کاپا تنها برای متغیرهایی مورد استفاده قرار می‌گیرد که هم سطح سنجش آنها یکی باشد و هم تعداد طبقات آنها با یکدیگر برابر باشد. مقدار شاخص کاپا که به کاپای کوهن معروف است، بین صفر تا یک نوسان دارد. در واقع، ضریب کاپا اندازه‌ای عددی بین -1 تا $+1$ است که هر چه به $+1$ نزدیک‌تر باشد بیانگر وجود توافق متناسب و مستقیم است. اندازه‌های نزدیک به -1 نشان‌دهنده وجود توافق وارون و اندازه‌های نزدیک به صفر عدم توافق را نشان

می‌دهد (حبیب‌پور، ۱۳۹۴)؛ لذا هرگاه ضریب کاپا کوچک‌تر از صفر باشد شدت توافق ضعیف، بین ۰/۰۱ تا ۰/۲۰ نسبتاً ضعیف، بین ۰/۲۱ تا ۰/۴۰ متوسط، بین ۰/۴۱ تا ۰/۶۰ نسبتاً زیاد، بین ۰/۶۱ تا ۰/۸۰ زیاد و بین ۰/۸۱ تا ۱۰۰ تقریباً کامل است.

ب) تحلیل یافته‌های سؤال‌های زیست‌شناسی آزمون سراسری ۱۳۹۵ رشته علوم تجربی توسط گروه ارزشیابان

نتایج حاصل از بررسی ۵۰ سؤال درس زیست‌شناسی در آزمون ورودی سال ۹۵ بر اساس سطوح مختلف حیطة شناختی بیان می‌دارد که: میزان توافق کاپا بین ارزشیاب ۱ و ارزشیاب ۲ معادل ۰/۴۸ است. این مقدار بیانگر وجود توافق نسبتاً زیاد بین دو سری اندازه‌گیری نظرات ارزشیاب ۱ و ارزشیاب ۲ خواهد بود. همچنین در $p\text{-value} < 0.001$ این اندازه‌گیری توافق، معنادار است. با وجود احتمال شانس و تصادفی بودن ۲۱ سؤال از سؤال‌های زیست‌شناسی که با کد فهمیدن از سطوح حیطة شناختی تشخیص‌داده‌شده است، اگر بخواهیم احتمالات شانسی و تصادفی برای کد فهمیدن را برداریم، بین ارزشیاب ۱ و ارزشیاب ۲ در ۱۳ سؤال درس زیست‌شناسی اتفاق نظر وجود دارد.

میزان توافق کاپا، میان ارزشیاب ۱ و ارزشیاب ۳ معادل ۰/۲۶ است. این مقدار بیانگر وجود توافق متوسط بین دو سری اندازه‌گیری نظرات ارزشیاب ۱ و ارزشیاب ۳ خواهد بود. همچنین در $p\text{-value} < 0.001$ این اندازه توافق، معنادار است. با وجود احتمال شانس و تصادفی بودن ۲۸ سؤال از سؤال‌های زیست‌شناسی که با کد فهمیدن از سطوح حیطة شناختی تشخیص‌داده‌شده است، اگر بخواهیم احتمالات شانسی و تصادفی برای کد فهمیدن را برداریم، میان ارزشیاب ۱ و ارزشیاب ۳ در ۲۶ سؤال درس زیست‌شناسی اتفاق نظر وجود دارد.

میزان توافق کاپا میان ارزشیاب ۲ و ارزشیاب ۳ معادل ۰/۰۶۲ است. این مقدار بیانگر وجود توافق نسبتاً ضعیف بین دو سری اندازه‌گیری نظرات ارزشیاب ۲ و ارزشیاب ۳ خواهد بود. همچنین در $p\text{-value} < 0.06$ این اندازه توافق، معنادار است. با وجود احتمال شانس و تصادفی بودن ۲۲ سؤال از سؤال‌های زیست‌شناسی که با کد فهمیدن از سطوح حیطة شناختی تشخیص‌داده‌شده است، اگر بخواهیم احتمالات شانسی و تصادفی برای کد فهمیدن را برداریم، میان ارزشیاب ۲ و ارزشیاب گروه ۳ در ۲۰ سؤال درس زیست‌شناسی اتفاق نظر وجود دارد.

پ) تحلیل یافته‌های سؤال‌های زیست‌شناسی آزمون سراسری ۱۳۹۶ رشته علوم تجربی توسط گروه ارزشیابان

میزان توافق کاپا بین ارزشیاب ۱ و ارزشیاب ۲ معادل $0/42$ است. این مقدار بیانگر وجود توافق نسبتاً زیاد بین دو سری اندازه‌گیری نظرات ارزشیاب ۱ و ارزشیاب ۲ خواهد بود. همچنین در $p\text{-value} < 0.001$ این اندازه توافق، معنادار است. با وجود احتمال شانس و تصادفی بودن ۱۴ سؤال از سؤال‌های زیست‌شناسی که با کد فهمیدن از سطوح حیطة شناختی تشخیص داده شده است، اگر بخواهیم احتمالات شانسی و تصادفی برای کد فهمیدن را برداریم، بین ارزشیاب ۱ و ارزشیاب ۲ در ۹ سؤال درس زیست‌شناسی اتفاق نظر وجود دارد.

میزان توافق کاپا بین ارزشیاب ۱ و ارزشیاب ۳ معادل $0/32$ است. این مقدار بیانگر وجود توافق متوسط بین دو سری اندازه‌گیری نظرات ارزشیاب ۱ و ارزشیاب ۳ خواهد بود. همچنین در $p\text{-value} < 0.001$ این اندازه توافق، معنادار است. با وجود احتمال شانس و تصادفی بودن ۲۶ سؤال از سؤال‌های زیست‌شناسی که با کد فهمیدن از سطوح حیطة شناختی تشخیص داده شده است، اگر بخواهیم احتمالات شانسی و تصادفی برای کد فهمیدن را برداریم، بین ارزشیاب ۱ و ارزشیاب ۳ در ۲۳ سؤال درس زیست‌شناسی اتفاق نظر وجود دارد.

میزان توافق کاپا بین ارزشیاب ۲ و ارزشیاب ۳ معادل $0/14$ است. این مقدار بیانگر وجود توافق نسبتاً ضعیف بین دو سری اندازه‌گیری نظرات ارزشیاب ۲ و ارزشیاب ۳ خواهد بود. همچنین در $p\text{-value} < 0.001$ این اندازه توافق، معنادار است. با وجود احتمال شانس و تصادفی بودن ۱۷ سؤال از سؤال‌های زیست‌شناسی که با کد فهمیدن از سطوح حیطة شناختی تشخیص داده شده است، اگر بخواهیم احتمالات شانسی و تصادفی برای کد فهمیدن را برداریم، بین ارزشیاب ۲ و ارزشیاب ۳ در ۱۵ سؤال درس زیست‌شناسی اتفاق نظر وجود دارد.

ت) تحلیل یافته‌های سؤال‌های زیست‌شناسی آزمون سراسری ۱۳۹۷ رشته علوم تجربی توسط گروه ارزشیابان

میزان توافق کاپا بین ارزشیاب ۱ و ۲ معادل $0/54$ است. این مقدار بیانگر وجود توافق نسبتاً زیاد بین دو سری اندازه‌گیری نظرات ارزشیاب ۱ و ۲ خواهد بود. همچنین در $p\text{-}$

0.001 < value این اندازه توافق، معنادار است. با وجود احتمال شانس و تصادفی بودن ۱۹ سؤال از سؤال‌های زیست‌شناسی که با کد فهمیدن از سطوح حیطة شناختی تشخیص داده شده است، اگر بخواهیم احتمالات شانسی و تصادفی برای کد فهمیدن را برداریم، بین ارزشیاب ۱ و ۱۲ در ۱۲ سؤال درس زیست‌شناسی اتفاق نظر وجود دارد.

میزان توافق کاپا بین ارزشیاب ۱ و ۳ معادل ۰/۴۳ است. این مقدار بیانگر وجود توافق نسبتاً زیاد بین دو سری اندازه‌گیری نظرات بین ارزشیاب ۱ و ۳ خواهد بود. همچنین در $p\text{-value} < 0.001$ این اندازه توافق، معنادار است. با وجود احتمال شانس و تصادفی بودن ۲۸ سؤال از سؤال‌های زیست‌شناسی که با کد فهمیدن از سطوح حیطة شناختی تشخیص داده شده است، اگر بخواهیم احتمالات شانسی و تصادفی برای کد فهمیدن را برداریم، بین ارزشیاب ۱ و ۳ در ۲۲ سؤال درس زیست‌شناسی اتفاق نظر وجود دارد.

میزان توافق کاپا بین ارزشیاب ۲ و ۳ معادل ۰/۱۵ است. این مقدار بیانگر وجود توافق نسبتاً ضعیف بین دو سری اندازه‌گیری نظرات بین ارزشیاب ۲ و ۳ خواهد بود. همچنین در $p\text{-value} < 0/009$ این اندازه توافق، معنادار است. با وجود احتمال شانس و تصادفی بودن ۱۹ سؤال از سؤال‌های زیست‌شناسی که با کد فهمیدن از سطوح حیطة شناختی تشخیص داده شده است، اگر بخواهیم احتمالات شانسی و تصادفی برای کد فهمیدن را برداریم، بین ارزشیاب ۲ و ۳ در ۱۷ سؤال درس زیست‌شناسی اتفاق نظر وجود دارد.

سؤال ۲: میزان عملکرد داوطلبان به پاسخگویی سؤال‌های درس زیست‌شناسی

آزمون ورودی رشته علوم تجربی در سال‌های ۹۷-۱۳۹۵ چگونه است؟
 نتایج آمار توصیفی وضعیت پاسخگویی به سؤال‌های زیست‌شناسی سال ۹۵ نشان می‌دهد که میانگین پاسخ‌های صحیح برای تمام سؤال‌های کمتر از حد متوسط (۵۰ درصد) گزارش شد. کمترین میانگین مربوط به سؤال ۱۹۸ با میانگین ۲/۳ و بیشترین میانگین پاسخ‌های صحیح مربوط به سؤال ۱۸۵ با میانگین ۳۶ بود. میانگین کل پاسخ‌های صحیح هم کمتر از ۱۰ درصد و معادل ۹/۵۲ گزارش شد.

نتایج آمار توصیفی وضعیت پاسخگویی به سؤال‌های زیست‌شناسی سال ۹۶ بیان می‌دارد که میانگین پاسخ‌های صحیح برای تمام سؤال‌های کمتر از حد متوسط (۵۰ درصد) است. کمترین میانگین مربوط به سؤال ۱۷۳ با میانگین ۴/۵ و بیشترین میانگین پاسخ‌های صحیح مربوط به سؤال ۱۷۷ با میانگین ۳۴/۵ بود. میانگین کل پاسخ‌های صحیح هم ۱۷/۳۶ گزارش شد.

نتایج آمار توصیفی وضعیت پاسخگویی به سؤال‌های زیست‌شناسی سال ۹۷ مشخص می‌کند که میانگین پاسخ‌های صحیح برای تمام سؤال‌های کمتر از حد متوسط (۵۰ درصد) گزارش شد. کمترین میانگین مربوط به سؤال ۱۹۱ با میانگین ۴/۳ و بیشترین میانگین پاسخ‌های صحیح مربوط به سؤال ۱۶۸ با میانگین ۶۰/۸ بود. میانگین کل پاسخ‌های صحیح هم ۲۱/۸۶ گزارش شد.

۳. نتیجه گیری

ارزیابی سؤال‌های درس زیست‌شناسی آزمون سراسری رشته علوم تجربی سال‌های ۹۵ تا ۹۷ نشان می‌دهد که سؤال‌های زیست‌شناسی در سطح پائین حیطه شناختی بلوم یعنی فهمیدن است و میزان پاسخگویی داوطلبان در حد نامطلوبی است (کمتر از ۵۰ درصد). یکی از رویکردها، استفاده از طبقه‌بندی دامنه‌های شناختی بلوم است. در این پژوهش نیز سعی شده است، یکی از مهم‌ترین آزمون‌های سرنوشت‌ساز دانش‌آموزان یعنی آزمون سراسری مورد بررسی قرار گیرد. زیرا در سال‌های اخیر توجه بسیار زیادی به آزمون‌های چهارگزینه‌ای در دوره‌های متوسطه اول و دوم برای آمادگی دانش‌آموزان به آزمون سراسری سازمان سنجش و اندازه‌گیری شده است. طبقه‌بندی بلوم به معنای مطلق یا قطعی نیست؛ ولی به‌عنوان یک راهنمای کلی برای کمک به معلمان و دانش‌آموزان در توسعه و شناسایی سؤال‌های مربوط به زیست‌شناسی در سطوح مختلف بلوم است. همچنین طبقه‌بندی بلوم می‌تواند به‌منظور کمک به شناسایی سؤال‌های امتحان توسط دانش‌آموزان مورداستفاده قرار گیرد که به‌عنوان بزرگ‌ترین چالش علمی مطرح است. باین‌حال، هنگامی که سؤال‌های چالش‌برانگیز بیان می‌شود، دانش‌آموزان نیاز به راهنمایی دارند تا چگونگی عادت مطالعه خود را تغییر و بهتر کنند و خود را برای پاسخ‌گویی به این نوع سؤال‌ها آماده کنند؛ لذا نیاز به طراحی یک سری فعالیت‌ها و سؤال‌های آزمون در سطوح مختلف طبقه‌بندی بلوم است (کرو، ۲۰۰۸، سیوارامن^۱ و کریشنا^۲، ۲۰۱۵: ۱/۶-۴). در سال‌های اخیر در ایران، تغییراتی در نحوه مطالعه دانش‌آموزان دوره دوم متوسطه صورت گرفته است که ناشی از امتحانات آزمون سراسری است؛ بنابراین شایسته است

سؤال‌های طوری انتخاب شوند که دانش‌آموزان به مباحث حفظی و طوطی‌وار هدایت نشوند و درک عمیقی نسبت به محتوای درسی داشته باشند.

به‌منظور پاسخ به سؤال ۱ پژوهش حاضر، نتایج حاصل از بررسی سؤال‌های درس زیست‌شناسی سه سال آزمون سراسری به‌وسیله دبیران زیست‌شناسی و اساتید دانشگاه فرهنگیان بر مبنای سطوح مختلف حیطه شناختی نشان داد که سؤال‌های زیست‌شناسی سال‌های ۹۵ تا ۹۷ دارای درصد بالایی در سطح فهمیدن است و درصد سطوح بالای طبقه‌بندی حیطه شناختی بلوم از درصد بسیار پایینی برخوردار است. هر چند در این سؤال‌ها از اشکال و نمودارها استفاده شده است؛ ولی سؤال‌هایی که منطبق بر شکل، نمودار و غیره باشند نمی‌توانند در سطوح بالاتر از کاربرد قرار گیرند (برادی^۱، ۲۰۰۵). در حالیکه می‌توان با تعمق بیشتر به سؤال‌های تجزیه و تحلیل و حتی سطح ارزشیابی نیز پرداخت.

ارزشیابی، دانش‌آموزان را از انتظارات دوره مطلع می‌کند که نشان‌دهنده پیامدهای مهمی برای نتایج یادگیری دانش‌آموزان است. اما طراحی سؤال‌های امتحانی که نیازمند مهارت‌های تفکر بالاتر است مطمئناً یک کار چالش‌برانگیز برای یک مربی، به‌خصوص با استفاده از فرمت چند گزینه‌ای است. با این حال، ارزیابی در سطوح بالاتر حیطه شناختی می‌تواند عامل اصلی در تحریک دانش‌آموزان و به‌طور مؤثر درک عمیق از مواد درسی باشد، درکی که نه‌تنها کاربرد، تجزیه و تحلیل و ارزشیابی را پشتیبانی می‌کند، بلکه همچنین بهتر است حفظ کردن حقایق اصلی را نیز باعث می‌گردد. در مقابل، اتخاذ یک رویکرد معمولی برای امتحانات (و شاید آسان‌تر برای مربی) یعنی یادآوری واقعی مواد درسی به دانش‌آموزان لطمه وارد می‌کند. از آنجائی که این نوع امتحانات به‌احتمال کمتری تفکر انتقادی را پرورش می‌دهد و دانش را می‌سجد، به نظر نمی‌رسد حتی افزایش یادگیری و حفظ اطلاعات واقعی در امتحانات عالی‌تر در طول دوره به میزان لازم تحریک برانگیز باشد (جنسن^۲، مک دانیل^۳، ودارد^۴، کومر^۵، ۲۰۱۴: ۲۶/۳۰۷-۳۲۹).

-
1. Brady
 2. Jensen
 3. Mc Danile
 4. Woodard
 5. Kummer

این پژوهش مانند پژوهش‌های میرزایی تبار (۱۳۹۴)، سلیمان‌زاده (۱۳۹۳)، آبکار (۱۳۹۱) و لورد (۲۰۰۷) همسو است و با پژوهش زارعی و اوجی نژاد، صارمی (۱۳۹۵) ناهم‌سوست؛ بنابراین می‌توان گفت، طراحان سؤال بیشتر به سطوح پایین مخصوصاً سطح فهمیدن پرداخته‌اند و توجه بسیار کمی در سطوح بالاتر نظیر تجزیه و تحلیل، ترکیب و ارزشیابی معطوف کرده‌اند. احتمالاً تخصص و تجربه ناکافی طراحان سؤال در باره انواع ارزشیابی و سطوح مختلف حیطه شناختی و یا نبودن شرایط و امکانات لازم از دلایل مهم آن است که در این زمینه نیاز به پژوهش بیشتری است. در طی پژوهش آ نیز^۱ (۲۰۱۷) در دانشکده کشاورزی پاکستان، اساتید از چندین روش تدریس برای آموزش دانشجویان استفاده کرده بودند؛ اما امتحانات محدود به سطح پایین یادگیری بود. اکثر معلمان هیچ‌گونه اطلاعاتی در مورد طبقه‌بندی بلوم نداشتند. نتایج حاصل از تجزیه و تحلیل یافته‌ها بیان داشت که بیشتر سؤالات در سطح پایین (دانش، درک و کاربرد) بودند. فقط چند سؤال در سطح بالاتر (تجزیه و تحلیل، ترکیب و ارزیابی) قرار داشتند. آزمون‌ها باید به‌وسیله معلمان و در سطوح بالاتر طبقه‌بندی بلوم طراحی شود تا واقعاً مهارت‌های فرایند علمی را آزمایش کنند. شواهد حاصل از دستیابی به نتایج یادگیری مطلوب نشان می‌دهد که تنها این فرم ارزیابی مناسب خواهد بود. مطالعات متعددی بیان می‌کنند که در واقع هدایت یادگیری دانش‌آموزان باعث تمرکز تلاش‌های مطالعه دانش‌آموزان بر روی این مهارت‌های موردنظر و در نهایت منجر به فهم عمیق مفهومی می‌شود. ارزیابی مؤثر می‌تواند صرف زمان لازم و مؤثر خارج از کلاس را به‌جای وقایع حفظ حافظه در دانش‌آموزان تغییر دهد (جنسن، ۲۰۱۴: ۳۰۷-۳۲۹).

علاوه بر این، بر مبنای پژوهش برخی از پژوهشگران، سؤال‌های چند گزینه‌تفکر و سطوح بالاتر را ارزیابی نمی‌کند و احتمال حدس بالاتر را فراهم می‌کند که توانمندی تست را برای دانش‌آموزان با توانمندی کم کاهش می‌دهد (کرانباخ^۲، ۱۹۸۸: ۳-۱۷). همچنین، سؤال‌های چندگزینه‌ای در فرایند سطوح شناختی بالاتر شکست می‌خورد و آزمون‌دهنده به حدس زدن ترغیب می‌شود (کمپبل^۳، ۲۰۰۰) که استعداد‌های مربوط به

1. Anees
2. Cronbach
3. Campbell

فهم دانش‌آموزان یا تسلط بر محتوای درسی را به گمراهی می‌کشاند (چان^۱، کندی^۲، ۲۰۰۲، ۹۵۷-۹۷۱). اگر چه سؤال‌های چند گزینه‌ی می‌تواند برای اندازه‌گیری مهارت‌های استدلال طراحی شده باشند، ولی تعدادی معتقدند که نمی‌توانند فرایندهای شناختی به طور مؤثر و مهم به‌مانند سؤال‌های باز را فراهم کنند که برای توانائی‌های سطوح بالاتر مانند تجزیه و تحلیل، ترکیب و ارزشیابی حائز اهمیت هستند (استیگینس^۳، ۲۰۰۵). احتمالاً یکی از دلایل دیگر در سؤال‌های آزمون ورودی رشته علوم تجربی نیز همین باشد که سؤال‌های چند گزینه‌ی برای ارزیابی سطوح شناختی بالاتر داوطلبان آزمون مناسب نمی‌باشد و شایسته است از شیوه دیگری برای ارزیابی داوطلبان استفاده کرد.

بر اساس این پژوهش و مطالعات انجام شده در رشته‌های مختلف ذکر شده در بالا، برای ارزیابی توانائی‌های شناختی بالاتر از جمله ترکیب و ارزشیابی نیاز به ارزیابی از نوع دیگر است که امیدواریم در سال‌های آتی تغییراتی در نحوه و نوع ارزیابی دانش‌آموزان در آزمون سراسری صورت گیرد که با اهداف آموزشی دو سند برنامه درسی ملی و سند تحول بنیادین همخوانی داشته باشد. در واقع این اسناد، سمت‌وسوی همه برنامه‌های آموزشی از جمله علوم تجربی از پایه اول تا دوازدهم را مشخص کرده است تا دانش‌آموزان شایستگی لازم در جهت دانش‌اندوزی را کسب کنند.

نتایج حاصل از داده‌ها در باره سؤال ۲ پژوهش بیان می‌دارد که میزان پاسخگویی داوطلبان به سؤال‌های آزمون زیست‌شناسی در سال ۱۳۹۵ نامطلوب است (کمتر از ۵۰ درصد). این در حالی است که میانگین سؤال‌های آزمون در حیطه شناختی بلوم در سطح فهمیدن است. با وجود اینکه میزان پاسخگویی در حد نامطلوب (کمتر از ۵۰ درصد) بوده است؛ ولی از سال ۹۵ تا ۹۷ میزان پاسخگویی افزایش یافته است. البته لازم به ذکر است در این پژوهش فقط نمره آزمون زیست‌شناسی بررسی شده است که با ضریب بالا در پذیرش داوطلب دخالت دارد و عوامل دیگر از جمله دروس دیگر، حالت‌های روحی و روانی داوطلب در هنگام آزمون، شرایط استفاده از کلاس‌های آزمون و امکانات آموزشی دیگر بعد از دیپلم، سهمیه‌بندی، انتخاب رشته و عوامل دیگر به‌نوعی در پذیرش داوطلب در دانشگاه‌ها مداخله می‌کنند.

1. Chan
2. Kenedy
3. Stiggins

در رشته‌های مختلف مانند هنر، علاوه بر آزمون کتبی، آزمون عملی نیز برگزار می‌شود. از آن جایی که رشته علوم تجربی شامل رشته‌های بسیار مهمی در زندگی انسان‌ها از جمله پزشکی، دندانپزشکی، آزمایشگاهی، داروسازی و دیگر رشته‌های مورد توجه در جامعه و مرتبط با تعلیم و تربیت مانند شغل معلمی است، لذا شایسته است افراد منتخب در آزمون سراسری به نحو مطلوب و در راستای اهداف برگزیده شوند و علاوه بر آزمون کتبی از آزمون‌های عملی و مهارتی دیگر استفاده گردد که منطبق بر اهداف آموزشی برنامه درسی ملی و سند تحول بنیادین است. همچنین هر دانشگاه برتر بتواند شاخص‌های مختلف برای گزینش دانشجو بر اساس ظرفیت، معدل، سوابق تحصیلی و نیاز جامعه در گزینش دانشجو تعریف کند تا افراد شایسته در جایگاه مناسب دانشگاه به تحصیل مشغول شوند و در کسب شغل مفید و مورد علاقه خود در جامعه مبادرت ورزند. علاوه بر این انتظار می‌رود به منظور طراحی سؤالات به نقش راهبردهای یادگیری - یاددهی دانش‌آموزان در پاسخ به سؤال‌های آزمون ورودی دانشگاه‌ها، میزان آگاهی طراحان سؤال آزمون ورودی دانشگاه‌ها در باره اهداف برنامه درسی ملی توجه لازم شود و به منظور بهبود سؤالات در هر سال به بررسی سؤال‌های آزمون از روش‌های مختلف از جمله نظریه سؤال - پاسخ پرداخته شود.

- Allen, D., & Tanner, K. (2005). Infusing active learning into the large-enrollment biology class: seven strategies, from the simple to complex. **Cell biology education**, 4(4), 262-268.
- Anees, S. (2017). Analysis of Assessment Levels of Students' Learning According to Cognitive Domain of Bloom's Taxonomy. Online Submission.
- Bissell, A. N., & Lemons, P. P. (2006). A new method for assessing critical thinking in the classroom. **BioScience**, 56(1), 66-72.
- Brady, A. M. (2005). Assessment of learning with multiple-choice questions. **Nurse Education in Practice**, 5(4), 238-242.
- Campbell, J. R. (2000). **Cognitive processes elicited by multiple-choice and constructed-response questions on an assessment of reading comprehension.**
- Chan, N., & Kennedy, P. E. (2002). Are multiple-choice exams easier for economics students? A comparison of multiple-choice and "equivalent" constructed-response exam questions. **Southern Economic Journal**, 957-971.
- Cronbach, L. J. (1988). Five perspectives on validity argument. **Test validity**, 3-17.
- Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in bloom: implementing Bloom's taxonomy to enhance student learning in biology. **CBE—Life Sciences Education**, 7(4), 368-381.
- Davis, B. G. (1993). Tools for Teaching (First).
- Hirsch, P. J. (2016). High Stakes Testing and its Effect on Teacher Methodologies (Doctoral dissertation, Caldwell College).
- Huang, X. H. & Lee, J. C. K. (2015). Mapping the Interaction between High-stake Testing and Curriculum Reform: A Case Study in Mainland China. **Curriculum and Teaching**, 30 (2), 81- 98.
- Hyder, I., & Bhamani, S. (2016). Bloom's taxonomy (cognitive domain) in higher education settings: Reflection brief. **Journal of Education and Educational Development**, 3(2), 288-300.
- Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the test... or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding. **Educational Psychology Review**, 26(2), 307-329.
- Lord, T., & Baviskar, S. (2007). Moving students from information recitation to information understanding-Exploiting Bloom's Taxonomy in creating science questions. **Journal of College Science Teaching**, 36(5), 40.

- Manthey, G. (2006). More than Just the Facts. **Leadership**, 36(2), 12-13.
- Morrison, Gary R., Ross, Steven M. and Kemp, Jerrold E. (2004). **Designing Effective Instruction**. 4th ed. US: Wiley.
- Udovic, D., Morris, D., Dickman, A., Postlethwait, J., & Wetherwax, P. (1996). **The Workshop Biology Curriculum Handbook**.
- Paul, D. V., Naik, S. B., Rane, P., & Pawar, J. D. (2012, July). Use of an evolutionary approach for question paper template generation. **In 2012 IEEE Fourth International Conference on Technology for Education** 144-148. IEEE.
- Sivaraman, S. I., & Krishna, D. (2015). Blooms taxonomy–application in exam papers assessment. **International Journal OF Multidisciplinary Sciences AND Engineering**, 6(9), 1-4.
- Stiggins, R. J. (2005). **Student-involved assessment for learning**. Prentice Hall.
- Yanchinda, J., Yodmongkol, P., & Chakpitak, N. (2016). Measurement of Learning Process by Semantic Annotation Technique on Bloom's Taxonomy Vocabulary. **International Education Studies**, 9(1), 107-122.