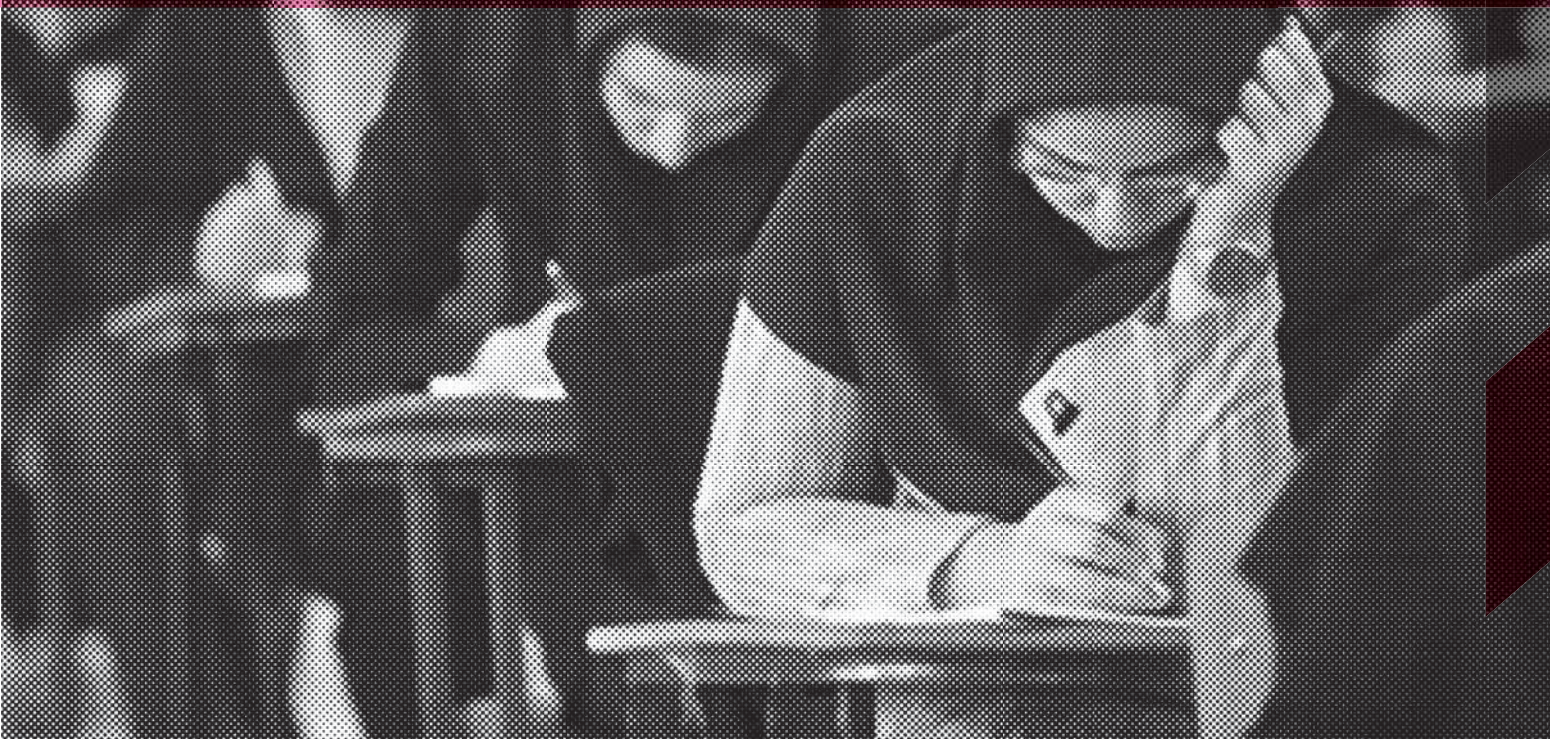


پاکت  
ویژه

# تحلیل نمرات خام پذیرفته‌شدگان آزمون ورودی دانشگاه‌های ایران در درس عمومی، اختصاصی و زبان انگلیسی

(مورد مطالعه: آزمون سراسری سال ۱۳۹۴)

رضایروی  
دانشجوی دکتری سنجش و اندازه‌گیری دانشگاه علامه طباطبائی  
Reza.Payravi@yahoo.com  
دکتر محمد رضا فلسفی نژاد  
دانشیار گروه سنجش و اندازه‌گیری دانشگاه علامه طباطبائی  
falsafinejad@yahoo.co.uk



## چکیده

پژوهش پیش رو به منظور تحلیل نمره‌های خام پذیرفته‌شدگان آزمون سراسری سال ۱۳۹۴ با هدف ارزشیابی کمی این آزمون انجام پذیرفت. هدف غایی این پژوهش، شناسایی کاستی‌های احتمالی این آزمون و پیشنهاد راه‌کارهایی جهت افزایش دقت و کاهش خطای اندازه‌گیری است. در این پژوهش، جامعه آماری و همچنین نمونه آماری شامل تمامی داوطلبان پذیرفته‌شده آزمون سراسری سال ۱۳۹۴، که در پنج گروه آزمایشی و از طریق شرکت در آزمون، پذیرش و به دانشگاه معرفی شده‌اند و حجم آن برابر با ۲۷۸۶۸۶ نفر می‌باشد. نتایج و یافته‌های این پژوهش نشان می‌دهد، میانگین نمرات پذیرفته‌شدگان در آزمون عمومی برابر با ۲۴/۵۰ درصد، متوسط دشواری سؤال‌ها حدود ۲۵٪ و متوسط واریانس سؤال‌ها مساوی با ۱۸٪ است. در آزمون اختصاصی نیز، میانگین نمرات این افراد برابر با ۱۴/۳۵ درصد، متوسط دشواری سؤال‌ها حدود ۱۴٪ و متوسط واریانس سؤال‌ها مساوی با ۱۲٪ است. به بیان دیگر، پذیرفته‌شدگان آزمون سراسری از هر ۴ سؤال عمومی به یک سؤال و از هر ۷ سؤال اختصاصی به یک سؤال پاسخ درست داده‌اند. در این میان، تحلیل‌ها اشاره دارد که این مقادیر برای آزمون ورودی دانشگاه که یک آزمون کشوری هنجاری با نرخ پذیرش ۳۰ درصد محسوب می‌شود، دارای یک درجه استاندارد اطمینان‌بخش، به این منظور که گفته شود «آزمون حداکثر قدرت تشخیص در تمام سطوح عملکرد آزمودنی‌ها را دارا می‌باشد»، نیست. همچنین تحلیل داده‌ها نشان داد که میانگین برخی خرده‌آزمون‌ها از قبیل میانگین درس ریاضی در گروه آزمایشی علوم انسانی با مقدار حدود ۲ درصد، بسیار پایین و نزدیک به صفر است. در این خرده‌آزمون‌ها، آزمودنی‌های بسیار قوی تا آزمودنی‌های بسیار ضعیف، نمره‌ای کم و بیش یکسان کسب کرده‌اند. لذا، نمرات این دروس، واریانس معناداری در نمره کل آزمون داوطلبان ایجاد نمی‌کند. بنابراین، انتظار می‌رود با اصلاح ساختار آزمون و بهبود پارامترهای فنی و روان‌سنجی سؤال‌ها، بتوان رتبه‌بندی داوطلبان را به نفع افراد شایسته‌تر تغییر داد.

**کلیدواژه‌ها:** آزمون ملی پیشرفت تحصیلی، افت تحصیلی، عملکرد تحصیلی، آزمون سراسری

## Abstract

This paper is aimed at analyzing the raw scores of university applicants admitted based on the university entrance exam in 2015. The aim of this research was to identify weaknesses of the test. The research method was of the applied, descriptive type and the statistical sample included all the applicants admitted to university programs (Total Numbers: 278686). To analyze data, the statistical measures such as mean, item difficulty, and item variance were applied.

The results of this research showed that the mean, average level of difficulty and variance of the items of the general test were 0.24.5%, 0.25 and 0.18 respectively. The same measures of the items in the discipline -specific test were 14.35% and 0.12 respectively. According to the Classical Test Theory (CTT), the acceptable levels of these measures are as follows:

- 1) The average level of difficulties for items, is equal to 0.50
- 2) Items difficulties should be in the range of 0.30 to 0.70
- 3) Items variances should be in the range of 0.21 to 0.25

The gap between the results and the acceptable levels suggests that the discrimination power of the tests was far less than optimal. The paper makes some suggestion for change to improve the quality of the tests.

**Key Words:** achievement test, academic failure, performance, the Iranian university entrance exam

## مقدمه

علمی داوطلبان آزمون. آزمون سرنوشت‌ساز آزمون است که در مقیاس گسترده اجرا شده و بر اساس نتایج آن تصمیماتی خطیر درباره افراد اتخاذ می‌گردد به طوری که این تصمیمات دارای تبعات مهم شخصی، اجتماعی و سیاسی برای افراد هستند (ریت<sup>۵</sup>، ۲۰۱۶؛ پاپهام<sup>۶</sup>، ۲۰۰۵). بدیهی است، چنانچه اندازه‌گیری با دقت و ثبات<sup>۷</sup> کافی انجام نپذیرد، دست کم، بخشی از پذیرفته‌شدگان آزمون سراسری وارد آن دسته از رشته‌هایی خواهند شد که یا انگیزه و علاقه کافی برای تلاش ندارند و یا توانایی پیشرفت تحصیلی مناسب در آن رشته‌ها را نخواهند داشت. به این ترتیب، نخستین خسارت ناشی از انگیزه پایین و یا کم توانی در تحصیل عبارت است از افت تحصیلی، ترک تحصیل و یا اخراج دانشجویان از دانشگاه (پروند، ۱۳۸۱، جانزاده، ۱۳۸۹).

آزمون سراسری ورود به دانشگاه‌ها و مؤسسات آموزش عالی کشور به منظور فراهم ساختن بستر رقابت عادلانه<sup>۱</sup> و منصفانه داوطلبان ورود به دانشگاه و گزینش افراد مستعد و توزیع آنان در رشته‌های مختلف، همه ساله از سوی سازمان سنجش آموزش کشور در پنج گروه آزمایشی علوم ریاضی، علوم تجربی، علوم انسانی، هنر و گروه زبان‌های خارجی برگزار می‌گردد.

در آزمون سراسری سال ۱۳۹۴ کشور، از میان ۱۰۷۲۸۱۳ متقاضی که ۹۱۸۳۸۷ نفر از آنان در جلسه آزمون حاضر بوده‌اند، تعداد ۲۷۸۶۸۶ نفر (قریب ۳۰ درصد) از طریق آزمون پذیرفته شده و به دانشگاه راه یافتند. جدول شماره (۱) آمار و وضعیت کلی داوطلبان آزمون سراسری سال ۱۳۹۴ را نشان می‌دهد.

جدول شماره (۱): آمار داوطلبان آزمون سراسری سال ۱۳۹۴

تعداد ثبت‌نام‌کننده	حاضر در جلسه آزمون	پذیرفته شده با آزمون	پذیرفته شده با سوابق تحصیلی	مجموع پذیرفته شده‌ها
نفر ۱۰۷۲۸۱۳	نفر ۹۱۸۳۸۷	نفر ۲۷۸۶۸۶	نفر ۱۴۹۷۹۲	نفر ۴۲۸۴۷۸

منبع: کارنامه آماری آزمون سراسری سال ۱۳۹۴

به اعتقاد متخصصان اندازه‌گیری، قوی‌ترین اتهام وارده به آزمون و سنجش و به طریق اولی، مهم‌ترین مسئله برای آزمودنی و جامعه، سنجش عادلانه و منصفانه است (همبلتون<sup>۸</sup> و همکاران، ۱۹۹۱). بی‌تردید، عادلانه‌بودن برای آزمون‌های سرنوشت‌ساز به عنوان آزمون‌هایی که به منظور تصمیم‌گیری برای آینده آزمودنی‌ها به کار می‌روند از اهمیت بیشتری برخوردار است (پاپهام، ۲۰۰۵). به قول استوکینگ<sup>۹</sup> و لوئیس<sup>۱۰</sup> (۲۰۰۰)، مسائل مرتبط با طرح سؤال، ساخت آزمون و به طور کلی سنجش، مشابه بادکنکی است که وقتی یک طرف آن را می‌فشاریم، سمت دیگر آن متورم می‌شود. یعنی، هنگامی که به یک جنبه آزمون و سنجش توجه می‌شود از سایر جنبه‌های آن غفلت می‌شود.

همان‌طور که آشکار است، هدف کلی آزمون‌های هنجار-مرجع<sup>۱۱</sup> همانند آزمون سراسری، شامل توصیف تفاوت‌های فردی، جداسازی آزمودنی‌ها از یکدیگر و رتبه‌بندی منصفانه و عادلانه آن‌هاست. در واقع، مهم‌ترین ویژگی این آزمون‌ها این است که میان آزمودنی‌ها تفکیک ایجاد کند و تفاوت‌های آن‌ها را در حوزه مورد اندازه‌گیری، تا حد ممکن به طور دقیق و کامل نشان دهد. واضح است، چنانچه شرکت‌کنندگان در یک آزمون، نمره صفر، یا نمره کامل و یا نمره برابر و یکسانی کسب کرده باشند، یعنی واریانس نمرات آنها صفر و یا نزدیک به صفر باشد، هیچ هدف سودمندی برای آزمون‌ساز نخواهد داشت (فرگوسن<sup>۱۲</sup> و تاکانه<sup>۱۳</sup>، ۱۹۸۹). لذا، آزمون‌ساز در فرآیند ساخت آزمون تلاش می‌کند سؤالاتی را در آزمون بگنجانند که واریانس نسبتاً بالایی داشته

سؤالات داوطلبان آزمون سراسری به دو بخش «سؤالات عمومی و سؤالات اختصاصی» تقسیم می‌شود. در بخش عمومی، تمامی داوطلبان می‌بایست به ۱۰۰ سؤال (شامل سؤالات دروس زبان و ادبیات فارسی، زبان عربی، فرهنگ و معارف اسلامی و زبان خارجی) در مدت ۷۵ دقیقه پاسخ دهند و در بخش آزمون اختصاصی، داوطلبان هر گروه آزمایشی، به سؤال‌های دروس مرتبط با رشته تحصیلی خود در وقت مقرر پاسخ می‌دهند (به جدول شماره ۴ مراجعه کنید). هدف از اجرای این پژوهش در حوزه نمره آزمون، توصیف و تحلیل نمره‌های خام درصدی پذیرفته‌شدگان آزمون سراسری سال ۱۳۹۴، ارزشیابی کمی این آزمون با استفاده از داده‌های واقعی و مقایسه وضعیت موجود با وضعیت مطلوب طراحی و اجرا گردید. هدف غایی این پژوهش نیز، شناسایی کاستی‌های احتمالی آزمون سراسری و پیشنهاد راهکارهایی جهت افزایش دقت<sup>۱۴</sup> و کاهش خطای اندازه‌گیری<sup>۱۵</sup> است. همچنین، گروه‌های مخاطب این مطالعه از یک طرف، داوطلبان آزمون سراسری و از طرف دیگر، برنامه‌ریزان آموزش و پرورش، معلمان، طراحان سؤال و سازمان سنجش آموزش کشور هستند تا بدین وسیله میزان و مسیر تلاش خود را هر چه بهتر و بیشتر تقویت و اصلاح نمایند.

## بیان مسئله

بی‌گمان، نقش آزمون سراسری به عنوان یک آزمون سرنوشت‌ساز<sup>۴</sup>، عبارت است از اندازه‌گیری دقیق و رتبه‌بندی

باشند. به این ترتیب، واریانس کل آزمون که از واریانس تک تک سوالات و همبستگی میان آن‌ها تشکیل می‌شود، بالا خواهد بود (مگنسون<sup>۱۴</sup>، ۱۹۶۷).

چنانچه یک آزمون، واریانس زیادی تولید کند، در واقع، چنین آزمونی توانسته تغییرات بیشتری را نشان دهد و در این صورت، بهتر می‌تواند افراد را از یکدیگر تفکیک کند. آشکار است که ایجاد واریانس مطلوب در یک آزمون، وابسته به انتخاب سوالات مناسب و توانمند است. مطابق نظریه کلاسیک آزمون<sup>۱۵</sup>، سؤال مناسب چندین قاعده و ضابطه دارد که از جمله مهم‌ترین آن‌ها این است که سطح ضریب دشواری<sup>۱۶</sup> سؤال حدود ۵۰٪ باشد (گالیکسن<sup>۱۷</sup>، ۱۹۸۷). زمانی که دشواری یک سؤال برابر با ۵۰٪ باشد، واریانس سؤال به بیشینه مقدار ممکن که برابر با ۲۵٪ است، می‌رسد (کروگر<sup>۱۸</sup> و آلچینا<sup>۱۹</sup>، ۲۰۰۸). به این ترتیب، آزمون‌سازان کار کشته پیشنهاد می‌کنند فقط سؤال‌هایی در آزمون قرار گیرند که «متوسط دشواری» آن‌ها فاصله و انحراف زیادی از ۵۰٪ نداشته باشند. بنابراین، ضرورت دارد و انتظار می‌رود با اصلاح ساختار آزمون سراسری و بهبود پارامترهای فنی و روان‌سنجی سوالات، بتوان رتبه‌بندی<sup>۲۰</sup> داوطلبان را به نفع افراد شایسته‌تر تغییر داد. بنا به گفته اوسترلیند<sup>۲۱</sup> (۲۰۰۲)، یکی از مهم‌ترین جنبه‌های ارزیابی آزمون، همواره، توصیف و تحلیل نمرات آزمودنی‌ها است. بنابراین، در این گفتار تلاش خواهیم کرد به بررسی و مطالعه نمرات خام دروس عمومی و دروس اختصاصی پذیرفته‌شدگان آزمون سراسری سال ۱۳۹۴ کشور بپردازیم.

**نقش آزمون سراسری به عنوان یک آزمون سرنوشت‌ساز، عبارت است از اندازه‌گیری دقیق و رتبه‌بندی علمی داوطلبان آزمون. آزمون سرنوشت‌ساز آزمونی است که در مقیاس گسترده اجرا شده و بر اساس نتایج آن تصمیماتی خطیر درباره افراد اتخاذ می‌گردد به طوری که این تصمیمات دارای تبعات مهم شخصی، اجتماعی و سیاسی برای افراد هستند**

## سؤال‌های پژوهش

به اذعان برنامه‌ریزان آموزش و پرورش، اندازه‌گیری، مهم‌ترین بخش فرآیند آموزش محسوب می‌شود به طوری که گرونلند<sup>۲۳</sup> و واف<sup>۲۴</sup> (۲۰۱۳) عقیده دارند، سنجش پیشرفت تحصیلی یادگیرنده، نقشی برجسته و بی‌بدیل در آموزش ایفا می‌کند. نتایج سنجش حاوی اطلاعات ارزشمندی است که می‌تواند در مراحل مختلف آموزش مورد استفاده قرار گیرد (کوئین<sup>۲۴</sup>،

۲۰۱۰). نمرات آزمون به منظور اتخاذ تصمیمات دقیق، آگاهانه و عالمانه، اطلاعاتی سازنده در اختیار فراگیران، آموزشیاران، آزمون‌سازان و سیاست‌گذاران قرار می‌دهد. بدیهی است برای دستیابی به این اطلاعات سازنده، داده‌های مربوط به نمرات آزمون باید به‌طور مناسب، تحلیل و تفسیر شوند (کوکی‌نووا<sup>۲۵</sup>، ۲۰۰۴). صاحب‌نظران اندازه‌گیری از قبیل هالادینا<sup>۲۶</sup> (۲۰۰۴) اعتقاد دارند، به کار بستن اصول ساخت آزمون، منجر به تفسیر بهتر و استفاده بیشتر از نمرات آزمون می‌گردد. بدیهی است، عدم به‌کارگیری قواعد ساخت و اصول تحلیل سؤال و آزمون، موجب می‌شود که آزمون‌ها در مقام عمل، مؤلفه‌ای جز خطا را اندازه‌گیری نکنند. از این رو، متخصصان ساخت آزمون، شاخص‌های مختلفی را برای تحلیل آزمون ارائه می‌کنند از جمله دامنه مطلوب دشواری را برای آزمون‌های نرم مرجع همانند آزمون ورودی دانشگاه (با در نظر گرفتن نرخ پذیرش)، بین مقادیر ۳۰٪ تا ۷۰٪ می‌دانند (گرونلند، ۱۹۷۷؛ شولتز<sup>۲۷</sup> و ویتنی<sup>۲۸</sup>، ۲۰۰۵). بنا به اعتقاد آلن<sup>۲۹</sup> و ین<sup>۳۰</sup> (۲۰۰۱)، چنانچه حداکثر تشخیص تمام سطوح عملکرد آزمودنی‌ها خواسته آزمون‌ساز باشد، بهترین دامنه دشواری سؤال‌ها در دامنه ۳۰٪ تا ۷۰٪ قرار می‌گیرد.

مطابق گزارش‌های فنی سازمان سنجش آموزش کشور (به‌عنوان مثال، به کارنامه آماری آزمون سراسری سال ۱۳۹۴ مراجعه کنید)، و اظهارات برخی از متخصصان تعلیم و تربیت و اندازه‌گیری آموزشی، به نظر می‌رسد شاخص‌های روان‌سنجی آزمون سراسری از استانداردهای لازم فاصله یافته است. لذا این پژوهش به منظور توصیف و ارزیابی ویژگی‌های آماری و سنجشی نمرات خام دروس عمومی و دروس اختصاصی پذیرفته‌شدگان آزمون سال ۱۳۹۴ طراحی گردید تا به سؤال‌های زیر پاسخ دهد:

۱. آیا متوسط عملکرد پذیرفته‌شدگان در آزمون عمومی و آزمون اختصاصی از ۳۰ درصد کمتر است؟
۲. آیا متوسط واریانس سوالات عمومی و اختصاصی پذیرفته‌شدگان، خارج از دامنه بهینه ۲۱٪ تا ۲۵٪ است؟
۳. آیا متوسط ضریب دشواری سوالات عمومی و اختصاصی آزمون سراسری سال ۱۳۹۴، از ۳۰٪ کمتر است؟
۴. آیا نسبت پاسخ «نادرست و سفید» پذیرفته‌شدگان به سوالات عمومی و اختصاصی از معیار قابل قبول (۵۰ درصد) بیشتر است؟

## روش پژوهش

در این پژوهش، جامعه آماری شامل تمامی پذیرفته‌شدگان آزمون سراسری سال ۱۳۹۴ در پنج گروه آزمایشی است که از طریق شرکت در آزمون، پذیرش و به دانشگاه معرفی شده‌اند. حجم این جامعه برابر با ۲۷۸۶۸۶ نفر است. در این

مطالعه، نمرات آزمون کلیه اعضای جامعه مورد بررسی قرار گرفت و فرآیند نمونه‌گیری طی نگردید. جدول شماره (۲)، تعداد حاضرین و تعداد پذیرفته‌شدگان آزمون را به تفکیک گروه آزمایشی نشان می‌دهد. همچنین ابزار جمع‌آوری داده‌ها، آزمون سراسری سال ۱۳۹۴ در پنج گروه آزمایشی است. شیوه جمع‌آوری داده‌ها نیز مراجعه به نمرات استخراج شده از برگه‌های پاسخنامه پذیرفته‌شدگان است. روش این پژوهش از لحاظ هدف، کاربردی و از حیث گردآوری داده‌ها، توصیفی-پیمایشی است.

جدول شماره (۲): تعداد حاضرین در جلسه آزمون و تعداد پذیرفته‌شدگان

تعداد گروه	حاضر در جلسه	پذیرفته شده
علوم ریاضی	۱۶۷۸۴۰	۹۳۹۴۶
علوم تجربی	۴۵۲۰۷۳	۹۳۳۳۰
علوم انسانی	۱۶۰۸۸۹	۷۹۹۰۰
هنر	۴۳۶۹۲	۴۹۳۰
زبان	۹۳۸۷۳	۶۵۸۰
جمع کل	۹۱۸۳۸۷ نفر	۲۷۸۶۸۶ نفر

(نرخ پذیرش آزمون تقریباً برابر با ۳۰ درصد است).

## یافته‌های پژوهش الف. بخش آزمون عمومی

جدول شماره (۳)، میانگین نمرات خام پذیرفته‌شدگان آزمون سراسری سال ۱۳۹۴ را در آزمون عمومی و به تفکیک پنج گروه آزمایشی نشان می‌دهد. بدیهی است مقادیر درون جدول، بیانگر میانگین پاسخ‌های درست پذیرفته‌شدگان هر رشته در هر یک از مواد امتحانی است. همچنین، مقادیر کناری جدول که در انتهای هر سطر آورده شده است، میانگین دروس عمومی هر یک از گروه‌های آزمایشی را به صورت کلی و اعداد خارج از جدول که در انتهای پایین هر ستون آمده، میانگین یک درس واحد را در تمام گروه‌های آزمایشی و به صورت یکجا ارائه می‌کند.

چنانچه شرکت‌کنندگان در یک آزمون، نمره صفر، یا نمره کامل و یا نمره برابر و یکسانی کسب کرده باشند، یعنی واریانس نمرات آنها صفر و یا نزدیک به صفر باشد، هیچ هدف سودمندی برای آزمون‌ساز نخواهد داشت

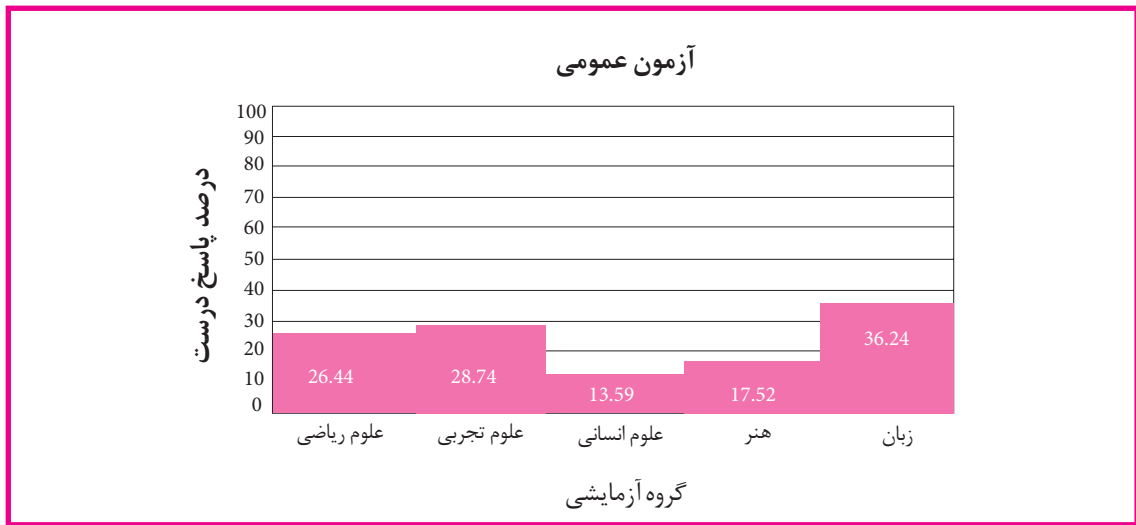
جدول شماره (۳):

میانگین نمرات پذیرفته‌شدگان آزمون سراسری سال ۱۳۹۴ در دروس آزمون عمومی

درس گروه	زبان و ادبیات فارسی	زبان عربی	فرهنگ و معارف	زبان انگلیسی	میانگین کل. %
علوم ریاضی	۱۸/۴۳	۲۱/۸۶	۳۹/۴۳	۲۶/۰۵	۲۶/۴۴
علوم تجربی	۲۵/۹۶	۲۲/۵۲	۴۱/۰۰	۲۵/۵۰	۲۸/۷۴
علوم انسانی	۱۴/۸۴	۱۰/۸۸	۲۱/۲۴	۷/۴۱	۱۳/۵۹
هنر	۱۳/۸۴	۱۲/۲۴	۲۳/۷۴	۲۰/۲۸	۱۷/۵۲
زبان	۲۸/۱۷	۲۲/۵۹	۴۴/۷۸	۴۹/۴۳	۳۶/۲۴
میانگین کل	۲۰/۲۵	۱۸/۰۰	۳۴/۰۰	۲۵/۷۳	۲۴/۵۰

نمره خام دروس آزمون سراسری به صورت درصدی با دامنه تغییرات بین ۳۳/۳۳- و ۱۰۰ محاسبه می‌شود<sup>۳۱</sup>. به‌عنوان مثال، مطابق داده‌های جدول فوق، میانگین نمره درس فرهنگ و معارف اسلامی پذیرفته‌شدگان گروه آزمایشی علوم تجربی، برابر با نمره ۴۱ است. به این معنی که، پذیرفته‌شدگان گروه آزمایشی علوم تجربی در آزمون سراسری سال ۱۳۹۴، به‌طور متوسط به ۴۱ درصد سؤالات فرهنگ و معارف اسلامی (پس از اعمال قاعده نمره منفی) پاسخ درست داده‌اند. همچنین در ماده امتحانی زبان عربی، نمره ۱۸ نشان می‌دهد، میانگین پذیرفته‌شدگان پنج گروه آزمایشی آزمون سال ۱۳۹۴ در درس زبان عربی برابر با ۱۸ درصد است. به این ترتیب، می‌توان اظهار داشت: به‌طور متوسط، این افراد به ۱۸ درصد سؤال‌های درس عربی (۴/۵ سؤال از ۲۵ سؤال آزمون) پاسخ درست داده‌اند. نمودار شماره ۱، تصویر «میانگین‌های کل» را نشان می‌دهد.





علوم انسانی با میانگین ۱۳/۵۹ درصد است.  
 ۵. بیسن میانگین کل گروه آزمایشی علوم ریاضی و گروه آزمایشی علوم تجربی در دروس عمومی، تفاوت محسوسی وجود ندارد، هر چند که میانگین گروه علوم تجربی، اندکی بیشتر است.  
 ۶. ساده‌ترین ماده امتحانی در میان دروس عمومی برای پذیرفته‌شدگان سال ۱۳۹۴، درس فرهنگ و معارف اسلامی با میانگین ۳۴ درصد، و دشوارترین ماده امتحانی در میان این دروس، درس زبان عربی با میانگین ۱۸ درصد است.  
 ۷. از مهم‌ترین نکات قابل توجه اینکه، عملکرد پذیرفته‌شدگان آزمون سال ۱۳۹۴، در درس «زبان خارجی» بهتر از عملکرد آنها در درس «زبان و ادبیات فارسی» بوده است.

### ب. بخش آزمون اختصاصی

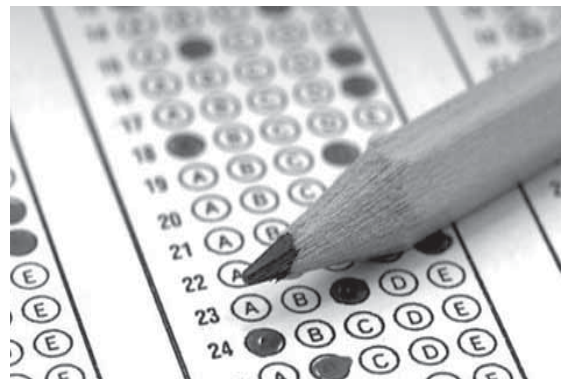
بخش دوم سؤالات آزمون سراسری سؤالات اختصاصی است. در این بخش بر خلاف بخش اول آزمون، تعداد سؤال، تعداد مواد امتحانی و مدت پاسخگویی پنج گروه آزمایشی با یکدیگر متفاوت است. جدول شماره (۴) تفاوت این پارامترها را به تفکیک گروه آزمایشی نشان می‌دهد.

جدول شماره (۴): تعداد مواد امتحانی و تعداد سؤال آزمون اختصاصی

مدت پاسخگویی	تعداد سؤال	تعداد درس	ویژگی گروه
۱۷۵	۱۳۵	۳ درس	علوم ریاضی
۱۷۵	۱۷۰	۵ درس	علوم تجربی
۱۶۵	۱۸۰	۸ درس	علوم انسانی
۱۶۵	۱۶۰	۷ درس	هنر
۱۰۵	۷۰	۱ درس	زبان

از داده‌های جدول شماره (۳) و نمودار (۱)، چند جمع‌بندی قابل استخراج و بیان است که عبارت‌اند از:

۱. میانگین مجموع نمره‌های ۴ درس عمومی پذیرفته‌شدگان آزمون سراسری، برابر با ۲۴/۵۰ درصد است به این معنی که پذیرفته‌شدگان آزمون، به‌طور متوسط، به حدود ۲۵٪ سؤالات پاسخ درست و به ۷۵٪ سؤالات، پاسخ سفید و یا پاسخ نادرست داده‌اند. بنابراین، متوسط دشواری سؤال‌ها برابر با ۲۵٪ و متوسط واریانس سؤال‌ها مساوی با ۱۸٪ است.
۲. میانگین کل چهار درس عمومی پذیرفته‌شدگان آزمون ۱۳۹۴ در دامنه ۱۸ تا ۳۴ درصد قرار می‌گیرد.
۳. قوی‌ترین گروه آزمایشی در دروس عمومی، گروه آزمایشی زبان با میانگین ۳۶/۲۴ درصد است. (باید توجه داشت که داوطلبان گروه آزمایشی زبان عمدتاً از داوطلبان قوی‌تر سایر گروه‌ها تشکیل می‌گردد. به عبارت دیگر، داوطلبان با انگیزه و قوی‌تر گروه‌های آزمایشی مختلف علاوه بر گروه اصلی خود، متقاضی گروه زبان نیز می‌شوند. لذا این نتیجه‌گیری، منطقی است).
۴. ضعیف‌ترین گروه آزمایشی در دروس عمومی، گروه آزمایشی



جدول شماره (۵) تا شماره (۹)، میانگین دروس اختصاصی  
آزمون سراسری را ( به صورت درصد) نشان می دهد :

جدول شماره (۵) : میانگین نمرات اختصاصی گروه آزمایشی علوم ریاضی

میانگین کل ۱۰/۱۱	شیمی	فیزیک	ریاضی	مواد امتحانی
	۷/۳۳	۱۴/۴۰	۸/۱۶	میانگین

جدول شماره (۶) : میانگین نمرات اختصاصی گروه آزمایشی علوم تجربی

میانگین کل ۱۰/۱۱	شیمی	فیزیک	زیست شناسی	ریاضی	زمین شناسی	مواد امتحانی
	۱۵/۳۱	۸/۴۲	۱۳/۴۹	۱۰/۴۲	۲/۸۹	میانگین

جدول شماره (۷) : میانگین نمرات اختصاصی گروه آزمایشی علوم انسانی

میانگین کل ۱۲/۰۵	روان شناسی	فلسفه و منطق	علوم اجتماعی	تاریخ و جغرافیا	زبان عربی	ادبیات فارسی	اقتصاد	ریاضی	مواد امتحانی
	۲۴/۲۹	۳/۹۰	۱۲/۴۴	۱۵/۰۹	۱۴/۳۶	۹/۹۰	۱۴/۵۵	۱/۹۲	میانگین

جدول شماره (۸) : میانگین نمرات اختصاصی گروه آزمایشی هنر

میانگین کل ۹/۸	خواص مواد	خلاقیت موسیقی	خلاقیت نمایشی	خلاقیت تصویری	ترسیم فنی	ریاضی و فیزیک	درک هنر	مواد امتحانی
	۱/۷۸	۷/۵۷	۵/۵۳	۲۷/۵۶	۷/۱۹	۹/۸۹	۹/۱۸	میانگین

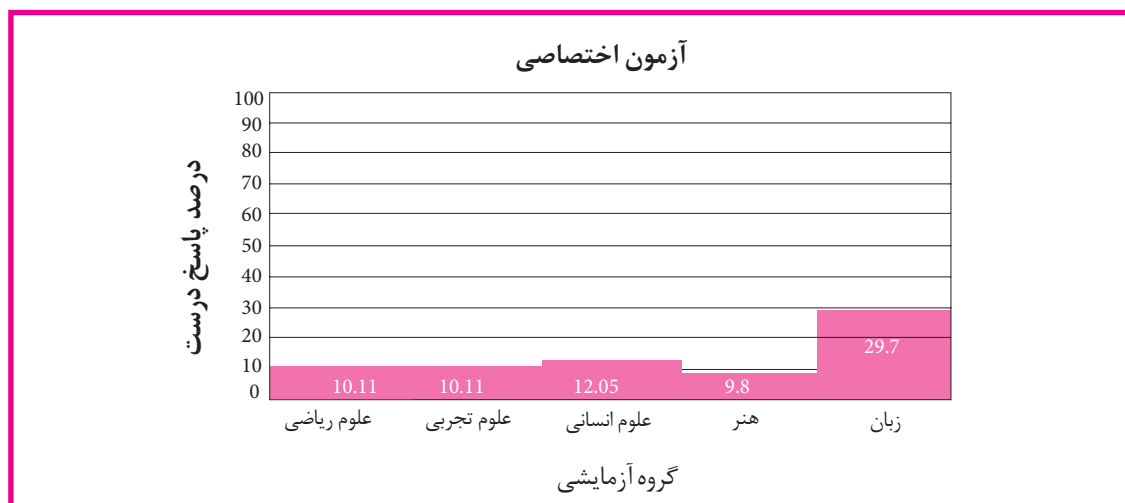
جدول شماره (۹) : میانگین اختصاصی گروه زبان انگلیسی

میانگین کل ۲۹/۷	انگلیسی اختصاصی	ماده امتحانی
	۲۹/۷	میانگین

دشواری سؤالها برابر با ۱۴٪ و متوسط واریانس سؤالها مساوی با ۱۲٪ است. نمودار شماره (۲) میانگین اختصاصی ۵ گروه آزمایشی را به صورت تصویری نشان می دهد.

بررسی داده های جداول بالا نشان می دهد که میانگین مجموع نمره های دروس اختصاصی پذیرفته شدگان پنج گروه آزمایشی روی هم، برابر با ۱۴/۳۵ درصد است. همچنین، متوسط ضریب

نمودار ۲: میانگین اختصاصی پذیرفته شدگان آزمون سراسری ۱۳۹۴



خواص مواد دشوارترین درس و خلاقیت تصویری ساده‌ترین ماده امتحانی است.

۳. عملکرد پذیرفته‌شدگان گروه هنر در دروس اختصاصی، ضعیف‌تر از عملکرد آن‌ها در دروس عمومی است.

### د) گروه آزمایشی زبان

۱. میانگین اختصاصی پذیرفته‌شدگان گروه آزمایشی زبان انگلیسی برابر با ۲۹/۷ است.

۲. در میان میانگین نمرات اختصاصی پنج گروه آزمایشی، بزرگ‌ترین میانگین متعلق به گروه زبان است. ناگفته پیداست که این گروه آزمایشی دارای تنها یک ماده امتحانی اختصاصی است و همانند سایر گروه‌ها دارای دروس امتحانی متعددی نیست. لذا، در این گروه آزمایشی دستیابی به عملکرد تحصیلی مناسب و بالاتر ساده‌تر است. از سوی دیگر، داوطلبان این گروه آزمایشی در مقایسه با داوطلبان سایر گروه‌های آزمایشی، دارای تجانس بیشتری هستند به این دلیل که داوطلبان گروه زبان عمدتاً از داوطلبان قوی‌تر سایر گروه‌ها تشکیل می‌شوند. بنابراین، بدیهی است که انتظار داشته باشیم میانگین این گروه در مقایسه با سایر گروه‌ها از مقدار بیشتری برخوردار باشد.

**مسائل مرتبط با طرح سؤال، ساخت آزمون و به‌طور کلی سنجش، مشابه با دکنکی است که وقتی یک طرف آن را می‌فشاریم، سمت دیگر آن متورم می‌شود. یعنی، هنگامی که به یک جنبه آزمون و سنجش توجه می‌شود از سایر جنبه‌های آن غفلت می‌شود**

### نتیجه‌گیری و پیشنهادها

در این بخش، نتیجه‌گیری از این مطالعه را با بررسی چهار سؤال مطروحه پژوهش ادامه می‌دهیم:

#### ۱) سؤال پژوهشی نخست

«میانگین کل» نمرات خام پذیرفته‌شدگان آزمون سراسری سال ۱۳۹۴ در دروس عمومی، برابر با ۲۴/۵۰ درصد است، به این معنی که این آزمودنی‌ها به‌طور متوسط، تقریباً به ۲۵ درصد سؤال‌ها، پاسخ درست و به ۷۵ درصد سؤال‌ها، پاسخ نادرست و یا سفید داده‌اند. آشکار است که پذیرفته‌شدگان از هر ۴ سؤال، به یک سؤال پاسخ درست و به ۳ سؤال، پاسخ نادرست و یا سفید داده‌اند. به سخن دیگر، به‌طور متوسط، ۲۵ درصد آزمودنی‌ها به هر یک از سؤال‌ها عمومی، پاسخ درست داده‌اند.

سایر یافته‌ها به تفکیک گروه آزمایشی به شرح زیر است:

### الف) گروه آزمایشی علوم ریاضی و علوم تجربی

۱. همان‌طور که از داده‌های جدول شماره ۵ و ۶ بر می‌آید، میانگین دروس اختصاصی قبول‌شدگان در علوم ریاضی برابر با ۱۰/۱۱ و در علوم تجربی نیز برابر با ۱۰/۱۱ است. به عبارت دیگر، پذیرفته‌شدگان این دو رشته به‌طور متوسط، تقریباً به ۱۰ درصد سؤال‌ها اختصاصی پاسخ درست و به ۹۰ درصد سؤال‌ها، پاسخ غلط و یا پاسخ سفید داده‌اند. (خواننده توجه دارد که میانگین دروس عمومی این دو گروه آزمایشی نیز تقریباً برابر یکدیگر است).

۲. پذیرفته‌شدگان گروه آزمایشی علوم ریاضی از میان دروس اختصاصی، در درس فیزیک (با میانگین ۱۴/۴۰ درصد) عملکرد بهتری داشته‌اند، در حالی که پذیرفته‌شدگان علوم تجربی از میان دروس اختصاصی، در درس شیمی (با میانگین ۱۵/۳۱) وضعیت مناسب‌تری دارند.

۳. عملکرد قبول‌شدگان علوم ریاضی و علوم تجربی در دروس عمومی بهتر از دروس اختصاصی است.

### ب) گروه آزمایشی علوم انسانی

۱. همان‌طور که در جدول شماره ۷ مشاهده می‌شود، میانگین دروس اختصاصی داوطلبان پذیرفته‌شده علوم انسانی برابر با ۱۲/۰۵ است. به بیان دیگر، این افراد حدود ۱۲ درصد سؤال‌ها اختصاصی را به‌طور درست و ۸۸ درصد سؤال‌ها را نادرست و یا سفید پاسخ داده‌اند. در این گروه آزمایشی، کمترین میانگین مربوط به درس ریاضی با مقدار ۱/۹۲ درصد است. به این معنی که پذیرفته‌شدگان این گروه آزمایشی از میان ۲۰ سؤال ریاضی، به‌طور متوسط به کمتر از یک سؤال پاسخ درست داده‌اند!

۲. پس از ریاضی، دشوارترین ماده امتحانی، فلسفه و منطق با میانگین ۲/۹ است. از طرف دیگر، روان‌شناسی با میانگین ۲۴/۲۹، به‌طور نسبی، ساده‌ترین ماده امتحانی این مجموعه محسوب می‌شود.

۳. عملکرد پذیرفته‌شدگان علوم انسانی در دروس عمومی و دروس اختصاصی تقریباً یکسان و برابر است.

### ج) گروه آزمایشی هنر

۱. میانگین پذیرفته‌شدگان گروه آزمایشی هنر از ۷ ماده امتحانی، برابر با ۹/۸ است. به سخن دیگر، این افراد به‌طور متوسط، به حدود ۱۰ درصد سؤال‌ها پاسخ درست و به ۹۰ درصد سؤال‌ها پاسخ غلط و یا پاسخ سفید داده‌اند.

۲. در گروه آزمایشی هنر، کمترین نمره مربوط به درس خواص مواد با میانگین ۱/۷۸ درصد و بیشترین نمره نیز مربوط به درس خلاقیت تصویری با مقدار ۲۷/۵۶ درصد است. بنابراین،



از سوی دیگر، «میانگین کل» نمرات خام پذیرفته‌شدگان در دروس اختصاصی برابر با ۱۴/۳۵ درصد است. لذا می‌توان گفت که این پذیرفته‌شدگان به‌طور متوسط، تقریباً به ۱۴ درصد سؤال‌ها، پاسخ درست و به ۸۶ درصد مابقی، پاسخ نادرست و یا سفید داده‌اند. به بیان دیگر، قبول‌شدگان آزمون سراسری، از هر ۷ سؤال اختصاصی، به ۱ سؤال پاسخ درست داده‌اند. بنابراین می‌توان گفت متوسط عملکرد پذیرفته‌شدگان در آزمون عمومی و آزمون اختصاصی کمتر از ۳۰ درصد است. لذا، پاسخ سؤال نخست، مثبت است.

## ۲) سؤال پژوهشی دوم

متوسط واریانس سؤالات آزمون عمومی برابر با ۰/۱۸ و متوسط واریانس سؤالات آزمون اختصاصی مساوی با ۰/۱۲ است. بنابراین واریانس سؤالات آزمون عمومی و آزمون اختصاصی خارج از دامنهٔ بهینه ۰/۲۱ تا ۰/۲۵ قرار می‌گیرد، لذا پاسخ سؤال دوم پژوهش مبنی بر اینکه آیا متوسط واریانس سؤالات عمومی و اختصاصی پذیرفته‌شدگان، خارج از دامنه بهینه قرار دارد، مثبت است.

## ۳) سؤال پژوهشی سوم

متوسط ضریب دشواری سؤالات آزمون عمومی پذیرفته‌شدگان سال ۱۳۹۴ برابر با ۰/۲۵ و متوسط دشواری سؤالات آزمون اختصاصی مساوی با ۰/۱۴ است. لذا، ضریب دشواری هر دو آزمون عمومی و آزمون اختصاصی پذیرفته‌شدگان از ۰/۳۰ کمتر است. این نتایج حاکی از دشواری بسیار زیاد سؤالات آزمون سال ۱۳۹۴ و عدم تناسب و همخوانی بین سطح دشواری سؤالات و سطح توانایی آزمودنی‌هاست. بنابراین، پاسخ سؤال سوم پژوهش دال بر اینکه آیا متوسط دشواری سؤالات عمومی و اختصاصی پذیرفته‌شدگان آزمون سراسری سال ۱۳۹۴ از ۰/۳۰ پایین‌تر است، مثبت است.

## ۴) سؤال پژوهشی چهارم

همان‌طور که گفته شد، در بخش آزمون عمومی، میانگین پذیرفته‌شدگان تقریباً برابر با ۲۵ درصد است، به این معنی که آزمودنی‌ها به‌طور متوسط، به ۲۵ درصد سؤال‌ها پاسخ درست و به ۷۵ درصد سؤال‌ها، پاسخ نادرست و یا سفید داده‌اند. همچنین، در بخش آزمون اختصاصی، میانگین پذیرفته‌شدگان مساوی با ۰/۱۴ است. به این ترتیب، آزمودنی‌ها به ۱۴ درصد سؤالات پاسخ درست و به ۸۶ درصد، پاسخ نادرست و یا سفید داده‌اند. لذا، پاسخ سؤال چهارم پژوهش نیز مثبت است. تحلیل داده‌ها نشان می‌دهد که میان شاخص‌های حاصل از این پژوهش (وضع موجود) با مقادیر مورد انتظار (وضع مطلوب) که پیش‌تر بدان اشاره شد، تفاوت زیادی وجود دارد. بررسی

جدول شماره (۱۰) نشان می‌دهد که این مقادیر برای آزمون ورودی سال ۱۳۹۴ دانشگاه که یک آزمون کشوری هنجاری با نرخ پذیرش ۳۰ درصد محسوب می‌شود، دارای یک درجه استاندارد اطمینان‌بخش به این منظور که گفته شود «آزمون حداکثر قدرت تشخیص در تمام سطوح عملکرد آزمودنی‌ها را دارا می‌باشد»، نیست. در صورتی که مطابق چارچوب نظریه کلاسیک آزمون (CTT)، ضرورت دارد متوسط ضریب دشواری سؤالات آزمون حدود ۰/۵۰، دشواری سؤالات در دامنه ۰/۳۰ تا ۰/۷۰ و نیز واریانس سؤالات در دامنه ۰/۲۱ تا ۰/۲۵ قرار گیرد. (خواننده نکته‌بین توجه دارد چنانچه این شاخص‌ها در جامعهٔ شرکت‌کنندگان آزمون سراسری و نه فقط قبول‌شدگان محاسبه و بررسی گردد، سطح میانگین نمرات و نیز سایر مقادیر، ضعیف‌تر و پایین‌تر خواهد بود).

علاوه بر چارچوب نظریه، به منظور این که شاخص و ملاکی عملی به دست آوریم تا بتوانیم داده‌های آزمون سال ۱۳۹۴ را با آن مقایسه کنیم، داده‌های آزمون سراسری سال ۱۳۸۲ به عنوان نمونه در جدول (۱۰) آورده شده تا امکان تحلیل و قضاوت بهتر فراهم گردد.

جدول شماره (۱۰): مقایسهٔ پذیرفته‌شدگان در آزمون سراسری

۱۳۸۲ و ۱۳۹۴

شاخص	آزمون		عمومی		اختصاصی	
	۱۳۸۲	۱۳۹۴	۱۳۸۲	۱۳۹۴	۱۳۸۲	۱۳۹۴
میانگین نمرات	۴۷/۸۰٪	۲۴/۵۰٪	۴۰/۷۰٪	۱۴/۳۵٪		
متوسط دشواری سؤالات	۰/۴۸	۰/۲۵	۰/۴۱	۰/۱۴		
متوسط واریانس سؤالات	۰/۲۵	۰/۱۸	۰/۲۴	۰/۱۲		

بنابراین، نقش آزمون سراسری به عنوان یک آزمون سرنوشت‌ساز، عبارت است از اندازه‌گیری دقیق و رتبه‌بندی علمی تمامی داوطلبان آزمون. آزمون سرنوشت‌ساز، آزمونی است که بر اساس نتایج آن تصمیماتی پرخطر دربارهٔ افراد اتخاذ می‌گردد به‌طوری که این تصمیمات دارای پیامدهای مهم شخصی، اجتماعی و سیاسی برای افراد هستند. بدیهی است، چنانچه اندازه‌گیری با دقت و ثبات کافی انجام نپذیرد، دست‌کم، بخشی از پذیرفته‌شدگان آزمون سراسری وارد آن دسته از رشته‌هایی خواهند شد که یا انگیزه و علاقه کافی برای تلاش ندارند و یا توانایی پیشرفت تحصیلی مناسب در آن رشته‌ها را نخواهند داشت و به این ترتیب، بخشی از سرمایه‌های مادی و انسانی جامعه تلف خواهد شد.

با توجه به تحلیل داده‌های ارائه شده و در راستای ارتقای کیفیت آزمون‌های ملی به منظور بهبود دقت اندازه‌گیری پیشرفت تحصیلی داوطلبان آزمون سراسری، سلسله اقدامات زیر پیشنهاد می‌گردد:

## ۱. منطبق ساختن سطح دشواری سؤال‌ها با متوسط توانایی داوطلبان

در این بخش، بحث را با طرح یک سؤال ساده دنبال می‌کنیم: «فرض کنید تمام یا تعداد بسیاری از آزمودنی‌های یک آزمون به صورت کاملاً تصادفی به سؤال‌های آزمون پاسخ دهند. در این وضعیت، واریانس نمره‌های آزمون چگونه خواهد شد؟» بدیهی است، در چنین شرایطی با وجود پاسخ‌های تصادفی، انتظار می‌رود ضریب دشواری سؤال‌ها برابر با صفر و یا نزدیک به صفر و همبستگی بین سؤال‌ها نیز صفر و یا نزدیک به صفر باشد. بنابراین، واریانس آزمون نیز، صفر و یا نزدیک به صفر خواهد بود (فرگوسن و تاکانه، ۱۹۸۹). واضح و روشن است که این نوع کیفیت آزمون، مورد قبول و پذیرش آزمون‌ساز نیست. با در نظر گرفتن این مقدمه، چنانچه ضریب دشواری سؤالات آزمون سراسری بسیار فراتر از متوسط توانایی داوطلبان آزمون باشد، دست‌کم، انتخاب پاسخ بخشی از سؤال‌ها، از حالت انتخاب عالمانه<sup>۳۲</sup> به سمت پاسخ تصادفی و انتخاب کورکورانه<sup>۳۳</sup> (یا پاسخ سفید) سوق می‌یابد. در این شرایط، واریانس سؤال‌های آزمون کاهش یافته و مسئله تفکیک داوطلبان از یکدیگر به بدترین وضعیت ممکن دچار می‌گردد. در واقع «یک آزمون، زمانی دقیق‌ترین اندازه‌گیری را از توانایی آزمودنی فراهم می‌آورد که دشواری آزمون منطبق بر سطح توانایی آزمودنی باشد» (همبلتون و همکاران، ۱۹۹۱، ص ۱۴۵، ترجمه فلسفی‌نژاد، ۱۳۸۹، ص ۱۸۵). لذا، متخصصان اندازه‌گیری همواره توصیه می‌کنند تا حد امکان، سطح دشواری آزمون با متوسط توانایی آزمودنی‌ها منطبق گردد.

## ۲. کاهش تعداد سؤال‌های آزمون عمومی

در حال حاضر، هر دو بخش «عمومی و اختصاصی» آزمون سراسری در یک نشست نیم‌روزه برگزار می‌گردد. لذا با در نظر گرفتن محدودیت زمانی، آزمون‌ساز ناگزیر است آزمون را با ملاحظات و محدودیت زمان تنظیم کند. همان‌طور که گفته شد، آزمون عمومی از ۱۰۰ سؤال تشکیل می‌شود که آزمودنی باید در مدت ۷۵ دقیقه به آن‌ها پاسخ دهد که به نظر می‌رسد تعداد ۱۰۰ سؤال نسبتاً قدرتمند با مدت پاسخگویی ۷۵ دقیقه برای بخش آزمون عمومی، بسیار سختگیرانه است. بنابراین، نخستین قدم در ارتقای دقت اندازه‌گیری و بهبود ضریب دشواری و واریانس سؤال‌های آزمون سراسری با در نظر گرفتن

پارامتر زمان، کاستن از تعداد سؤال‌های آزمون عمومی و آزمون اختصاصی تارسیدن به یک نقطه بهینه است. پژوهش‌های انجام یافته (از جمله ایبل<sup>۳۴</sup>، ۱۹۶۹ و آسموس<sup>۳۵</sup>، ۱۹۸۱) در خصوص ارتباط میان تعداد سؤال و ضریب پایایی آزمون، نشان می‌دهد که پایایی آزمون‌های عینی چهارگزینه‌ای در مرز ۱۴۰ سؤال می‌تواند به نقطه بهینه ۹۰٪ برسد. لذا، کاهش تعداد سؤالات بخش عمومی آزمون از ۱۰۰ سؤال فعلی به ۶۰ تا ۷۵ سؤال با حفظ مدت پاسخگویی ۷۵ دقیقه، اقدامی منطقی و عاقلانه خواهد بود.

برخلاف آزمون سرعت که در آن سرعت عمل بالای آزمودنی اهمیت دارد، آزمون ورودی دانشگاه به عنوان یک آزمون پیشرفت تحصیلی، یک آزمون قدرت-زمان‌دار<sup>۳۶</sup>، محسوب می‌شود (نانالی<sup>۳۷</sup> و برنشتاین<sup>۳۸</sup>، ۱۹۹۴). در آزمون قدرت زمان‌دار، سرعت عمل آزمودنی اولویت ندارد و فرض بر این است عمدتاً آزمودنی‌ها فرصت کافی در اختیار داشته باشند و تمامی سؤال‌ها را مرور کرده تا هر داوطلب با اتکا به میزان توانایی خود بتواند به تمام یا بخشی از سؤال‌ها، پاسخ درست بدهد (اگرچه ضرورت دارد آزمون‌ساز سؤالات را از لحاظ دشواری به‌گونه‌ای تنظیم نماید تا هیچ فردی به کف<sup>۳۹</sup> نمره و یا سقف<sup>۴۰</sup> نمره آزمون دست نیابد). در واقع، بایستی فرصت کافی در اختیار داوطلبان قرار گیرد تا بتوانند حداکثر تفاوت‌های فردی خود را بروز دهند. بدیهی است، چنانچه زمان مناسب برای بروز و معرفی تفاوت‌های فردی موجود نباشد، عملکرد داوطلبان نزدیک به یکدیگر بوده و به این ترتیب، واریانس آزمون، کاهش و خطای اندازه‌گیری افزایش یافته، و به دنبال آن، رتبه‌بندی داوطلبان غیرپایا<sup>۴۱</sup> و غیردقیق<sup>۴۲</sup> گردیده و هدف آزمون‌ساز که همانا تفکیک مطلوب آزمودنی‌ها از یکدیگر است محقق نخواهد شد.

**فرض بر این است عمده آزمون‌ها فرصت کافی در اختیار داشته باشند تمامی سؤال‌ها را مرور کرده تا هر داوطلب با اتکا به میزان توانایی خود بتواند به تمام یا بخشی از سؤال‌ها، پاسخ درست بدهد (اگرچه ضرورت دارد آزمون‌ساز سؤالات را از لحاظ دشواری به‌گونه‌ای تنظیم نماید تا هیچ فردی به کف نمره و یا سقف نمره آزمون دست نیابد). در واقع، بایستی فرصت کافی در اختیار داوطلبان قرار گیرد تا بتوانند حداکثر تفاوت‌های فردی خود را بروز دهند**

### ۳. کاهش تعداد سؤال‌های آزمون اختصاصی

به استثنای گروه آزمایشی زبان که تعداد سؤال‌های اختصاصی فعلی آن منطقی به نظر می‌رسد، تعداد سؤال‌های اختصاصی گروه‌های آزمایشی علوم ریاضی، علوم تجربی، علوم انسانی و هنر بسیار زیاد بوده و کاهش تعداد آن‌ها اقدامی اجتناب‌ناپذیر می‌نماید. جدول شماره (۱۱) تعداد سؤال و مدت پاسخگویی پیشنهادی برای هر ماده امتحانی را به تفکیک در ۴ گروه آزمایشی و بر اساس اهمیت دروس و میزان محتوا نشان می‌دهد.

### ۴. کاهش تعداد دروس امتحانی

اگرچه تعداد دروس امتحانی گروه‌های آزمایشی علوم ریاضی و زبان‌های خارجی، منطقی و منصفانه به نظر می‌رسد، لیکن تعداد مواد امتحانی گروه‌های علوم تجربی، علوم انسانی و گروه هنر دارای تنوع زیاد و بسیار متعدد است. جدول شماره ۱۲ تعداد دروس امتحانی را به تفکیک گروه آزمایشی نشان می‌دهند. به عنوان مثال، در گروه علوم انسانی، ۴ درس عمومی به علاوه ۸ درس اختصاصی، مجموعاً ۱۲ درس آزمون را تشکیل

جدول شماره (۱۱): تعداد سؤال‌ها و مدت پاسخگویی پیشنهادی برای دروس اختصاصی

ویژگی گروه	مواد امتحانی	تعداد سؤال فعلی	تعداد سؤال پیشنهادی	مدت پاسخگویی پیشنهادی
علوم ریاضی	۱- ریاضی	۵۵	۳۵	۹۰ سؤال ۱۸۰ دقیقه
	۲- فیزیک	۴۵	۳۰	
	۳- شیمی	۳۵	۲۵	
علوم تجربی	۱- زمین‌شناسی	۲۵	۲۰	۱۳۰ سؤال ۱۸۰ دقیقه
	۲- ریاضی	۳۰	۲۵	
	۳- زیست‌شناسی	۵۰	۳۵	
	۴- فیزیک	۳۰	۲۵	
	۵- شیمی	۳۵	۲۵	
علوم انسانی	۱- ریاضی	۲۰	۱۵	۱۴۰ سؤال ۱۵۰ دقیقه
	۲- اقتصاد	۱۵	۱۵	
	۳- ادبیات فارسی	۳۰	۲۰	
	۴- زبان عربی	۲۰	۲۰	
	۵- تاریخ و جغرافیا	۳۰	۲۰	
	۶- علوم اجتماعی	۲۰	۱۵	
	۷- فلسفه و منطق	۲۵	۲۰	
	۸- روان‌شناسی	۲۰	۱۵	
هنر	۱- درک عمومی هنر	۳۰	۲۰	۱۱۰ سؤال ۱۵۰ دقیقه
	۲- درک ریاضی فیزیک	۳۰	۱۵	
	۳- ترسیم فنی	۲۰	۱۵	
	۴- خلاقیت تصویری	۲۰	۱۵	
	۵- خلاقیت نمایشی	۲۰	۱۵	
	۶- خلاقیت موسیقی	۲۰	۱۵	
	۷- خواص مواد	۲۰	۱۵	

- ing Company, Inc. Reading, Mass.  
 Nunnally, J. C.; & Bernstein, I.H. (1994). *Psychometric theory*. New York: McGraw-Hill.  
 Ritt, M. (2016). *The impact of high-stakes testing on the learning environment*. Master of social work clinical research papers. Paper 658.  
 Shultz, K.S.; Whitney, D.J. (2005). *Measurement theory in action*. Thousand Oaks, CA: sage.  
 Stocking, M. L. & Lewis, C. (2000). Methods of controlling the exposure of items in CAT. In W. J. Van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: Theory and practice*. Netherlands: Kluwer Academic Publishers.  
 Osterlind, S. J. (2002). Constructing test items: Multiple-choice, constructed-response, performance, and other formats. Boston: Kluwer.  
 Popham, W.J. (2005). High-stakes tests: harmful, permanent, fixable. *American Educational Research*. 6, p85.  
 Quinn, G.W. (2010). *Improving test scores in five easy steps*. Rowman & Littlefield Education.

می‌دهد. در این گروه، برخی مواد امتحانی همانند درس «زبان و ادبیات فارسی» و یا درس «زبان عربی» هم در بخش آزمون عمومی و هم در بخش آزمون اختصاصی وجود دارد که می‌تواند دو درس را به یک درس کاهش داد.

جدول شماره (۱۲): تعداد مواد امتحانی گروه‌های آزمایشی

تعداد سؤال	تعداد درس	ویژگی گروه
۲۳۵	۷ درس	علوم ریاضی
۲۷۰	۹ درس	علوم تجربی
۲۸۰	۱۲ درس	علوم انسانی
۲۶۰	۱۱ درس	هنر
۱۷۰	۵ درس	زبان

#### پی‌نوشت‌ها

1. Fair
2. Accuracy
- 3 Measurement error
- 4 High-stakes test
5. Ritt
6. Popham
7. Stability
- 8 . Hambleton
9. Stocking
10. Lewis
11. Norm- referenced
12. Ferguson
13. Takane
14. Magnusson
15. Classical test theory
16. Difficulty
17. Gulliksen
18. Crocker
19. Algina
20. Ranking
21. Osterlind
22. Gronlund
23. Waugh
24. Quinn
25. Kokinova
26. Haladyna
27. Shultz
28. Whitney
29. Allen
30. Yen

۳۱. در آزمون سراسری به هر پاسخ درست آزمودنی، ۳ نمره مثبت و به هر پاسخ نادرست، ۱ نمره منفی تعلق می‌گیرد. بنابراین، چنانچه داوطلب به تمام سوالات یک درس پاسخ درست بدهد، نمره خام درصدی وی در آن درس برابر با ۱۰۰ و چنانچه به تمام سوالات یک درس پاسخ نادرست بدهد، نمره وی در آن درس مساوی با ۲۳/۳- خواهد شد و مابقی احتمالات در این دامنه قرار می‌گیرد.

32. Wisely
33. Blind
34. Ebel
35. Asmus
36. Timed- power test
37. Nunnally
38. Bernstein
39. Floor
40. Ceiling
41. Unreliable
42. Inaccurate

#### منابع فارسی

- پروند، محمد حسن. (۱۳۸۱). مقدمات برنامه‌ریزی آموزشی و درسی. تهران: دنیای پژوهش.  
 جانزاده، علیرضا (۱۳۸۹). بررسی عوامل اثرگذار بر افت تحصیلی دانشجویان دانشگاه‌های دولتی. آموزش عالی ایران. سال سوم، شماره سوم.  
 دفتر طرح و آمار سازمان سنجش آموزش کشور (۱۳۹۵). کارنامه آماری آزمون سراسری سال ۱۳۹۴. تهران: انتشارات سازمان سنجش آموزش کشور.  
 دفتر طرح و آمار سازمان سنجش آموزش کشور (۱۳۸۳). کارنامه آماری آزمون سراسری سال ۱۳۸۲. تهران: انتشارات سازمان سنجش آموزش کشور.  
 همبلیتون، رونالد. و سوامیناتان، هارپهاران. و راجرز، جین (۱۹۹۱). مبانی نظریهٔ پرسش- پاسخ. ترجمه: محمدرضا فلسفی‌نژاد (۱۳۸۹). تهران: انتشارات دانشگاه علامه طباطبایی.

#### منابع انگلیسی

- Allen, M. J., & Yen, W. M. (2001). *Introduction to measurement theory*. Monterey, CA: Brooks.  
 Asmus, E. P. (1981). The Effect of Altering the Number of Choices per Item on Test Statistics: Is Three Better than Five? *Bulletin of the Council for Research in Music Education*. 65, pp 1-15.  
 Croker, L. M., & Algina, J. (2008). *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth.  
 Ebel, R.L. (1969). Expected reliability as a function of choices per item. *Educational and Psychological Measurement*. 29, p565.  
 Ferguson, G. A., & Takane, Y. (1989). *Statistical analysis in psychology and education*. New York: McGraw-Hill.  
 Gronlund, N.E. (1977). *Constructing achievement tests*. Englewood Cliffs, New Jersey: Prentice-Hall. Inc.  
 Gronlund, N.E., & Waugh, C.K. (2013). *Assessment of student achievement*. Upper saddle River, NJ: Pearson Education.  
 Gulliksen, H. (1987). *Theory of mental tests*. New York: John Wiley & Sons.  
 Haladyna, T.M. (2004). *Developing and validating multiple-choice test items*. Mahwah, NJ: Lawrence Erlbaum Association.  
 Hambleton, R.K.; Swaminathan, H. & Rogers, H.J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: sage.  
 Kokinova, M. (2004). *Psychometric analysis of student's performance on basic medical science tests*. Kent: Kent State University.  
 Magnusson, D. (1967). *Test theory*. Addison-Wesley Publish-