

داده‌های گمشده در آزمون‌های سراسری ورود به دانشگاه: مبانی نظری و شواهد مبتنی بر داده‌های واقعی

مریم چگینی *
ابراهیم خدایی **
ولی‌الله فرزاد ***
بلال ایزانلو ****

چکیده:

داده‌های گمشده پدیده رایج مطالعات تجربی و سنجش‌های آموزشی و روان‌شناسی هستند که به هرگونه بدون پاسخ ماندن سؤال اطلاق می‌شود. روش‌های آماری بسیاری برای محاسبات مجموعه داده‌های دارای پاسخ گمشده وجود دارد که تحت تأثیر مکانیسم گمشدگی، علل و میزان آنها هستند. هدف از نگارش این مقاله، بررسی وضعیت موجود داده‌های گمشده در آزمون سراسری ورودی دانشگاه‌هاست. با استفاده از روش‌های آمار توصیفی و نرم‌افزارهای SPSS و R نشان داده شد میزان داده‌های گمشده در سال‌های مورد بررسی افزایش یافته است (در دامنه‌ای بین ۲٫۲٪ تا ۹۱٫۶٪) و شاخص‌های آماری آزمون تحت تأثیر میزان داده‌های گمشده قرار دارد. همچنین نشان داده شد همبستگی مثبت و بالایی بین تعداد پرسش‌های بدون پاسخ در درس‌های مختلف (0.78, 0.41=r)؛ و همبستگی منفی و بالایی بین تعداد پاسخ‌های درست و تعداد پاسخ‌های گمشده (0.85, -0.56=r) وجود دارد. نتایج به دست آمده از این مقاله، نشان از ثبات رخداد داده‌های گمشده در ابعاد مختلف شایستگی و غیرقابل اغماض بودن داده‌های گمشده در تحلیل‌های آماری دارد.

واژگان کلیدی: داده‌های گمشده، الگوی گمشدگی، مکانیسم گمشدگی، آزمون‌های پیشرفت تحصیلی

* دانشجوی دکتری سنجش و اندازه‌گیری، دانشگاه تهران، دانشکده روان‌شناسی و علوم تربیتی

** دانشیار دانشگاه تهران، دانشکده روان‌شناسی و علوم تربیتی (نویسنده مسئول: khodaie@ut.ac.ir)

*** استاد دانشگاه خوارزمی

**** استادیار دانشگاه خوارزمی

مقدمه

در هر آزمون ویژگی‌های روان‌شناختی، سؤال آزمون واحدی از اندازه‌گیری با محرک‌ها و فرم مشخصی از پاسخ است که به وسیله آن عملکرد فرد در برخی سازه‌های روان‌شناختی (مانند توانایی، استعداد، صفت) استنباط می‌شود (اوسترلیند^۱، ۱۹۹۰). در اندازه‌گیری‌های آموزشی و روان‌شناختی از آزمودنی خواسته می‌شود به تعدادی سؤال پاسخ دهد. اختصاص عدد به واحدهای مشاهده شده^۲ بر اساس برخی قوانین صریح و روشن انجام می‌گیرد. این شیوه گاهی نمره‌گذاری نامیده می‌شود. پاسخ‌های آزمودنی به این سؤال‌ها با استفاده از مدل‌های IRT مقیاس‌گذاری می‌شود که استنباط‌هایی از سطوح توانایی فرد فراهم می‌کند (رز^۳، ۲۰۱۳). گاهی آزمودنی‌ها در پاسخ به سؤال‌های ارائه شده شکست می‌خورند (کهلر، پل و کارستنسن^۴، ۲۰۱۵) و سؤال‌ها را بدون پاسخ رها می‌کنند.

داده‌های گمشده^۵ ویژگی ذاتی اغلب سنجش‌ها در سطح بین‌المللی، ملی و ایالتی است. در مفهوم کلی، داده گمشده به هر پاسخ خالی در مجموعه کلی سؤال‌های در دسترس آزمودنی منتسب می‌شود. این مسئله ممکن است به دلایل مختلفی رخ دهد از جمله اینکه در اغلب موارد، هیچ کدام از افرادی که در آزمون شرکت می‌کنند برای همه سؤال‌ها تلاش نمی‌کنند (لادلو و آلری^۶، ۱۹۹۹). داده‌های گمشده برنامه‌ریزی نشده، باید از داده‌های گمشده برنامه‌ریزی شده متمایز شوند. طرح‌هایی از داده‌های گمشده برنامه‌ریزی شده وجود دارد که به دلیل تعداد کم شرکت‌کننده‌ها یا کاهش هزینه‌ها به کار می‌روند (گراهام، تیلور، آلچوسکی و کامسیل^۷، ۲۰۰۶؛ گراهام، ۲۰۰۹؛ رز، ۲۰۱۳). در ارزیابی‌های مقیاس بزرگ آموزشی^۸ (LSAs)، داده‌های گمشده برنامه‌ریزی شده به دلیل طراحی چندماتریسی^۹ رایج هستند، مواردی که تنها زیرمجموعه‌ای از صدها سؤال به هر آزمودن‌دهنده ارائه می‌شود (توماس، راگانانان،

1. Osterlind

2. observational units

3. Rose

4. Köhler, Pohl & Carstensen

5. Missing Data

6. Ludlow & O'leary

7. Graham, Taylor, Olchowski & Cumsille

8. Large-scale assessment

9. Multi-matrix designs

چنکر، کاتزوف و جانسون^۱؛ ۲۰۰۶؛ فری، هارتیگ و راپ^۲؛ ۲۰۰۹؛ ژانگ^۳، ۲۰۱۳). داده‌های گمشده برنامه‌ریزی نشده در یک محیط آموزشی ممکن است مربوط به از دست دادن مشاهدات باشد که یا به این دلیل است که آزمودنی سؤال را حذف کرده^۴ یا آزمون را به علت محدودیت زمانی آن کامل نکرده است^۵. آزمودنی ممکن است یک سؤال را به اشتباه یا گذشتن عمدی از سؤال^۶؛ به دلیل اینکه بسیار دشوار بوده یا واضح نبوده؛ حذف کند. همچنین، آزمودنی ممکن است سؤال‌های پایانی آزمون را پاسخ ندهد؛ به دلیل اینکه زمان آزمون به پایان رسیده یا سؤال‌ها برای آنها اجرا نشده اند (لرد^۷، ۱۹۸۰؛ لادلوو و آلری، ۱۹۹۹؛ ژانگ، ۲۰۱۳).

همچنین آزمودنی‌ها راهبردهای مختلف آزمون دادن را به کار می‌گیرند و این راهبردها بر تعداد سؤال‌هایی که هر آزمودنی برای آن تلاش می‌کند، تأثیر دارد. برای مثال، نظام‌های آموزشی متفاوت در سطح ملی، عادت‌ها، قوانین و خط مشی‌هایشان در مورد راهبردهای آزمون گرفتن متفاوت است. این موضوع به‌ویژه در مورد حدس زدن صحیح است. از طرفی، سیستم‌هایی وجود دارد که زمانی که آزمودنی پاسخ سؤال را نمی‌داند صراحتاً حدس زدن را تشویق می‌کند و زمانی که آزمودنی‌ها از این راهبرد مطلع می‌شوند و آن را دنبال می‌کنند، انتظار می‌رود حداقل سؤال‌های حذف شده و به پاسخ نرسیده وجود داشته باشد. از سویی دیگر، سیستم‌هایی وجود دارد که حدس را تشویق نمی‌کنند و تنها پاسخ به سؤال‌هایی را لازم می‌دانند که آزمودنی دانش آنها را دارد. اجتناب‌ناپذیر است که رهنمودها و خط مشی‌های سیستم‌ها بر گستره سؤال‌های حذف شده و به پاسخ نرسیده تأثیرگذار است تا جایی که آزمودنی‌ها بسته به دستورالعمل‌هایی که به آنها داده شده بیشتر یا کمتر برای سؤال‌ها تلاش خواهند کرد. در نتیجه، اگر گروهی از آزمودنی‌ها به‌طور سیستماتیک برای سؤال‌های کمتری تلاش کنند، پس سطح توانایی‌شان ممکن است به‌طور جدی کم برآورد شود. به‌علاوه، سؤال‌های پایانی در آزمون‌شان ممکن است دشوارتر از آنچه واقعاً هستند به

1. Thomas, Raghunathan, Schenker, Katzoff & Johnson

2. Frey, Hartig & Rupp

3. Zhang

4. Omitted

5. Not-reached

6. Skipped

7. Lord

نظر برسند (لادلوو و آلری، ۱۹۹۹؛ ژانگ، ۲۰۱۳؛ پل، گراف و رز، ۲۰۱۴). از این رو، انواع مختلفی از داده‌های گمشده می‌تواند هم‌زمان در مجموعه‌ای واحد از داده‌ها رخ دهد و احتمالاً به روش‌های برخورد متفاوتی نیاز دارند (رز، ۲۰۱۳).

به‌طور کلی، حضور مقادیر گمشده در داده‌ها سه مشکل عمده ایجاد می‌کند: نخست اینکه قسمتی از اطلاعات از دست می‌رود، در نتیجه با کاهش حجم نمونه و متعاقباً با کاهش دقت و کارایی برآوردها روبه‌رو هستیم. از سوی دیگر، حضور این مقادیر به ایجاد مجموعه داده‌ای نامتعادل منجر می‌شود، زیرا همه افراد دارای تعداد یکسانی از اندازه‌گیری متغیرها نخواهند بود. بنابراین امکان استفاده مستقیم از مدل‌ها و روش‌هایی که برای داده‌های کامل قابل اجرا هستند، وجود ندارد. علاوه بر موارد فوق، از آنجا که ممکن است اطلاعات گمشده مربوط به افرادی باشد که تفاوت‌های اساسی با افراد دیگر (افراد با اطلاعات کامل) داشته باشند، پتانسیل سوگیری وجود دارد. برطرف کردن سوگیری در این موارد مشکل است زیرا دلایل دقیق گمشدگی معمولاً ناشناخته‌اند (روبین^۲، ۱۹۸۷؛ فیتز موریس، داویدیان، وربک و مولنبرگ^۳، ۲۰۰۸).

امروزه بسیاری از پژوهشگران روش‌هایی را برای مقابله با مشکل داده‌های گمشده ارائه کرده‌اند که بیشتر این روش‌ها به‌نوعی مکانیسم گمشدگی است؛ لذا باید نوع مکانیسم گم شدن معلوم باشد تا بتوان روش تحلیل مناسبی را انتخاب کرد. متأسفانه برخی از محققان بین الگوی گمشدگی^۴ و مکانیسم گمشدگی^۵ تمایزی قائل نمی‌شوند و این دو تعریف را به جای هم استفاده می‌کنند. یک الگوی گمشدگی نشان‌دهنده ترکیب داده‌های گمشده و مشاهده شده در مجموعه داده‌ها است، در حالی که مکانیسم گمشدگی، رابطه ممکن بین مقادیر گمشده و مشاهده شده را توضیح می‌دهد.

الگوی گمشدگی تنها مکان گمشدگی در داده‌ها را به ما نشان می‌دهد: الف- الگوی تک‌متغیره^۶؛ که مقادیر گمشده در یک متغیر رخ داده است. ب- الگوی گمشدگی

1. Pohl, Grafe & Rose

2. Rubin

3. Fitzmaurice, Davidian, Verbeke & Molenberghs

4. Missing Pattern

5. Missing Mechanism

6. Univariate pattern

بدون پاسخ واحد^۱؛ که این الگو بیشتر در تحقیقات پیمایشی^۲ رخ می‌دهد. ج- الگوی گمشدگی یکنواخت^۳؛ معمولاً در مطالعات طولی^۴ رخ می‌دهد که فرد شرکت‌کننده از مطالعه خارج می‌شود و دیگر به مطالعه برنمی‌گردد. د- الگوی گمشدگی داده عمومی (کلی)؛ که مقادیر گمشده در سرتاسر مجموعه داده‌ها به روش تصادفی پخش شده‌اند. ه- الگوی گمشدگی با برنامه^۵؛ برای گردآوری تعداد زیاد پرسشنامه‌ها با سؤال‌های زیاد مورد استفاده قرار می‌گیرد که به‌طور هم‌زمان هزینه پاسخ‌دهنده‌ها را نیز کاهش می‌دهد. و- الگوی متغیر پنهان^۶؛ فقط در تحلیل متغیرهای پنهان، مانند مدل‌های معادلات ساختاری^۷ دیده می‌شود (اندرس^۸، ۲۰۱۰).

روبین در سال ۱۹۷۶ سه مکانیسم مهم برای گمشدگی داده‌ها معرفی کرد (روبین، ۱۹۸۷). وقتی گم‌شدن هیچ ارتباطی با داده‌ها نداشته باشد، گمشدگی کاملاً تصادفی^۹ (MCAR) رخ می‌دهد. چنانچه گم‌شدن تنها به داده‌های مشاهده شده بستگی داشته باشد ولی به داده‌های مشاهده نشده وابسته نباشد، گمشدگی تصادفی^{۱۰} (MAR) نامیده می‌شود و در حالتی که مکانیسم گم‌شدن به داده‌های مشاهده نشده نیز وابسته باشد، گمشدگی غیر تصادفی^{۱۱} (MNAR) است (ابراهیم و مولنبرگز^{۱۲}، ۲۰۰۹). همچنین پاسخ‌های گمشده به دو گروه قابل اغماض^{۱۳} و غیر قابل اغماض^{۱۴} تقسیم می‌شوند. به‌طوری که روبین (۲۰۰۴) بیان کرد، زمانی که سوگیری در پاسخ وجود دارد، بدون پاسخ‌ها غیر قابل اغماض هستند (ژانگ، ۲۰۱۳).

1. Unit Non Response Pattern

2. Survey Research

3. Monotone Pattern

4. Longitudinal Study

5. Planned Missing Data Pattern

6. Latent Variable Pattern

7. Structural Equation models

8. Enders

9. Missing completely at random

10. Missing at random

11. Missing not at random

12. Ibrahim & Molenberghs

13. Ignorable

14. Nonignorable

سنجش‌های مقیاس بزرگ آموزشی، مانند برنامه سنجش بین‌المللی دانشجویان^۱ (PISA)، سنجش ملی پیشرفت تحصیلی^۲ (NAEP)، مطالعه بین‌المللی روندها در علوم و ریاضیات^۳ (TIMSS)، پیشرفت در مطالعه سواد خواندن^۴ (PIRLS) و مطالعه پنل آموزش ملی^۵ (NEPS)؛ هدفشان اندازه‌گیری دقیق شایستگی‌ها مانند ادراک خواندن و سواد ریاضی است. داده‌های آزمون‌های شایستگی معمولاً دارای پاسخ‌های گمشده در سؤال‌های آزمون هستند (پل و همکاران، ۲۰۱۴). مقدار پاسخ‌های گمشده در سنجش‌های مقیاس بزرگ، ناچیز و بی‌اهمیت نیست؛ برای نمونه در PISA (2006)، در همه کشورهای و در سه بعد (ریاضیات، خواندن و علوم) به‌طور متوسط ۱۰ درصد سؤال‌ها حذف شدند^۶ (در آلمان ۳۷٪) و ۴٪ به پاسخ نرسیدند^۷ (در آلمان ۱۵٪) (سازمان توسعه و همکاری اقتصادی^۸، ۲۰۰۹). برای ریاضیات و علوم در تیمز ۲۰۰۳ به‌طور متوسط ۳۷۳٪ سؤال‌ها در پایه هشتم و ۵۹۶٪ در پایه چهارم به پاسخ نرسیدند (مولیس، مارتین و دیاکن^۹، ۲۰۰۴). در آزمون ریاضی سنجش ملی پیشرفت تحصیلی (NAEP) ۱۹۹۰، میزان به پاسخ نرسیده‌ها بیشتر از حذف شده‌ها بود. در پایه ۱۲، ۱۳ تا از ۱۴۴ سؤال اجرا شده نرخ حذف بیشتر ۱۰ درصد داشتند و نرخ گمشدگی در به پاسخ نرسیده‌ها بیشتر از ۱۵ درصد بود (پل و همکاران، ۲۰۱۴).

تعداد زیاد مقادیر گمشده، تهدیدی برای اعتبار استنباط‌ها و نتایج هستند؛ برای مثال، استنباطی که از داده‌های ناقص حاصل می‌شود ممکن است توانایی فرد را از آنچه در داده‌های کامل به دست می‌آید منحرف کند (هلمن و گلس^{۱۰}، ۲۰۰۵؛ گلس و پیمنتل^{۱۱}، ۲۰۰۸؛ رز، ون داویر و اکسو^{۱۲}، ۲۰۱۰؛ رز، ۲۰۱۳). اگرچه بهترین روش

1. Program for International Student Assessment

2. National Assessment of Educational Progress

3. Trends in International Mathematics and Science Study

4. Progress in International Reading Literacy Study

5. National Educational Panel Study

6. Omitted Items

7. Not-reached Items

8. Organisation for Economic Co-operation and Development [OECD]

9. Mullis, Martin & Diaconu

10. Holman & Glas

11. Glas & Pimentel

12. Rose, von Davier & Xu

برای مدیریت داده‌های گمشده این است که با استفاده از طراحی مناسب مطالعه، وسایل اندازه‌گیری و دستورالعمل‌ها، مانع رخ دادن آنها شویم (مکنایت، مکنایت و فیگوردو^۱، ۲۰۰۷)، ایجاد روش‌های جایگزین برای سؤال‌های بدون پاسخ، به‌منظور به دست آوردن نتایج پایا در آزمون روان‌شناسی و آموزشی حیاتی است. اگر نتایج آزمون برای تصمیم‌گیری‌های مهم و گسترده^۲ به‌کار گرفته شود توجه به این موضوع لازم و ضروری است.

چندین راهبرد به‌طور رایج در برخورد با داده‌های گمشده در موقعیت آزمون^۳ استفاده می‌شوند. اینها شامل نمره‌گذاری بدون پاسخ‌ها به‌صورت صفر، برخورد با بدون پاسخ‌ها به‌عنوان گمشده و برآورد مهارت فرد تنها با سؤال‌هایی که پاسخ آنها در دسترس است یا ترکیبی از هر دو. انتخاب بهترین راهبرد برای رسیدگی به داده‌های گمشده باید بر دلایل گمشدگی و هدف خود آزمون مبتنی باشد. خاطرنشان می‌شود روشی برای برخورد با داده‌های گمشده وجود ندارد که در همه موارد برتر^۴ باشد (میسلوی و وو^۵، ۱۹۹۶).

مطالعه داده‌های گمشده سهم معنی‌داری در رویکرد کل‌نگر به سنجش مقیاس بزرگ به عهده دارند؛ چراکه راهبرد خاصی که برای نمره‌گذاری چنین سؤال‌هایی انتخاب می‌شود تأثیر قابل توجهی بر تفسیر نتایج دارد و این اثر در سنجش فردی یا گروهی مشهود است. از آنجایی که آزمون سراسری ورود به دانشگاه‌ها از جمله آزمون‌های مقیاس بزرگ آموزشی است و تاکنون مطالعه منتشرشده‌ای در خصوص داده‌های گمشده در این آزمون صورت نگرفته است؛ به‌منظور دستیابی به دانش اولیه از وضعیت داده‌های گمشده و علت وقوع آنها هدف از اجرای این پژوهش این است که با تحلیل توصیفی داده‌ها، الف- وضعیت فعلی داده‌های گمشده در آزمون سراسری بررسی شود؛ ۲- بررسی شود که آیا بین تعداد پاسخ صحیح و گمشده رابطه‌ای وجود دارد؟ ۳- آیا میزان پاسخ‌های گمشده تأثیری در شاخص‌های روان‌سنجی سؤال دارد؟ ۴- آیا داده‌های گمشده، قابل اغماض یا غیر قابل اغماض هستند؟

1. McKnight, McKnight & Figueredo

2. Far-reaching decisions

3. Testing situations

4. Preferable

5. Mislevy & Wu

داده‌ها و روش

در این پژوهش از داده‌های آزمون سراسری ورود به دانشگاه‌ها در سال‌های ۸۳، ۹۲ و ۹۵ استفاده شد. در این مجموعه، درس‌های عمومی (ادبیات، معارف اسلامی، زبان انگلیسی و عربی) و درس‌های تخصصی رشته‌های اصلی: علوم انسانی (ادبیات تخصصی، عربی اختصاصی، تاریخ و جغرافیا، علوم اجتماعی، منطق و فلسفه، روانشناسی، اقتصاد و ریاضیات)، علوم تجربی (ریاضیات، فیزیک، شیمی، زمین‌شناسی و زیست‌شناسی) و ریاضی و فیزیک (ریاضیات، فیزیک و شیمی) ارزیابی می‌شود. این آزمون چهارگزینه‌ای است که به پاسخ‌های غلط نمره منفی تعلق می‌گیرد و راهبردهای آزمون، پاسخ ندادن به سؤال‌هایی را که داوطلب از پاسخ درست آنها اطمینان ندارد، تشویق می‌کند.

جدول (۱) عنوان درس‌ها و تعداد سؤال‌ها

رشته تحصیلی	عنوان درس‌ها	تعداد سؤال‌ها
علوم انسانی	ادبیات عمومی	۲۵
	ادبیات فارسی تخصصی	۳۰
علوم تجربی	ادبیات عمومی	۲۵
	زیست‌شناسی	۵۰
ریاضی و فیزیک	ادبیات عمومی	۲۵
	ریاضیات	۵۵

نمونه‌ای به حجم ۱۰ هزار نفر به صورت تصادفی از داوطلبان هر سه رشته اصلی (علوم انسانی، ریاضی و علوم تجربی) در مطالعه شرکت داده شد. به منظور امکان مقایسه بین درس‌های عمومی و اختصاصی، در تمام رشته‌ها درس ادبیات عمومی انتخاب شد. از درس اختصاصی نیز در رشته علوم انسانی (ادبیات فارسی)، علوم تجربی (زیست‌شناسی) و ریاضی و فیزیک (درس ریاضیات) تعیین شد.

داده‌ها با استفاده از روش‌های توصیفی (میانگین و درصد)، روند پاسخ‌های صحیح، غلط و گم‌شده؛ همبستگی و ضرایب دشواری سؤال‌ها در هر درس؛ با استفاده از نرم‌افزار spss و R تحلیل شد.

یافته‌ها

در مفهوم کلی، داده گمشده به هر پاسخ خالی در مجموعه کلی سؤال‌های در دسترس آزمودنی متناسب می‌شود. مشکل عمده در مورد کسانی است که از سؤال می‌گذرند^۱ یا به دلیل محدودیت زمانی نمی‌توانند سؤال را پاسخ دهند (لادلو و آلری، ۱۹۹۹). چون الگو و مکانیسم گمشدگی در آزمون‌های سراسری نامشخص است لازم است ابتدا با مطالعه‌ای توصیفی، الگو و مکانیسم گمشدگی مشخص شود تا بر اساس آن بتوان راهبرد مناسبی برای برخورد با این نوع داده‌ها در نظر گرفت.

جدول (۲) تعداد و درصد داوطلبان هر رشته را در سه سال ۸۳، ۹۲ و ۹۵ به تفکیک جنس نشان می‌دهد. درصد داوطلبان زن در تمام رشته‌ها در هر سه سال به میزان قابل توجهی بالاتر از داوطلبان مرد است. با این وجود، در هر سه رشته درصد داوطلبان زن شرکت‌کننده در آزمون روند کاهشی را نشان می‌دهد. برای مثال، درصد داوطلبان زن رشته انسانی در سال ۸۳ (۶۷/۴)، سال ۹۲ (۶۳/۷) و سال ۹۵ (۶۳/۵) است. این روند در دو رشته ریاضی و فیزیک و علوم تجربی نیز برقرار است.

جدول (۲) توزیع فراوانی داوطلبان به تفکیک رشته، سال و جنسیت

رشته	جنسیت	۸۳		۹۲		۹۵	
		فراوانی	درصد	فراوانی	درصد	فراوانی	درصد
انسانی	♂	۶۷۴۰	۶۷/۴	۶۳۶۷	۶۳/۷	۶۳۵۳	۶۳/۵
	♀	۳۲۶۰	۳۲/۶	۳۶۳۳	۳۶/۳	۳۶۴۷	۳۶/۵
ریاضی و فیزیک	♂	۴۶۳۴	۴۳/۶	۳۹۱۶	۳۹/۲	۳۷۲۸	۳۷/۳
	♀	۵۳۶۶	۵۳/۷	۶۰۸۴	۶۰/۸	۶۲۷۲	۶۲/۷
تجربی	♂	۷۱۴۸	۷۱/۵	۶۶۹۸	۶۷	۶۵۴۹	۶۵/۵
	♀	۲۸۵۲	۲۸/۵	۳۳۰۲	۳۳	۳۴۵۱	۳۴/۵

^۱. Skipped

نمودارهای ۱ و ۲، گویای آن است که درصد پاسخ گمشده از سال ۸۳ تا ۹۵ در درس‌های عمومی و تخصصی رو به افزایش و درصد پاسخ صحیح روند کاهشی دارد. به طوری که در درس ادبیات عمومی رشته انسانی (سال ۸۳) درصد پاسخ صحیح ۱۶ سؤال بیشتر از درصد پاسخ گمشده آنهاست، اما در سال ۹۲ و ۹۵ تنها ۴ سؤال این شرایط را دارند. در آزمون ادبیات عمومی رشته انسانی، کمترین و بیشترین پاسخ گمشده به ترتیب (سال ۸۳) سؤال ۲۳ (۰/۷٪) و سؤال ۲۲ (۵۳/۴٪)؛ (سال ۹۲) سؤال ۱۱ (۲۱/۲٪) و سؤال ۱۳ (۷۳/۶٪) و (سال ۹۵) سؤال ۱ (۲۶/۳٪) و سؤال ۸ (۰/۷۸٪) مشاهده شده. در رشته تجربی، کمترین و بیشترین پاسخ گمشده ادبیات عمومی به ترتیب (سال ۸۳) سؤال ۲۵ (۳/۵٪) و سؤال ۱۵ (۶۵/۷٪)، (سال ۹۲) سؤال ۳ (۱۳/۸٪) و سؤال ۱۶ (۸۱/۱٪) و (سال ۹۵) سؤال ۱۷ (۰/۱۸٪) و سؤال ۱۴ (۸۳٪) مشاهده شده. در رشته ریاضی و فیزیک کمترین و بیشترین پاسخ گمشده ادبیات عمومی به ترتیب (سال ۸۳) سؤال ۳ (۲/۲٪) و سؤال ۱۶ (۵۴/۳٪)، (سال ۹۲) سؤال ۱۷ (۱۵/۸٪) و سؤال ۱۱ (۸۵٪) و (سال ۹۵) سؤال ۲۴ (۲۰/۵٪) و سؤال ۱۴ (۹۱/۶٪) دیده شد.

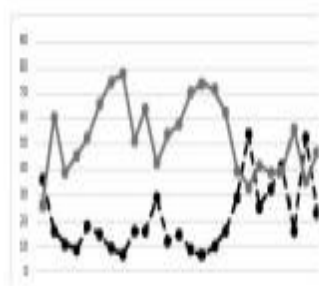
به منظور مقایسه بین سال‌های مختلف، میانگین پاسخ‌های صحیح و گمشده محاسبه شد (نمودار ۳). میانگین‌های محاسبه شده نتایج مشابه نمودار ۱ و ۲ (درصد پاسخ‌های صحیح و گمشده) را نشان داد (افزایش میانگین پاسخ‌های گمشده و کاهش میانگین پاسخ‌های صحیح از سال ۸۳ تا ۹۵). برای مثال، میانگین پاسخ‌های گمشده ادبیات عمومی رشته ریاضی و فیزیک سال ۸۳ (۵/۱)، سال ۹۲ (۱۲/۲) و سال ۹۵ (۱۳/۶)؛ و میانگین پاسخ‌های گمشده ریاضیات در رشته ریاضی و فیزیک سال ۸۳ (۳۲)، سال ۹۲ (۴۳/۷) و سال ۹۵ (۴۶).

در برخی مطالعات، علل پاسخ‌های گمشده بررسی شده و غالباً ویژگی‌های سؤال مانند دشواری یا فرمت پاسخ مدنظر بوده است. یافته‌های این مطالعات نسبتاً یکسان است. غالباً اولین سؤالی که آزمودنی‌ها را از ادامه پاسخگویی باز می‌دارد به احتمال بیشتر سؤال‌های باز پاسخ هستند تا چندگزینه‌ای (کورتز، لویس، اسکوس-کاگس و بارسین، ۱۹۹۳). به طور کلی، نتایج چندین مطالعه نشان می‌دهد که دشواری بیشتر

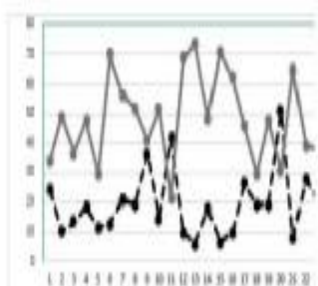
1. Koretz, Lewis, Skewes-Cox & Burstein

سؤال‌ها، از قلم افتادن بیشتر را در پی دارد (کورتز و همکاران، ۱۹۹۳؛ پل، هابركورن، هارت و ویگان^۱، ۲۰۱۲؛ رز و همکاران، ۲۰۱۰؛ ژانگ، ۲۰۱۳).
با توجه به اینکه سؤال‌های آزمون سراسری، تماماً سؤال‌های چهارگزینه‌ای هستند یافته‌های این مطالعه گویای این موضوع است که در گذشت زمان، دشواری سؤال‌ها در این آزمون افزایش یافته است.

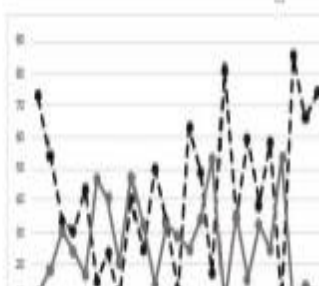
^۱. Pohl, Haberkorn, Hardt & Wiegand



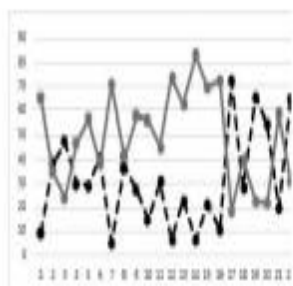
الف ۲



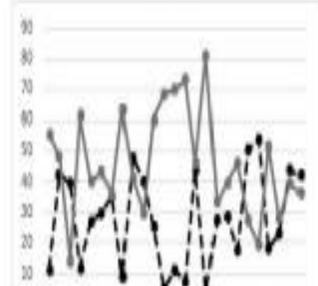
الف ۲



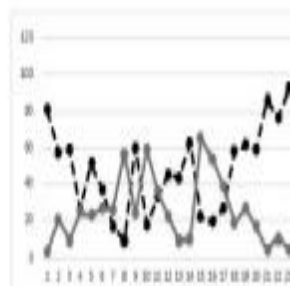
الف ۱



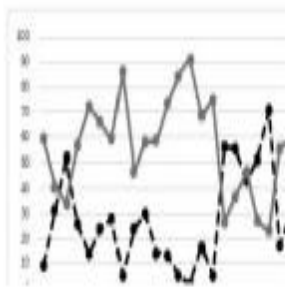
ب ۲



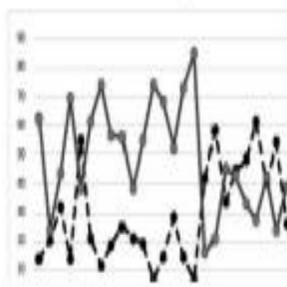
ب ۲



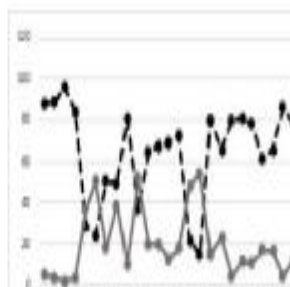
ب ۱



ج ۲

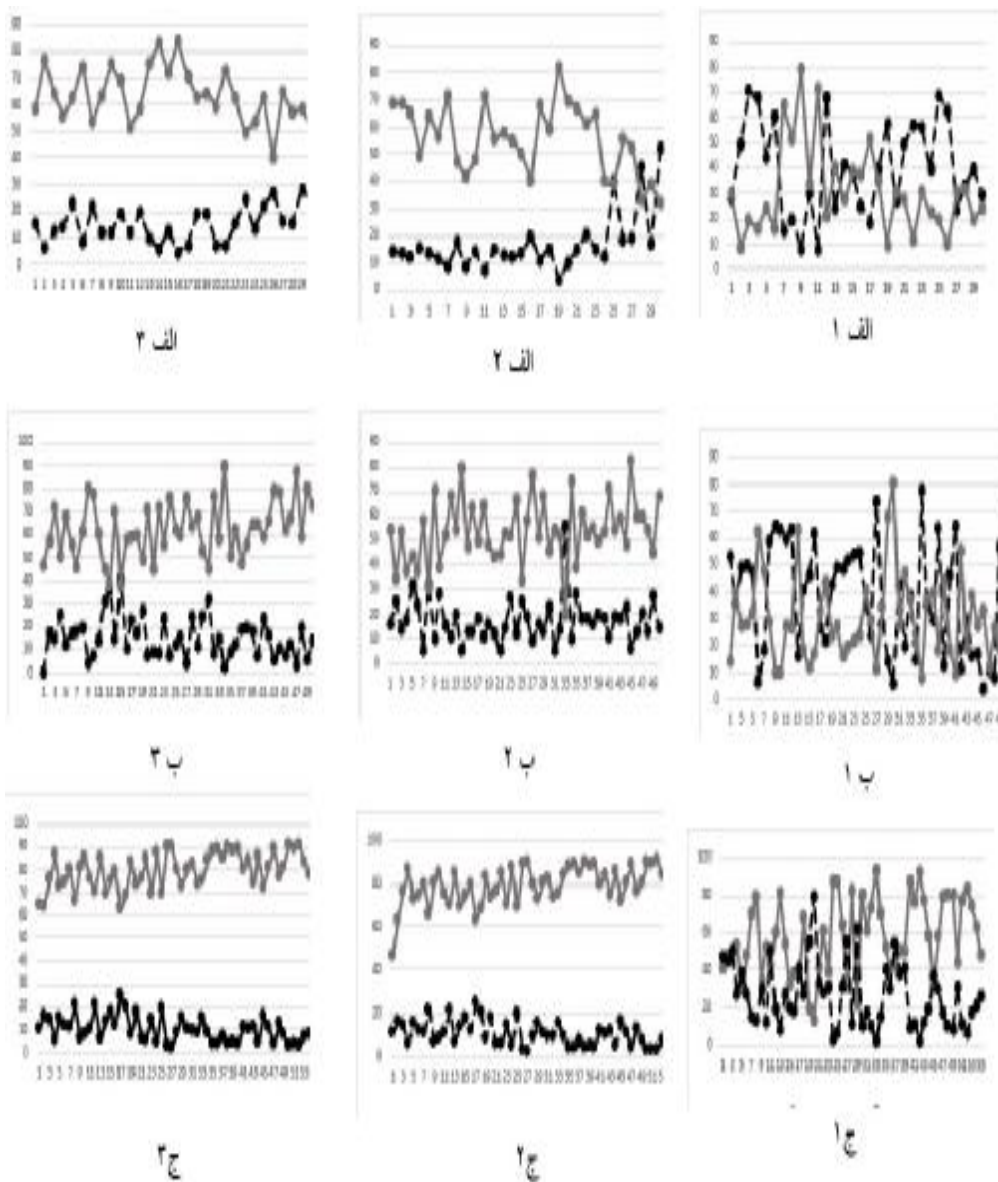


ج ۲

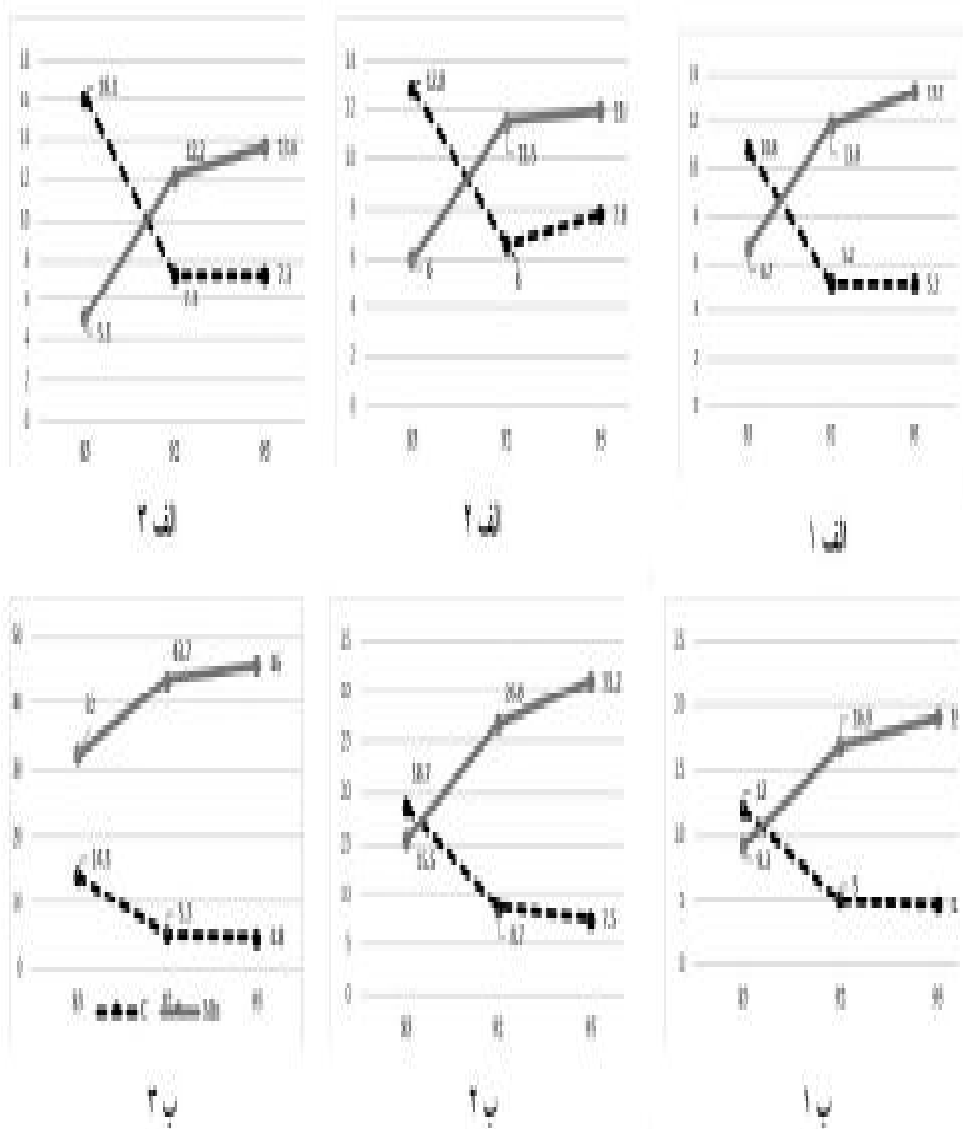


ج ۱

نمودار (۱) درصد پاسخ‌های صحیح و گمشته ادبیات عمومی در سال‌های مختلف هر رشته. الف ۱ (انسانی ۸۳)، الف ۲ (انسانی ۹۲)، الف ۳ (انسانی ۹۵) - ب ۱ (تجربی ۸۳)، ب ۲ (تجربی ۹۲)، ب ۳ (تجربی ۹۵) - ج ۱ (ریاضی ۸۳)، ج ۲ (ریاضی ۹۲)، ج ۳ (ریاضی ۹۵). (گمشته — صحیح —)



نمودار (۲) درصد پاسخ‌های صحیح و گمشده دروس تخصصی (ادبیات، زیست‌شناسی و ریاضی) در سال‌های مختلف هر رشته. الف ۱ (انسانی ۸۳)، الف ۲ (انسانی ۹۲)، الف ۳ (انسانی ۹۵) - ب ۱ (تجربی ۸۳)، ب ۲ (تجربی ۹۲)، ب ۳ (تجربی ۹۵) - ج ۱ (ریاضی ۸۳)، ج ۲ (ریاضی ۹۲)، ج ۳ (ریاضی ۹۵). (گمشده — صحیح —)



نمودار (۳) میانگین پاسخ‌های صحیح و گمشته ادبیات عمومی، ادبیات تخصصی، زیست‌شناسی، ریاضیات در سال‌های ۸۳، ۹۲ و ۹۵. الف ۱ (رشته انسانی)، الف ۲ (رشته تجربی)، الف ۳ (رشته ریاضی و فیزیک) - ب ۱ (ادبیات تخصصی، انسانی)، ب ۲ (زیست‌شناسی، تجربی)، ب ۳ (ریاضیات، ریاضی و فیزیک) (گمشته — صحیح —)

با توجه به اینکه پاسخ‌های گمشده در آزمون‌های مقیاس بزرگ آموزشی از جمله آزمون ورودی دانشگاه، از نوع پاسخ‌های حذف شده و پاسخ داده نشده به دلیل محدودیت زمانی است؛ جدول (۳) درصد این دو نوع پاسخ گمشده را در درس‌های انتخابی و سال‌های مختلف نشان می‌دهد. بر اساس داده‌های جدول (۳) در طی سال‌های متمادی در پاسخ‌های گمشده ناشی از حذف سؤال‌ها و محدودیت زمانی افزایش دیده می‌شود. برای مثال، در درس ادبیات عمومی رشته انسانی در سال‌های ۸۳، ۹۲ و ۹۵ درصد سؤال‌هایی که حداقل ۲۰ درصد پاسخ گمشده ناشی از حذف را داشته‌اند به ترتیب ۶۴ درصد، ۸۸ درصد و ۹۲ درصد است؛ و درصد سؤال‌هایی که حداقل نرخ ۵ درصد بدون پاسخ ناشی از محدودیت زمانی را نشان می‌دهند به ترتیب ۴ درصد، ۴۴ درصد و ۵۲ درصد است. این وضعیت در سایر درس‌ها نیز دیده می‌شود. همچنین با بررسی تمام درس‌ها و سؤال‌ها به تفکیک رشته و سال (به دلیل محدودیت فضا همه نمودارها نمایش داده نشده) مشخص شد بیشتر نرخ گمشدگی سؤال‌ها در آزمون ورودی دانشگاه ناشی از حذف سؤال‌هایی است که داوطلبان از پاسخ دادن به آنها گذشته‌اند.

جدول (۳) درصد سؤال‌های حذف شده و بدون پاسخ محدودیت زمانی در سال‌های مختلف

عنوان درس	شاخص	سال	۸۳	۹۲	۹۵
ادبیات عمومی رشته انسانی	سؤال‌های با حداقل نرخ ۲۰٪ حذف	سؤال‌های با حداقل نرخ ۵٪ محدودیت زمانی	٪۶۴	٪۸۸	٪۹۲
			٪۴	٪۴۴	٪۵۲
ادبیات عمومی رشته تجربی	سؤال‌های با حداقل نرخ ۲۰٪ حذف	سؤال‌های با حداقل نرخ ۵٪ محدودیت زمانی	۰	٪۸۴	٪۷۶
			۰	٪۱۶	٪۳۲
ادبیات عمومی رشته ریاضی و فیزیک	سؤال‌های با حداقل نرخ ۲۰٪ حذف	سؤال‌های با حداقل نرخ ۵٪ محدودیت زمانی	٪۲۸	۸۰	٪۸۴
			۰	٪۲۸	٪۲۸
ادبیات تخصصی رشته انسانی	سؤال‌های با حداقل نرخ ۲۰٪ حذف	سؤال‌های با حداقل نرخ ۵٪ محدودیت زمانی	٪۳۰	۸۶/۷٪	۹۳/۳٪
			۱۳/۳٪	۴۲/۳٪	٪۶۰
زیست‌شناسی رشته تجربی	سؤال‌های با حداقل نرخ ۲۰٪ حذف	سؤال‌های با حداقل نرخ ۵٪ محدودیت زمانی	٪۶۴	٪۹۶	٪۹۶
			٪۶	٪۳۶	٪۵۴
ریاضیات رشته ریاضی و فیزیک	سؤال‌های با حداقل نرخ ۲۰٪ حذف	سؤال‌های با حداقل نرخ ۵٪ محدودیت زمانی	۹۲/۷٪	۹۶/۴٪	۹۶/۴٪
			۶۷/۳٪	۶۹/۱٪	۶۹/۱٪

هنگامی که نسبت پاسخ‌های صحیح را محاسبه می‌کنیم نظرات متفاوتی در مورد مخرج کسر وجود دارد. در محاسبه ضریب دشواری برای هر سؤال عموماً از فرمول $P = \frac{N_r}{N} \times 100$ استفاده می‌شود (N_r افرادی هستند که پاسخ درست به سؤال داده‌اند و N همه افراد شرکت‌کننده در آزمون است). روش دیگر تقسیم کردن N_r به N_t (تعداد آزمودنی‌هایی که برای پاسخ تلاش کرده‌اند) است. N_t عمدتاً به موقعیت سؤال در آزمون و محدودیت زمانی آزمون بستگی دارد. مقدار N_t حتی اگر به‌طور رضایت بخشی برای سؤال‌های اولیه آزمون بزرگ باشد، ممکن است به‌طور قابل ملاحظه‌ای در سؤال‌های نزدیک به انتهای آزمون کاسته شود. از سوی دیگر، مقدار N_t گاهی اوقات ممکن است بسیار بزرگ باشد اگر، بعد از روبرو شدن با سؤال‌هایی کمی که

فراتر از توانایی فرد است، آزمودنی تنها به صورت سرسری پاسخ سؤال‌های باقی مانده را حدس بزند- چنین «تلاشی» در بهترین حالت تنها در نیمی از موارد درست خواهد بود و قطعاً از تلاش در قسمت اولیه و آسان‌تر آزمون متفاوت است (کنراد^۱، ۱۹۴۸؛ رز و همکاران، ۲۰۱۰).

در مورد آزمونی که سرعت عملکرد را اندازه می‌گیرد، استفاده از N پایه به p -value ای منجر خواهد شد که فراتر از دشواری ذاتی، موقعیت سؤال در آزمون را منعکس می‌کند. برای سؤال‌های چنین آزمونی استفاده از N_t ارجح است. از طرفی، در یک آزمون قدرت بدون محدودیت زمانی، N پایه ارجح است؛ اما در چنین آزمونی، باید N_t برابر یا بسیار نزدیک به N پایه باشد، به طوری که مزیت عملی N پایه نسبت به N_t عموماً کمتر خواهد شد. به طور کلی، به نظر می‌رسد استفاده از N_t نسبت به N پایه در تعیین سهولت یا دشواری سؤال- آزمون ارجح باشد. استفاده از N_t عموماً به P بالاتری منجر می‌شود (سؤال آسان‌تر از آنچه واقعاً هست نشان داده می‌شود)؛ در حالی که استفاده از N پایه عموماً گرایش دارد به اینکه P را بسیار کمتر محاسبه کند. تفاوت بیشتر بین N پایه و N_t ، احتمال خطا در P را بیشتر می‌کند.

در این مطالعه، هر دو نوع ضریب دشواری برای دو درس ادبیات عمومی رشته انسانی (جدول ۳) و زیست‌شناسی رشته تجربی (جدول ۴) محاسبه شده است ($P_1 = \frac{N_r}{N} \times 100$ و $P_2 = \frac{N_r}{N_t} \times 100$). یافته جدول‌های (۳ و ۴) و نمودار (۴) نشان می‌دهند دشواری سؤال‌ها به تفکیک و میانگین ضرایب دشواری در سه سال بررسی شده افزایش یافته است به عبارتی دشواری سؤال‌های آزمون در سال ۹۲ و ۹۵ اختلاف بسیاری با سال ۸۳ دارد. از سوی دیگر، داده‌ها نشان می‌دهند با استفاده از N_t در مجموع سؤال‌ها در هر سه سال آسان‌تر است و این به دلیل حضور داده‌های گمشده و پاسخ ندادن آزمودنی‌ها به سؤال‌هاست.

همچنین بر اساس نتایج داده‌های جدول‌های (۴ و ۵) حضور داده‌های گمشده و استفاده از N یا N_t در تعیین آسان‌ترین و دشوارترین سؤال آزمون تأثیرگذار بوده است به طوری که آسان‌ترین و دشوارترین سؤال ادبیات عمومی در سال ۸۳ استفاده از فرمول رایج ضریب دشواری به ترتیب (سؤال‌های ۲۳ و ۲۲) و مبتنی بر فرمول ضریب دشواری که N_t را در مخرج کسر دارد به ترتیب (سؤال‌های ۲۳ و ۹)؛ در سال

^۱. Conrad

۹۲ این ارقام عبارت‌اند از: به ترتیب سؤال‌های (۲۰ و ۱۳)؛ و بر اساس فرمول N_f ، به ترتیب (سؤال‌های ۲۰ و ۳)؛ و درنهایت در سال ۹۵ این ارقام عبارت‌اند از: به ترتیب سؤال‌های (۱۹ و ۱۵)؛ و بر اساس فرمول N_f ، به ترتیب (سؤال‌های ۱۹ و ۴) است. آسان‌ترین و دشوارترین سؤال زیست‌شناسی در سال ۸۳ با استاده از فرمول رایج ضریب دشواری به ترتیب (سؤال‌های ۴۱ و ۹؛ و ۵۰) و مبتنی بر فرمول ضریب دشواری که N_f را در مخرج کسر دارد به ترتیب (سؤال‌های ۱۲ و ۴۶)؛ در سال ۹۲ این ارقام عبارت‌اند از به ترتیب سؤال‌های (۳۳ و ۱۴)؛ و بر اساس فرمول N_f ، به ترتیب (سؤال‌های ۳۳ و ۲۱)؛ و درنهایت در سال ۹۵ این ارقام عبارت‌اند از به ترتیب سؤال‌های (۱۵ و ۱)؛ و بر اساس فرمول N_f ، به ترتیب (سؤال‌های ۱۵ و ۱).

در محاسبه ضریب دشواری، در کنار میزان سهولت یا دشواری سؤال، دو عامل محدودیت زمانی و آرایش سؤال‌ها در آزمون (در برخی از آزمون‌ها سؤال‌ها به ترتیب دشواری مرتب نمی‌شوند) در میزان این شاخص سؤال اثرگذار است. با توجه به اینکه در آزمون سراسری ورودی دانشگاه تمام سؤال‌های درس‌های مختلف، هم‌زمان در اختیار داوطلبان قرار می‌گیرد و سؤال‌های هر درس به ترتیب دشواری ارائه نمی‌شود. همان‌طور که در نتایج ارائه شده جدول‌های (۴ و ۵) و نمودار (۴) دیده می‌شود رخداد پاسخ‌های گم‌شده ضمن اینکه موجب افزایش اختلاف بین N پایه و N_f شده است، انحراف این آماره سؤال را به دنبال دارد. بنابراین استفاده از N_f نسبت به N پایه در تعیین سهولت یا دشواری سؤال‌های این آزمون ارجح است.

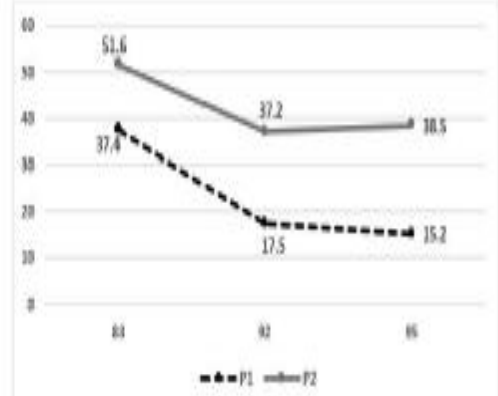
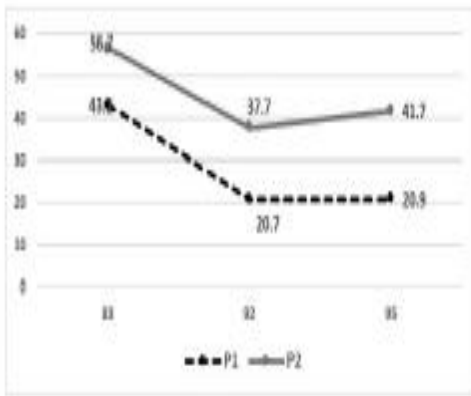
جدول (۴) ضریب دشواری سؤال‌های درس ادبیات عمومی رشته انسانی به تفکیک سال

سال	۸۳		۹۲		۹۵	
	P1	P2	P1	P2	P1	P2
۱	۷/۷۲	۱/۱/۴	۲/۴/۴	۶/۳/۵	۳/۵/۶	۴/۸/۳
۲	۷/۳/۵	۳/۶/۴	۱/۰/۲	۲/۰/۱	۶/۱	۴/۰/۳
۳	۷/۳/۳	۷/۸/۳	۸/۱/۵	۱/۵/۱	۱۱	۵/۷/۱
۴	۸/۰/۳	۶/۰/۴	۸/۷/۱	۳/۵/۳	۳/۷	۴/۵/۱
۵	۶/۳/۳	۱/۲/۵	۶/۱/۱	۵/۶/۱	۱/۷/۱	۳/۷/۳
۶	۵/۳/۱	۱/۷/۸	۲/۴/۱	۳/۳/۳	۸/۳/۱	۴/۲/۳
۷	۷/۳/۳	۴/۰	۵/۰/۰	۵/۶/۳	۵/۵	۱/۷/۸
۸	۸/۰	۳/۱/۱	۵/۰/۰	۳/۵/۳	۵/۵	۳/۷/۳
۹	۸/۰	۳/۱/۱	۵/۰/۰	۳/۵/۳	۵/۵	۳/۷/۳
۱۰	۶/۳/۱	۱/۷/۳	۳/۳/۱	۲/۵/۱	۱/۷/۱	۴/۴/۳
۱۱	۷/۵/۳	۶/۱/۵	۸/۴/۳	۵/۲/۵	۶/۷/۱	۳/۵/۳
۱۲	۴/۳/۳	۳/۵/۳	۷/۵	۲/۰/۳	۱/۴/۱	۸/۶/۲
۱۳	۸/۱/۱	۵/۶/۱	۲/۶	۲/۳/۱	۱/۳/۱	۲/۵/۱
۱۴	۸/۶	۵/۸/۷	۱/۷/۱	۳/۵/۳	۵/۷	۸/۵/۱
۱۵	۱/۵/۳	۵/۳/۸	۵/۶/۱	۲/۳/۱	۱/۶	۲/۳/۱
۱۶	۱/۷/۱	۳/۶/۳	۵/۵	۲/۶/۱	۳/۰/۱	۶/۳
۱۷	۱/۱/۷	۵/۸/۷	۲/۶	۳/۷/۳	۷/۵/۱	۴/۱/۴
۱۸	۶/۳/۳	۷/۱/۳	۲/۵/۱	۲/۷/۱	۲/۷/۱	۵/۶/۳
۱۹	۳/۵/۳	۱/۰/۸	۳/۵/۱	۸/۶/۳	۱/۳/۵	۵/۰/۷
۲۰	۵/۷/۳	۷/۱/۵	۱/۰/۵	۸/۷/۸	۵/۵/۲	۱/۴/۳
۲۱	۶/۷/۷	۶/۷/۸	۱/۷/۱	۲/۱	۳/۳/۳	۷/۱/۵
۲۲	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۲۳	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۲۴	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۲۵	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۲۶	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۲۷	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۲۸	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۲۹	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۳۰	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۳۱	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۳۲	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۳۳	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۳۴	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۳۵	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۳۶	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۳۷	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۳۸	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۳۹	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۴۰	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۴۱	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۴۲	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۴۳	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۴۴	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳
۴۵	۶/۷/۷	۵/۸/۱	۲/۱	۴/۴/۳	۴/۳/۳	۳/۷/۳

جدول (۵) ضریب دشواری سؤال‌های درس زیست‌شناسی رشته تجربی به تفکیک سال

سال	۸۳		۹۲	
	P1	P2	P1	P2
۱	۵/۳/۳	۱/۲/۱	۱/۶/۲	۳/۵/۳
۲	۶/۵/۳	۶/۵/۵	۲/۵/۱	۳/۷/۳
۳	۳/۵/۳	۳/۷/۱	۶/۳/۱	۳/۵/۳
۴	۵/۰	۶/۵/۶	۶/۱/۱	۲/۷/۳
۵	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۶	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۷	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۸	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۹	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۱۰	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۱۱	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۱۲	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۱۳	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۱۴	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۱۵	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۱۶	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۱۷	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۱۸	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۱۹	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۲۰	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۲۱	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۲۲	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۲۳	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۲۴	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۲۵	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۲۶	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۲۷	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۲۸	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۲۹	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۳۰	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۳۱	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۳۲	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۳۳	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۳۴	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۳۵	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۳۶	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۳۷	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۳۸	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۳۹	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۴۰	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۴۱	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۴۲	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۴۳	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۴۴	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵
۴۵	۴/۳	۳/۰/۸	۳/۳/۱	۵/۷/۵

۹۰		سال	۸۳		۹۲		۹۰	
P1	P2		P1	P2	P1	P2	P1	P2
۰	۰	۲۶	۲۳/۶	۶۱	۴۴/۳	۳۵/۱	۳۸/۲	
۷/۶۱	۳/۷۸	۲۱	۷۳/۸	۷/۷	۳/۷	۳/۳	۱/۷۱	
۷/۳۱	۵/۵۵	۷۱	۶/۷۲	۶/۶۱	۳/۵	۳/۱	۷/۲۱	
۲۵/۲	۷/۰۰	۲۱	۳/۳۱	۲/۳۳	۲/۳۱	۲/۰۳	۸/۵	
۱/۳۱	۲/۰۳	۳۰	۲/۰	۴/۱	۲/۱۳	۰/۳۱	۷/۲۰	
۶/۸۱	۷/۲۳	۱۸	۷/۲۳	۲/۰	۰/۱۱	۲/۱۳	۶/۷۰	
۷/۸۱	۱/۳۳	۳۲	۶/۲۱	۳/۶۱	۷/۳۱	۸/۷۱	۳/۵	
۲/۲۱	۵/۱۰	۳۳	۰/۷۱	۰/۱۶	۲/۰	۳/۲۱	۳/۰۳	
۷/۳	۲/۳۳	۳۵	۰/۰۱	۸/۲۱	۰/۱	۷/۲	۷/۷۱	
۱/۸	۱/۲۱	۳۸	۷/۸	۰/۷	۲/۷۱	۰/۶۳	۲/۷۱	
۶/۸۱	۲/۳۳	۶۸	۲/۰۳	۱/۶۰	۱/۸۳	۴/۱	۰/۳۳	
۷/۰۳	۰/۶۰	۸۱	۳/۰۳	۸/۸۱	۸/۵	۱/۲۱	۷/۶۳	
۳/۶۳	۸/۲۰	۷۸	۲/۳۶	۳/۶۱	۲/۶۳	۰/۲	۳/۰۳	
۰/۳۱	۲/۷۳	۲۸	۴/۱	۰/۲۱	۳/۷۳	۲/۸۱	۴/۱۰	
۳/۰۳	۳/۶۶	۰۳	۲/۶۳	۳/۶۰	۴/۲۱	۸/۸	۲/۲۱	
۲/۱۱	۳/۶۶	۱۳	۷/۳۶	۱/۱۸	۶/۰۱	۲/۳۱	۰/۰۰	
۳/۸۱	۶/۵۰	۲۳	۶/۱۱	۲/۰۲	۸/۷۱	۳/۶۱	۶/۷۳	
۰/۸۱	۴/۳۳	۴۳	۱/۰۳	۲/۸۳	۲/۷۱	۲/۰	۲/۰۲	
۳/۸۱	۱/۳۰	۳۳	۶/۱	۶/۵۱	۲/۶۳	۲/۰۱	۶/۲۳	
۶/۸	۲/۶۶	۰۳	۸/۱	۳/۶۱	۷/۰۲	۴/۸	۲/۲۱	
۰/۲	۴/۸۱	۶۳	۷/۲	۸/۰	۴/۶۳	۴/۱	۲/۵	
۷	۲/۷۲	۸۳	۲/۰۱	۳/۲۱	۸/۰۲	۶/۳	۶/۷۲	
۱/۳۱	۲/۰	۷۳	۰/۷	۰/۱۱	۱/۳۱	۰/۲	۱/۷۳	
۱/۸	۴/۲۱	۲۳	۷/۶۰	۱/۰۸	۱/۸۱	۰/۰	۷/۵	
۰/۲۱	۲/۳۳	۰۰	۳	۸/۲۱	۴/۳۱	۳/۳۱	۲/۲۳	



نمودار ۴: میانگین ضرایب دشواری زیست‌شناسی رشته تجربی (الف) و ادبیات عمومی رشته انسانی (ب). P1 (ضریب دشواری با فرمول رایج)، P2 (ضریب دشواری با فرمول N_E).

با توجه به اینکه نظام‌های آموزشی مختلف، راهبردهای آزمون گرفتن متفاوتی دارند که این راهبردها بر تعداد سؤال‌هایی که هر آزمودنی برای آن تلاش می‌کند تأثیر دارد؛ آزمون سراسری ورودی دانشگاه هم از این قانون مستثنا نیست. از آنجایی که این آزمون، آزمونی چهارگزینه‌ای است که به پاسخ‌های غلط نمره منفی اختصاص می‌یابد، بر اساس راهبردهای پیشنهاد شده از طرف نظام‌های آموزشی، داوطلبان به حذف پاسخ سؤال‌هایی که از پاسخ آن اطمینان ندارند یا پاسخ را نمی‌دانند تشویق می‌شوند؛ لذا آزمودنی می‌تواند با مدیریت آزمون از طریق فرایند خودانتخابی^۱ و حذف سؤال‌ها زمانی که به‌طور کامل از پاسخ درست مطمئن نبود؛ بالاترین $\hat{\theta}$ را که مایل است به دست آورد. حقیقت این است که آزمودنی که سال‌های خاصی را حذف می‌کند اطلاعات مهمی را دربردارد که او پاسخ این سؤال‌ها را نمی‌داند و شانس موفقیتش در هر سؤال تنها برابر C است و نمی‌توان این اطلاعات را نادیده گرفت. در برخورد با داده‌های گمشده، راهبردهای متفاوتی وجود دارد از جمله نمره‌گذاری بدون پاسخ‌ها به‌صورت غلط، برخورد با بدون پاسخ‌ها با نادیده گرفتن آنها، سؤال‌های اجرا نشده. انتخاب بهترین راهبرد برای رسیدگی به داده‌های گمشده باید مبتنی بر دلایل

^۱. Self-selection

گمشدگی و هدف خود آزمون باشد. باید به خاطر داشت که روشی برای برخورد با داده‌های گمشده وجود ندارد که در همه موارد برتر^۱ باشد (میسلوی و وو، ۱۹۹۶).

جدول (۶) الگوی پاسخ ۶ آزمودنی در درس ادبیات عمومی رشته انسانی سال ۹۲ را نشان می‌دهد. این افراد تعداد پاسخ درست یکسان و تعداد پاسخ غلط و گمشده متفاوت دارند. بررسی این جدول به سرعت نشان می‌دهد که چطور روش‌های نمره‌گذاری متفاوت داده‌های گمشده، بر برآورد توانایی‌شان اثر می‌گذارد. با استفاده از سه روش رایج نمره‌گذاری، نمره‌های این افراد برآورد شد (نمره ۱ = نمره داوطلب بر اساس نمره‌گذاری سازمان سنجش، در محاسبه نمره ۲ با پاسخ‌های گمشده به‌عنوان غلط برخورد شده است و در نمره ۳ پاسخ‌های گمشده ۱/۴ پاسخ صحیح تعلق گرفته است). همان‌طور که مشاهده می‌شود آزمودنی شماره ۱ که به تمام ۲۵ سؤال آزمون پاسخ داده با هر سه روش نمره‌گذاری، نمره‌ای پایین‌تر از سایر افراد کسب کرده است و آزمودنی شماره ۶ با بیشترین تعداد پاسخ گمشده در نمره‌گذاری اول و سوم، بالاترین برآورده نمره را نسبت به سایر افراد که پاسخ گمشده کمتری داشته‌اند به دست آورده است. تنها در روش دوم (نمره ۲) که با پاسخ‌های گمشده به‌عنوان غلط برخورد شده، تمام افراد برآورد توانایی یکسان دارند. بررسی جدول (۵) نشان می‌دهد افرادی که پاسخ گمشده بیشتری دارند با فرایند خودانتخابی و پاسخ ندادن به سؤال‌های بیشتر، آزمون را مدیریت کرده‌اند. اگر این شیوه به‌عنوان خط مشی نمره‌گذاری برای هر آزمون شناخته شود، نتیجه به‌احتمال بسیار زیاد این خواهد شد که در موقعیت آزمون، آزمودنی‌ها تشویق می‌شوند که نمره‌شان را با پاسخ به تعداد کمتری از سؤال‌ها به حداکثر برسانند؛ درحالی‌که بررسی عمیق‌تر داده‌ها نشان می‌دهد آزمودنی اول که برای پاسخ تمام سؤال‌ها تلاش کرده در حقیقت سؤال‌های دشوارتری را پاسخ داده است (میانگین ضرایب دشواری = $22/5$)؛ و میانگین ضرایب دشواری آزمودنی دوم، سوم، چهارم، پنجم و ششم با تعداد سؤال‌های گمشده بیشتر به ترتیب $23/8$ ؛ $24/6$ ؛ $25/6$ و $24/7$ است. بنابراین افراد با پاسخ به سؤال‌های ساده‌تر و بی‌پاسخی بیشتر بر اساس روش‌های متفاوت نمره‌گذاری (به‌استثنای روش دوم) نمره بالاتری کسب می‌کنند. این فرایند با توجه به نتایج جدول‌های (۴ و ۵) و نمودار (۴)،

1. Preferable

در افزایش ضریب دشواری سؤال‌ها نیز مؤثر بوده است. همچنین در افزایش تعداد پاسخ‌های گمشده مطابق آنچه در نمودارهای (۳) نمایش داده شد؛ مؤثر است.

جدول (۶) الگوی پاسخ ۶ نفر از افراد با نمره خام ۱۴ در درس ادبیات عمومی رشته انسانی

سال ۹۲

نمره ۳ (گمشده=۷۵/نمره صحیح)	نمره ۲ (گمشده=غلط)	نمره ۱	تعداد سؤالات پاسخ داده شده	تعداد پاسخ گمشده	تعداد پاسخ غلط	تعداد پاسخ صحیح	فرد
۸/۱۳	۸/۱۳	۳/۱۳	۵۲	۳۱	۰	۲۱	۱
۳/۷/۳	۸/۱۳	۳/۵/۳	۳۸	۳۱	۷	۰	۲
۳/۵	۸/۱۳	۷/۳	۰	۳	۵	۰	۳
۷/۷	۸/۱۳	۷/۱	۰	۳	۳	۰	۴
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۰
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۱
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۲
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۳
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۴
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۵
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۶
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۷
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۸
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۹
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۲۰
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۲۱
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۲۲
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۲۳
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۲۴
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۲۵
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۲۶
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۲۷
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۲۸
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۲۹
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۳۰
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۳۱
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۳۲
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۳۳
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۳۴
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۳۵
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۳۶
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۳۷
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۳۸
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۳۹
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۴۰
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۴۱
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۴۲
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۴۳
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۴۴
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۴۵
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۴۶
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۴۷
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۴۸
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۴۹
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵۰
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵۱
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵۲
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵۳
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵۴
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵۵
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵۶
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵۷
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵۸
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۵۹
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶۰
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶۱
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶۲
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶۳
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶۴
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶۵
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶۶
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶۷
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶۸
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۶۹
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷۰
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷۱
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷۲
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷۳
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷۴
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷۵
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷۶
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷۷
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷۸
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۷۹
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸۰
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸۱
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸۲
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸۳
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸۴
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸۵
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸۶
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸۷
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸۸
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۸۹
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹۰
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹۱
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹۲
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹۳
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹۴
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹۵
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹۶
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹۷
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹۸
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۹۹
۳/۱۳	۳/۱۳	۳/۱۳	۰	۳	۳	۰	۱۰۰

جدول‌های (۷ و ۸) رابطه داده‌های گمشده در ابعاد مختلف را نشان می‌دهد. ضرایب همبستگی بالا گویای این است که تمایل به پاسخ ندادن در حیطه‌های مختلف وابسته است، بدین معنی که افراد با تمایل بالاتر برای پاسخ ندادن به سؤال‌ها در یک بعد، همچنین تمایل بیشتری به پاسخ ندادن سؤال‌ها در ابعاد دیگر دارند. از آنجایی

که، ضریب همبستگی بالای $0/3$ ، با یک اثر متوسط در نظر گرفته می‌شود (کوهن^۱، ۱۹۸۸)؛ بنابراین از طرفی بالا بودن ضریب همبستگی نشان‌دهنده ثبات در تمایل به گمشدگی است و بیان می‌کند که گمشدگی ویژگی اختصاصی فرد است و کمتر به بعد شایستگی وابسته است. از سوی دیگر، تمایل به پاسخ ندادن در یک بعد می‌تواند به‌عنوان یک پیش‌بینی‌کننده نسبی این ویژگی در بعد دیگر شایستگی در نظر گرفته شود (کههر و همکاران، ۲۰۱۵). همچنین ضرایب همبستگی بالا در داده‌های گمشده بین درس‌های مختلف نشان‌دهنده ارتباط و همبستگی بین توانایی و دشواری آزمون است؛ به این معنی که افراد با توانایی پایین در درس‌های مختلف تمایل به بی‌پاسخی سؤال‌های بیشتری دارند. بر اساس داده‌های مشاهده شده، ضرایب همبستگی در دامنه ای از ۴۳-۷۸ در جدول ۷ و ۶۶-۶۱ در جدول ۸ قرار دارد.

در حالی که تمایل به پاسخ ندادن در ابعاد مختلف بالا و نسبتاً ثابت است؛ در بعد شایستگی یکسان [مثلاً، همبستگی بین ادبیات عمومی و اختصاصی رشته انسانی (۷۸-۶۷)]، این همبستگی نسبتاً بالاتر از ابعاد متفاوت است [همبستگی ادبیات عمومی با زیست‌شناسی (۶۷-۵۳) و ادبیات عمومی با ریاضی (۵۷-۴۳) است].

جدول (۷) ضرایب همبستگی بین داده‌های گمشده در ابعاد مختلف

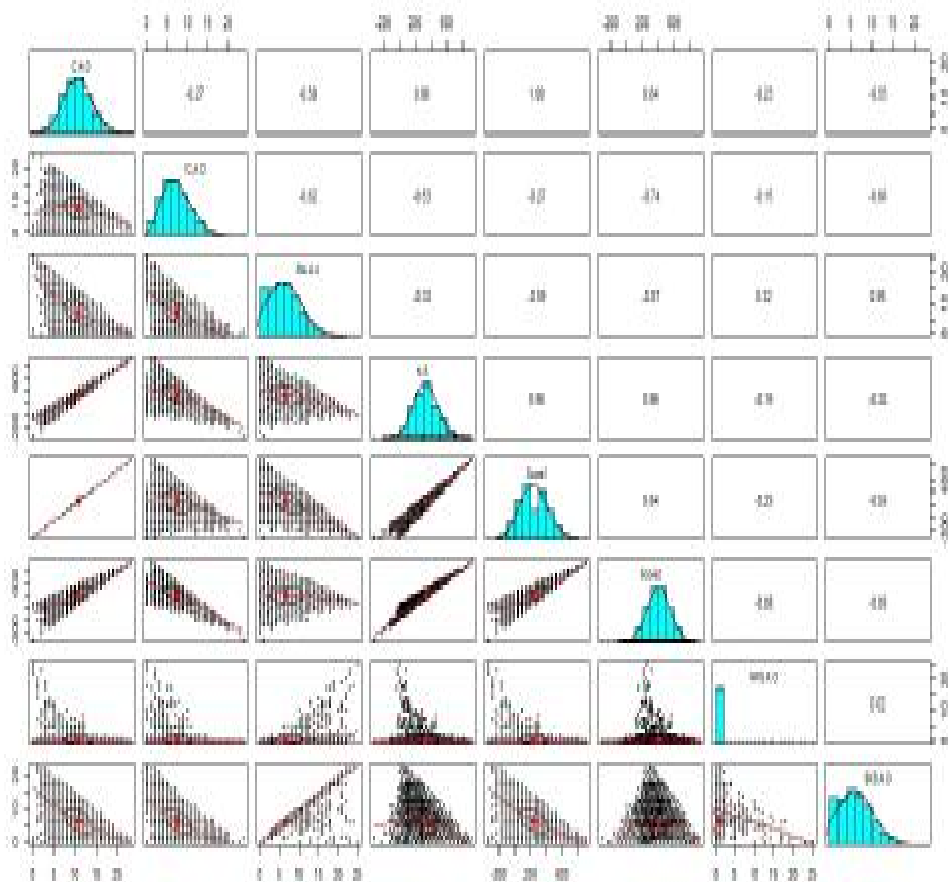
رشته	سال	ادبیات عمومی - ادبیات تخصصی	رشته	سال	ادبیات - ریاضی	رشته	سال	ادبیات - زیست‌شناسی
انسانی	۸۳	۰/۶۷	فنی	۸۳	۰/۴۳	فنی	۸۳	۰/۵۳
	۹۲	۰/۷۸		۹۲	۰/۵۷		۹۲	۰/۶۷
	۹۵	۰/۷۷		۹۵	۰/۵۶		۹۵	۰/۶۶

جدول (۸) ضرایب همبستگی بین داده‌های گمشده در ابعاد مختلف رشته تجربی سال ۹۵

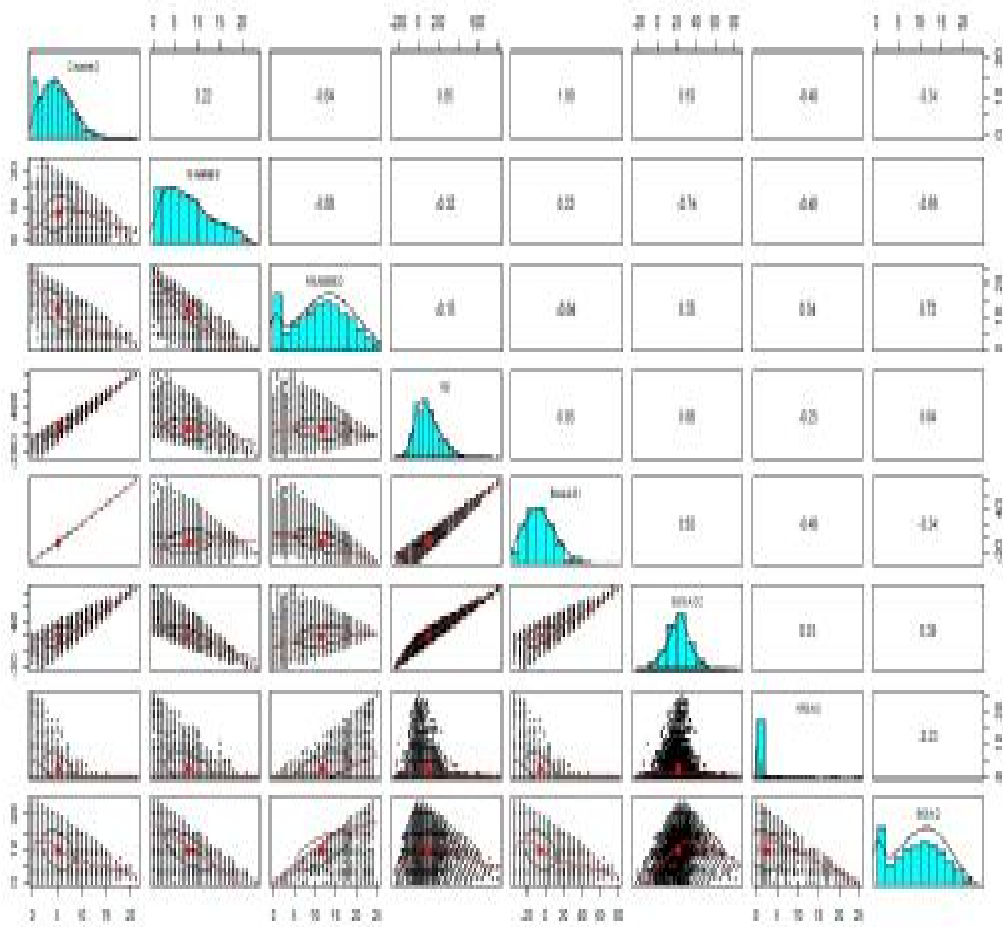
رشته	عنوان درس‌ها	ضریب همبستگی	عنوان درس‌ها	ضریب همبستگی
تجربی	ادبیات-عربی	۰/۶۲	عربی-زبان انگلیسی	۰/۵۳
	ادبیات-معارف اسلامی	۰/۶۴	زیست‌شناسی-ادبیات	۰/۶۶
	ادبیات-زبان انگلیسی	۰/۴۵	زیست‌شناسی-عربی	۰/۶۰
	معارف اسلامی-زبان انگلیسی	۰/۴۱	زیست‌شناسی-معارف اسلامی	۰/۵۹
	عربی-معارف اسلامی	۰/۵۴	زیست‌شناسی-زبان انگلیسی	۰/۴۹

با استفاده از نمودارهای ۱۰-۵ همبستگی بین نمره‌های درست، غلط، گمشده، حذف شده (SKS)، بدون پاسخ محدودیت زمانی (NRS) به صورت تصویری نمایش داده شد. ضرایب همبستگی بین نمره پاسخ‌های درست و گمشده در درس‌های مختلف و سال‌های متفاوت بالا و منفی است (در دامنه‌ای از ۰/۵۴- تا ۰/۸۵-)؛ یعنی هرچه تعداد پاسخ‌های درست، بیشتر می‌شود تعداد پاسخ‌های گمشده کمتر می‌شود و بالعکس. به عبارتی، ضرایب همبستگی بالا نشان از ارتباط پاسخ به سؤال‌های با توانایی آزمودنی‌ها دارد، آزمودنی‌هایی که توانایی بالاتری دارند پاسخ درست بیشتر و پاسخ گمشده کمتری دارند. افراد با تعداد پاسخ‌های گمشده بیشتر، پاسخ‌های غلط کمتری دارند که نشان از توانایی پایین فرد و مدیریت آزمون با فرایند خودانتخابی است (رز و همکاران، ۲۰۱۰؛ کالبرستون، ۲۰۱۱).

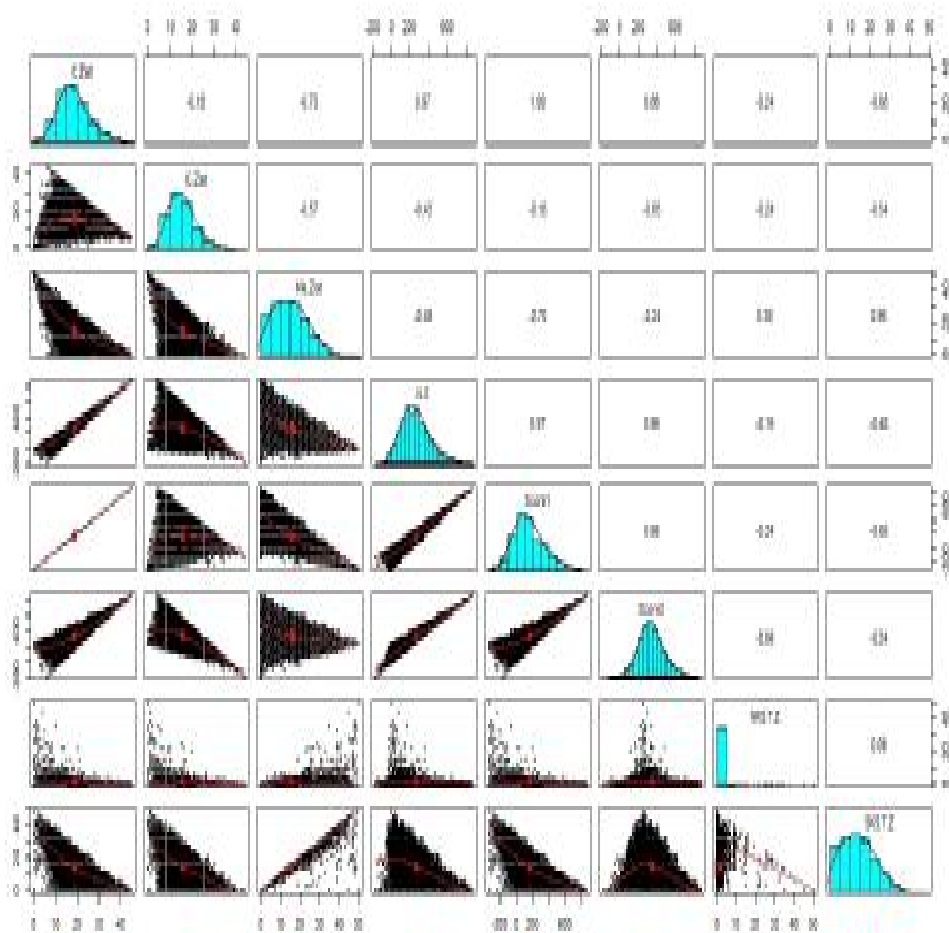
همچنین مقایسه برآورد نمرات نهایی داوطلبان با روش نمره‌گذاری سازمان سنجش (NS) گویای این است در سال‌های مختلف به موازات افزایش پاسخ گمشده، سطح و پراکندگی نمره‌های داوطلبان کاهش یافته و نمودارها از حالت تقریباً نرمال به نمودارهای چوله به راست تغییر کرده است. بنابراین این ارتباط و فرایند خودانتخابی سؤال‌ها توسط آزمودنی گویای این است که پاسخ گمشده در آزمون سراسری ورود به دانشگاه غیر قابل اغماض است.



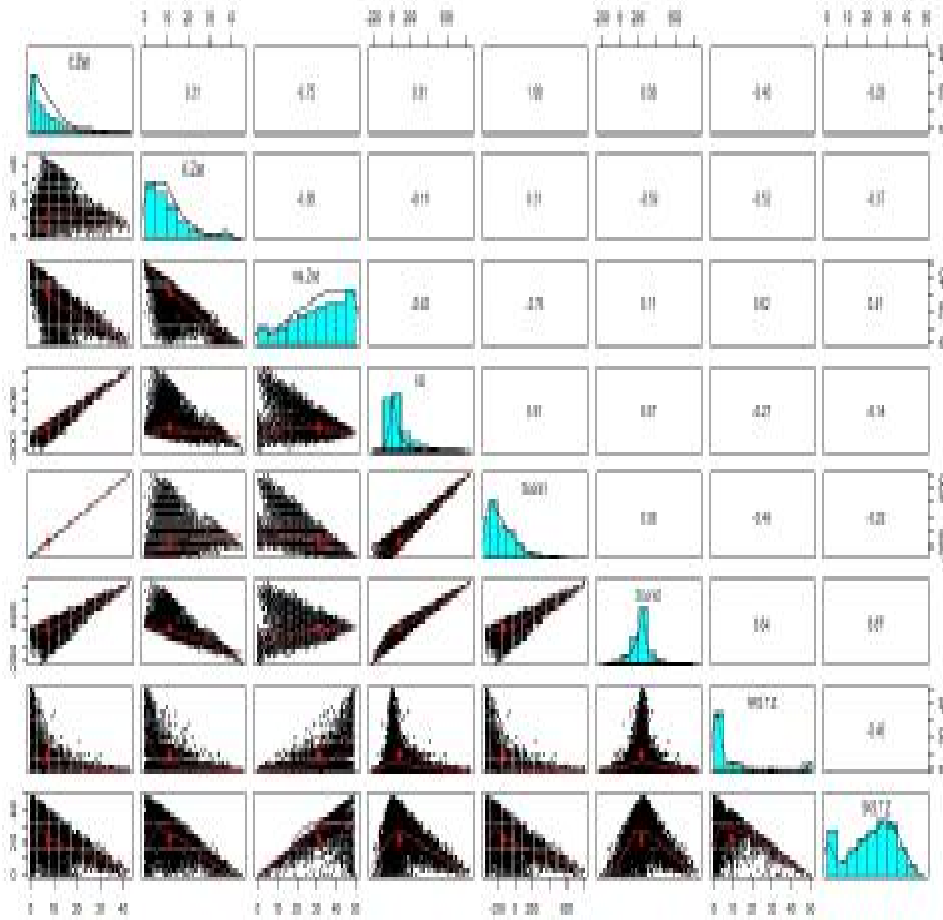
نمودار (۵) همبستگی میان پاسخ‌های مشاهده شده و نمره‌های حاصل از سه روش نمره‌گذاری
در درس ادبیات عمومی رشته انسانی سال ۸۳



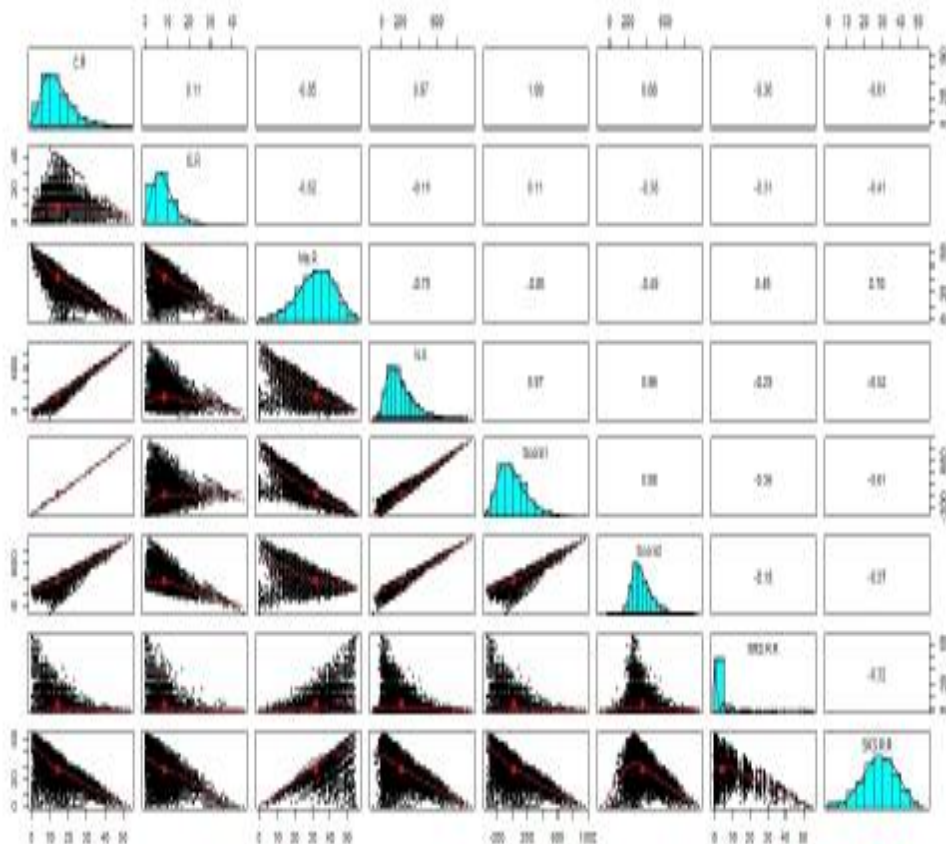
نمودار (۶) همبستگی میان پاسخ‌های مشاهده شده و نمره‌های حاصل از سه روش نمره‌گذاری در درس ادبیات عمومی رشته انسانی سال ۹۵



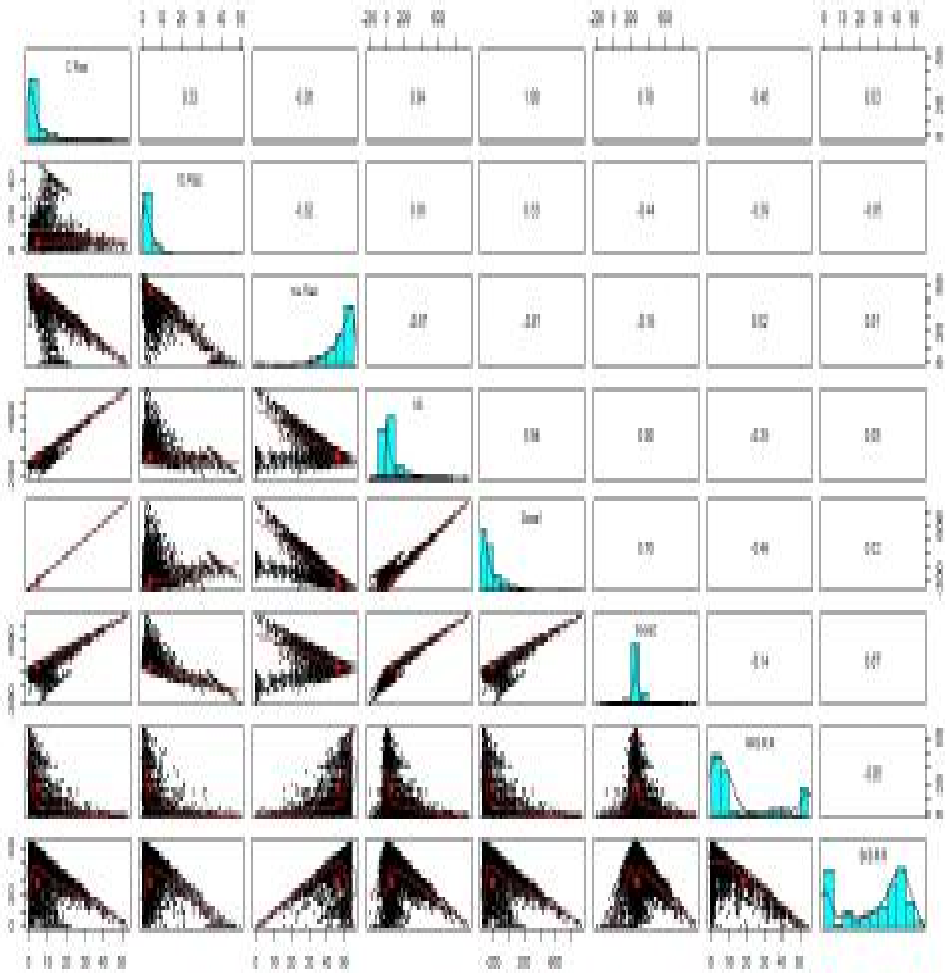
نمودار (۷) همبستگی میان پاسخ‌های مشاهده شده و نمره‌های حاصل از سه روش نمره‌گذاری در درس زیست‌شناسی رشته تجربی سال ۸۳



نمودار (۸) همبستگی میان پاسخ‌های مشاهده شده و نمره‌های حاصل از سه نمره‌گذاری در درس زیست‌شناسی رشته تجربی سال ۹۵



نمودار (۹) همبستگی میان پاسخ‌های مشاهده شده و نمره‌های حاصل از سه روش نمره‌گذاری در درس ریاضیات رشته ریاضی و فیزیک سال ۸۳



نمودار (۱۰) همبستگی میان پاسخ‌های مشاهده شده و نمره‌های حاصل از سه روش نمره‌گذاری در درس ریاضی رشته ریاضی و فیزیک سال ۹۵

بحث و نتیجه‌گیری

هر آزمون شامل مجموعه‌ای از سؤال‌هاست و هر سؤال آزمون در نوع خودش پیچیده است. افرادی که در آزمون شرکت می‌کنند دارای ویژگی‌های فردی مختلفی هستند.

حتی دوقلوهای همسان هم تفاوتی‌هایی را در دانش و مهارت‌هایشان نشان می‌دهند زیرا تجارب زندگی‌شان دقیقاً بعد از تولد همانند نیست. تعامل آزمودنی با سؤال‌های آزمون به مجموعه‌ای از پاسخ‌های صحیح، غلط و گم‌شده منجر می‌شود که فرایندهای بسیار پیچیده‌ای را نشان می‌دهند (ریکاس^۱، ۲۰۰۹).

مطالعات داده‌های گم‌شده سهم معنی‌داری در رویکرد کل‌نگر به سنجش‌های مقیاس‌بزرگ را به عهده دارند. در میان مطالعات مرتبط، اسچیت، وولف و کيفر^۲ (۱۹۹۳) توزیعی سه‌راهه (درست، غلط، حذف) با طرحی سه‌گوش^۳ را مورد بحث قرار دادند. سه‌گوشه مثلث ۱۰۰٪ حذف، ۱۰۰٪ غلط و ۱۰۰٪ پاسخ صحیح قرار می‌گرفت. آنها دریافتند که ممکن است آزمودنی‌ها به دلایل متفاوت به پاسخ دادن به سؤالات چندگزینه‌ای، تمایل داشته یا نداشته باشند.

در آزمون‌های آموزشی، پاسخ‌های گم‌شده در دو دسته طبقه‌بندی می‌شوند: پاسخ ندادن به سؤال به‌عنوان حذف شده یا پاسخ ندادن ناشی از محدودیت زمانی. راهبردهای متفاوت آموزشی، شکل سؤال‌های آزمون، دشواری سؤال‌ها، محدودیت زمانی و انگیزه آزمودنی‌ها از علل عمده وقوع پاسخ‌های گم‌شده در این آزمون‌ها هستند. به‌طوری که این علل بر گستره سؤال‌های بدون پاسخ تأثیرگذار است. از آنجایی که هدف از آزمون‌های آموزشی به‌ویژه برگزاری آزمون‌های حساس، ارزیابی و برآورد درست از توانایی افراد برای تصمیم‌گیری است، رخداد داده‌های گم‌شده مشکلات بسیاری در تحلیل و تفسیر داده‌ها ایجاد می‌کند. به‌طوری که کاهش حجم نمونه، از دست رفتن اطلاعات و کاهش دقت برآورد را به دنبال دارند؛ زیرا همه آزمودنی‌ها دارای تعداد یکسانی از اندازه‌گیری متغیرها نیستند و افراد با داده گم‌شده تفاوت‌های اساسی با افراد دارای داده‌های کامل دارند. با توجه به اینکه توافق مشخصی بین پژوهشگران در خصوص نمره‌گذاری و تحلیل این داده‌ها وجود ندارد لذا این موضوع نه‌تنها موجب سردرگمی ارزیابان در نحوه محاسبه پارامترهای سؤال و فرد می‌شود بلکه امکان مقایسه درست و شایسته بین داوطلبان را از بین می‌برد. همچنین در اغلب نرم‌افزارهای آماری امکان تحلیل داده‌های گم‌شده وجود ندارد و روش‌های آماری برای تحلیل داده‌های کامل برنامه‌ریزی شده‌اند. بنابراین در پژوهش

1. Reckase

2. Schmidt, Wolfe & Kifer

3. triangle

حاضر، داده‌های گمشده، توصیف و تأثیر آنها در برخی آماره‌ها در آزمون سراسری دانشگاه بررسی شده است.

یافته‌های مطالعه، روند کاهشی داوطلبان زن در سه سال تحت بررسی را نشان داد. همچنین بررسی داده‌ها گویای افزایش تعداد سؤال‌های دارای پاسخ‌های گمشده و میزان پاسخ‌های گمشده هر سؤال، از سال ۸۳ تا ۹۵ بود. به طوری که تعداد و میانگین پاسخ‌های درست، روند کاهشی و تعداد و میانگین پاسخ‌های گمشده، روندی افزایشی داشت. داده‌های جدول (۳)، نشان داد که بیشتر نرخ گمشدگی سؤال‌ها در این مطالعه به سؤال‌های حذف شده مربوط می‌شود. این نتایج گویای آن است که دشواری سؤال‌ها در سال‌های مختلف افزایش داشته است. بررسی متون و مطالعات پیشین نیز نشان‌دهنده آن بود که افزایش دشواری سؤال‌ها، گمشدگی بیشتر را به دنبال دارد (کورتز و همکاران، ۱۹۹۳؛ پل و همکاران، ۲۰۱۲؛ رز و همکاران، ۲۰۱۰؛ ژانگ، ۲۰۱۳).

پس از محاسبه ضریب دشواری سؤال‌ها با استفاده از دو روش رایج، نشان داده شد که افزایش پاسخ ندادن به سؤال‌ها توسط تعداد بسیاری از آزمودنی‌ها به متورم شدن شاخص دشواری سؤال منجر شده است و گویا سؤال‌ها سخت‌تر شده‌اند؛ درحالی‌که با استفاده از N_f در مخرج فرمول ضریب دشواری، سؤال‌های آزمون در هر سه سال از سهولت بیشتری برخوردار خواهند بود. بنابراین، با توجه به شرایط ویژه آزمون ورودی دانشگاه (ارائه هم‌زمان تمام سؤال‌ها در یک دفترچه و مرتب نشدن سؤال‌ها به ترتیب دشواری) و تفاوت زیاد تعداد افراد شرکت‌کننده در آزمون و تعداد افرادی که برای سؤال‌ها تلاش کرده‌اند؛ استفاده از N_f در محاسبه ضریب دشواری ارجح است. همچنین از آنجایی که اینکه راهبردهای نظام آموزشی بی‌پاسخی سؤال‌ها را تشویق می‌کند، پس ضمن اینکه هر آزمودنی آزمون ویژه خود را از سؤال‌های ارائه شده می‌سازد؛ تأثیر وقوع داده‌های گمشده در مخرج فرمول محاسبه دشواری، تورم و سوگیری مثبت برآورد شاخص دشواری به دلیل بی‌پاسخی را نشان می‌دهد.

در جدول‌های (۴ و ۵) و نمودارهای (۳ و ۴) افزایش پاسخ‌های گمشده، بیش برآورد پارامتر دشواری و سوگیری در برآورد پارامترهای فرد دیده می‌شود. بر اساس توضیحات بالا، شاید این افزایش و سوگیری کاذب و ناشی از دستورالعمل‌های نظام‌های آموزشی برای افزایش نمره آزمون و مدیریت آزمون با فرایند خودانتخابی توسط هریک از آزمودنی‌ها باشد. به طوری که بررسی الگوی پاسخ شش نفر از

آزمودنی‌ها با تعداد پاسخ صحیح یکسان نشان داد افراد با تعداد پاسخ گمشده بیشتر، برآورد توانایی بیشتری داشتند. درحالی‌که محاسبه میانگین دشواری سؤال‌های پاسخ داده شده، گویای این موضوع است که این افراد با حذف سؤال‌های دشوار و پاسخ به سؤال‌های آسان به نمره‌های بالاتر دست یافتند.

پاسخ زیرمجموعه متفاوتی از آزمودنی‌ها به هر سؤال، غیرمعرف بودن نمونه از نظر توزیع توانایی را به همراه دارد. این مسئله ضمن تأکید بر غیر قابل اغماض بودن داده‌های گمشده (رز، ۲۰۱۳)، در تحلیل پارامترهای فرد و سؤال، شواهدی مبنی بر عدم ارزیابی صحیح و عادلانه ارائه می‌کند (رز و همکاران، ۲۰۱۰).

نتایج جدول‌های (۷ و ۸) گویای وابستگی زیاد پاسخ‌های گمشده در ابعاد مختلف شایستگی است و بیان می‌کند ثبات در تمایل به گمشدگی ویژگی اختصاصی فرد است و کمتر به بعد شایستگی وابسته است. به طوری که می‌توان از میزان وقوع پاسخ‌های گمشده در یک درس، میزان پاسخ‌های گمشده درس دیگر را پیش‌بینی کرد. این نتایج با مطالعه کهلر و همکاران (۲۰۱۵) همسویی دارد. رز و همکاران (۲۰۱۰) در مطالعه داده‌های پیزا^۱ (۲۰۰۶) و کالبرستون (۲۰۱۱) نشان دادند که نسبت سؤال‌های درست و گمشده با هم مرتبط هستند و آزمودنی با تعداد پاسخ درست کمتر، تعداد پاسخ گمشده بیشتری دارد. یافته‌های مطالعه ما نیز همسو با مطالعه رز و کالبرستون نشان داد که بین تعداد سؤال‌های درست و گمشده، همبستگی بالا و منفی وجود دارد. همچنین افزایش تعداد پاسخ‌های گمشده به کاهش سطح نمره‌های داوطلبان و کاهش پراکندگی نمره‌های آنها منجر شده است.

همسو با مطالعه کهلر و همکاران (۲۰۱۵)، یافته‌های این مطالعه همبستگی بالای مقادیر گمشده در ابعاد مختلف و همبستگی بالا بین تعداد پاسخ‌های درست و گمشده را نشان داد؛ که گواه بر غیر قابل اغماض بودن داده‌های گمشده در آزمون سراسری است.

با توجه به اینکه داده‌های گمشده در آزمون سراسری از نوع داده‌های گمشده برنامه‌ریزی نشده و سیستماتیک است و احتمال حذف سؤال‌ها به دلیل دشواری و محدودیت زمانی وجود دارد، پس مکانیسم داده‌های گمشده در آزمون سراسری

۱. PISA

NMAR و غیر قابل اغماض بوده و کاهش حجم نمونه، دقت و کارایی برآوردها و معرف نبودن نمونه مورد بررسی را در برآورد شاخص‌ها و متغیرها به همراه دارد. بر اساس ادبیات پژوهش، رویکردهای مختلفی برای برخورد با داده‌های گمشده وجود دارد، اما دستورالعمل مشخصی برای اینکه کدام یک از آنها را استفاده کنیم وجود ندارد. پس انتخاب راهبرد مناسب برای مدیریت داده‌های گمشده، چالش مهمی را برای متخصصان اندازه‌گیری ایجاد کرده است (لادلوو و آلری، ۱۹۹۹). مطالعه تحقیقات گذشته نشان می‌دهد علی‌رغم اینکه برخورد با داده‌های گمشده به‌عنوان غلط یا نادیده گرفتن آنها، سوگیری در برآورد را به دنبال دارد (لرد، ۱۹۸۰؛ کالبرستون، ۲۰۱۱؛ رز، ۲۰۱۳؛ رز و همکاران، ۲۰۱۰)، همچنان استفاده از این رویکردها رایج است. بنابراین، با افزایش درصد پاسخ‌های گمشده در آزمون سراسری دانشگاه‌ها، در شرایط حساس آزمون^۱ (برای مثال، ارتقای درجه^۲ یا صدور گواهی نامه^۳) که پیامد راهبردهای مختلف نمره‌گذاری برای آزمودنی‌ها مهم است؛ ضمن اینکه لازم است در مطالعات بعدی، رویکردهای مختلف نمره‌گذاری داده‌های گمشده بررسی شود؛ همچنین می‌توان با افزایش مدت‌زمان آزمون از رخداد بالای داده‌های گمشده ممانعت کرد.

1. High-stakes testing

2. Grade promotion

3. Certification

منابع

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Conrad, H. (1948). *Characteristics and Uses of Item Analysis Data*. Psychological Monographs, General and Applied.
- Culbertson, M. (2011, April). Is it wrong? Handling missing responses in IRT. Speech presented at the Annual Meeting of the National Council on Measurement in Education. New Orleans, LA.
- Enders, C. K. (2010). *Applied Missing Data Analysis*. New York and London: Guilford Press.
- Fitzmaurice, G. M.; Davidian, M.; Verbeke, G. & Molenberghs, M. (2008). *Longitudinal data analysis*. London: Chapman & Hall.
- Frey, A.; Hartig, J. & Rupp, A. A. (2009). An NCME instructional module on booklet designs in large-scale assessments of student achievement: Theory and practice. *Educational Measurement: Issues and Practice*, 28 (3), 39–53. Doi: 10.1111/j.1745 - 3992.2009.00154.x
- Glas, C. A. W. & Pimentel, J. (2008). Modeling nonignorable missing data in speeded tests. *Educational and Psychological Measurement*, 68, 907-922.
- Graham, J. W.; Taylor, B.; Olchowski, A. & Cumsille, P. (2006). Planned missing data designs in psychological research. *Psychological Methods*, 11 (4), 323–343. Doi: 10.1037/1082-989X.11.4.323
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual review of psychology*, 60 (1), 549–576. Doi: 10.1146/annurev.psych.58.110405.085530
- Holman, R. & Glas, C. A. W. (2005). Modelling non-ignorable missing-data mechanisms with item response theory models. *British Journal of Mathematical & Statistical Psychology*, 58, 1-17.
- Ibrahim, J. G. & Molenberghs, M. (2009). Missing data methods in longitudinal studies: a review. *NIH*, 18, 1-43.
- Köhler, C.; Pohl, S. & Carstensen, C. (2015). Investigating mechanisms for missing responses in competence tests. *Psychological Test & Assessment Modeling*, 57 (4), 499-522
- Koretz, D.; Lewis, E.; Skewes-Cox, T. & Burstein, L. (1993). Omitted and not-reached items in mathematics in the 1990 National Assessment of Educational Progress (CRE Technical

- Report 347). Los Angeles: Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Lord, F. M. (1980). Applications of item response theory to practical testing problems. Hillsdale, NJ: Erlbaum.
- Ludlow, L. H. & O'leary, M. (1999). Scoring omitted and not-reached items: Practical data analysis implications. *Educational & Psychological Measurement*, 59, 615-630.
- McKnight, P. E.; McKnight, K. M. & Figueredo, A. J. (2007). Missing data: A gentle introduction. New York: Guilford Press.
- Mislevy, R. J. & Wu, P. K. (1996). Missing responses and IRT ability estimation: Omits, choice, time limits, and adaptive testing (ETS Research Rep. no. RR-98-30-ONR). Princeton, NJ: Educational Testing Service.
- Mullis, I. V. S.; Martin, M. O. & Diaconu, D. (2004). Item analysis and review. In M. O. Martin, I. V. S. Mullis, & S. J. Chrostowski (Eds.), TIMSS 2003 technical report (pp. 225-252). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Organisation for Economic Co-operation and Development. (2009). Pisa 2006 technical report. Paris, France: Author.
- Osterland, S. J. (1990) toward a uniform definition of a test item. *Educational Research Quarterly*, 14, 2-5
- Pohl, S.; Grafe, L. & Rose, N. (2014). Dealing with Omitted and Not-Reached Items in Competence Tests: Evaluating Approaches Accounting for Missing Responses in Item Response Theory Models. *Educational & Psychological Measurement*, 74 (3) 423-452.
- Pohl, S.; Haberkorn, K.; Hardt, K. & Wiegand, E. (2012). Technical Report of Rreading –Scaling results of Starting Cohort 3 in fifth grade (NEPS Working Paper No. 15). Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- Reckase, Mark D. (2009). Multidimensional Item Response Theory.
- Rose, N.; Von Davier, M. & Xu, X. (2010). Modeling nonignorable missing data with item response theory (IRT) (ETS Research Rep. no. RR-10-11), Princeton, NJ: Educational Testing Service.
- Rose, N. (2013). Item nonresponses in educational and psychological measurement (Unpublished doctoral dissertation). Friedrich-Schiller-University of Jena, Germany.
- Rubin, D. B. (1987). Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons.

- Rubin, D. B. (2004). Multiple imputations for nonresponse in surveys. Hoboken, NJ: Wiley-Interscience.
- Schmidt, W. H.; Wolfe, R. G. & Kifer, E. (1993). The identification and description of student growth in mathematics achievement. In L. Burstein (Ed.), *The IEA study of mathematics III: Student growth and classroom processes* (pp. 59–99). New York: Pergamon Press.
- Thomas, N.; Raghunathan, T.; Schenker, N.; Katzoff, M. & Johnson, C. (2006). An evaluation of matrix sampling methods using data from the national health and nutrition examination survey. *Survey Methodology*, 32 (2), 217.
- Zhang, J. (2013). Relationships between missing responses and skill mastery profiles of cognitive diagnostic assessment. Doctoral dissertation. University of Toronto.