



## Differential Item Functioning of Psychasthenia and Psychopathic Deviate Items in Entrance Employment Tests: A Comparison between the Mantel-Haenszel and Rasch Model

**Mahdi Rahmani Malek Abad, PhD student**

Assessment and Measurement, Faculty of Psychology and Educational Sciences, Allameh Tabataba'i University, Tehran, Iran

**Mohammad Reza Falsafinejad, Ph.D**

Associate Professor, Faculty of Psychology and Educational Sciences, Allameh Tabataba'i University, Tehran, Iran

**Asghar Minaei, Ph.D**

Associate Professor, Faculty of Psychology and Educational Sciences, Allameh Tabataba'i University, Tehran, Iran

**Noor Ali Farrokhi, Ph.D**

Associate Professor, Faculty of Psychology and Educational Sciences, Allameh Tabataba'i University, Tehran, Iran

### Abstract

Choosing healthy human resources is one of the main tasks of recruitment tests, provided that the measurement tools are fairness in test. Methods: This research is survey research and the statistical population was all the Recruitment candidates of Isfahan city. The sample size was 5997 people who were selected by one-step cluster method. The tools used was the MMPI-2. The fit of the Items with the model was confirmed using outfit MnSq and infit MnSq, and DIF according to gender was checked and compared with Rasch Model and Mantel-Haenszel by Winsteps version 3.64. Results: After fitting the data with Rasch model, DIF analysis was performed. The results showed that 38% of psychopathic deviate items and 22.9% of psychasthenia items were biased towards men. Mantel-Haenszel had a weaker performance than Rasch Model in some items. Conclusion: Psychopathic deviate and psychasthenia and items have been to the bias of men and they have naturally obtained higher scores in these two scales. Also, due to the strength of the Rasch model compared to the CTT, it is better to use this model to detect biased items.

**Keyword:** differential item functioning, Rasch model, Mantel-Haenszel, psychopathic deviate, psychasthenia, recruitment candidates.

## بررسی کنش افتراقی سوالات ضعف روانی و ضداجتماعی در آزمون‌های استخدامی بدو ورود: مقایسه روش متل‌هنزل و مدل راش

\*مهدی رحمانی ملک‌آباد

دانشجوی دکتری سنجش و اندازه‌گیری، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه علامه طباطبائی، تهران، ایران

محمد رضا فلسفی‌نژاد

دانشیار گروه سنجش و اندازه‌گیری، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه علامه طباطبائی، تهران، ایران.

اصغر مینائی

دانشیار گروه سنجش و اندازه‌گیری، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه علامه طباطبائی، تهران، ایران.

نورعلی فرخی

دانشیار گروه سنجش و اندازه‌گیری، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه علامه طباطبائی، تهران، ایران.

### چکیده

انتخاب نیروی انسانی سالم جزء وظایف اصلی آزمون‌های استخدامی است به شرط آنکه ابزارهای اندازه‌گیری دارای انصاف و عدالت در سنجش باشند. روش: پژوهش از نوع پیمایشی است و جامعه آماری کلیه داوطلبان استخدامی شهر اصفهان بود. حجم نمونه 5997 نفر بود که به روش خوشه‌ای یک‌مرحله‌ای انتخاب شد. ابزار مورد استفاده پرسشنامه MMPI-2 بود. برازش سوالات با مدل با استفاده از شاخص‌های outfit MnSq و infit MnSq تأیید شد و کنش افتراقی سوالات بر حسب جنسیت با دو رویکرد مدل راش و متل‌هنزل توسط نرم‌افزار Winsteps نسخه 3.64 بررسی شد. یافته‌ها: پس از آنکه برازش داده‌ها با مدل تأیید شد، تحلیل کنش افتراقی انجام شد. نتایج نشان داد که 38 درصد از سوالات ضداجتماعی و 22/9 درصد از سوالات ضعف‌روانی دارای تورش به جهت مردان است. روش متل‌هنزل در برخی سوالات عملکرد ضعیف‌تری نسبت به مدل راش داشت. نتیجه‌گیری: نتایج نشان داد که سوالات ضداجتماعی و ضعف‌روانی به ضرر مردان است و مردان به طور سوء‌گیرانه‌ای نمرات بالاتری را به دست خواهند آورد. همچنین بهتر است که از مدل راش به جای روش متل‌هنزل در تشخیص سوالات دارای تورش استفاده کرد.

**کلید واژه:** کنش افتراقی سوال، مدل‌راش، متل‌هنزل، ضداجتماعی، ضعف‌روانی، داوطلبان استخدامی.

## مقدمه

استخدام جزء یکی از مهمترین و پیچیده‌ترین فرایندهای هر سازمان است که در آن سازمان به دنبال به‌کارگیری افرادی توانمند است که هم از نظر علمی و مهارتی مناسب و هم از نظر روانی سالم باشند، چرا که اگر افراد از نظر روانی سالم نباشند در آینده می‌توانند با ایجاد تنش بین‌فردی یا قانون‌گریزی و سهل‌انگاری موجب کاهش بازدهی و کم‌اثر شدن فرایند کار، ایجاد خلل سازمانی و آسیب رساندن به همکاران از نظر روانی و حتی جسمی شوند، به‌ویژه اگر سازمان هدف مسئول آموزش و تعلیم نسل آینده کشور باشد. بدین جهت شناسایی و به‌کارگیری افراد سالم از نظر روان‌شناختی بسیار قابل اهمیت است ( Baker, Baum & Francis, 2023). یکی از ابعادی که توسط آزمون‌های بدو ورود به سازمان‌ها مورد سنجش قرار می‌گیرد، بررسی سطح ضداجتماعی و ضعف‌روانی افراد است. ویژگی افراد پرخاشگر در محیط کار عدم رعایت حقوق دیگران و احترام نگذاشتن به قوانین کار است. رفتار افراد دارای اختلال ضداجتماعی و ضداجتماعی خارج از استانداردهای اجتماعی، رفتاری و اخلاقی سازمان‌ها است، و این اختلالات باعث کاهش بازدهی سازمان می‌شود (Cao, Li, Van Der Wal & Taris, 2022). روان‌شناسان بالینی در اختلال شخصیت ضداجتماعی (Psychasthenia) دو صفت را مشخص کرده‌اند. اولین صفت بی‌تفاوتی، بی‌احساسی و خودخواهی است. به این صورت که فرد از کارهای خود احساس پشیمانی نمی‌کند و نمی‌تواند با دیگران همدلی داشته باشد. دومین ویژگی سبک زندگی خاصی است که به علت فکر نکردن قبل از عمل است و باعث می‌شود رفتارهای ضداجتماعی به وجود آید (Antikchi, Bigdeli & Sabahi, 2017).

بعد دیگری که در محیط کار می‌تواند به فرد یا سازمان آسیب برساند داشتن شخصیت ضعف‌روانی (Psychopathic Deviate) است. ضعف‌روانی، معمولاً احساس مبهم و نامطموع ترس و تشویش تعریف می‌شود. اجزای جسمی و هیجانی ضعف‌روانی شبیه اجزای ترس است، اما شدیدتر از آن است. این اشخاص نه تنها مشکلاتی در تصمیم‌گیری دارند، بلکه پس از آن نیز به‌خاطر احتمال اشتباهات و وقایع پیش‌بینی نشده‌ای

که ممکن است منجر به مصیبت گردد بسیار نگران می‌شوند. این اختلال از یک سو می‌تواند در قالب شکایت از خستگی فزاینده پس از فعالیت ذهنی بروز کند که همراه با کاهش نسبی در عملکرد شغلی و حتی انجام فعالیت‌های روزانه خواهد بود و از سوی دیگر می‌تواند فرد را دچار احساس ضعف جسمی یا بدنی و فرسودگی ناشی از کمترین فعالیت کند که با دردها و ناراحتی عضلانی و ناتوانی در شل کردن آنها همراه است (Queirolo, Bacci, Roccon, Zanette & Mucignat, 2023) و این امر در بین معلمین بیشتر اهمیت دارد زیرا معلمین هر روز با چالش‌های جدیدی در بین دانش‌آموزان مواجهه هستند. این دو مولفه ضداجتماعی و ضعف‌روانی در آزمون‌های بدو استخدام توسط مقیاس انحراف روانی-اجتماعی (Psychopathic Deviate) و ضعف‌روانی-اجتماعی (Psychasthenia) پرسشنامه MMPI-2 مورد سنجش قرار می‌گیرد. اما مسئله مهم دیگری که در سنجش این مقیاس‌ها موثر است، تفاوت‌های جنسیتی است. برخی پژوهش‌ها نشان داد که نتایج پرسشنامه MMPI-2 احتمالاً دارای سوگیری است، در حالی که دیگران احساس کردند که خود پرسشنامه MMPI-2 حاوی سؤالات جنسیتی و نژادپرستانه است که می‌تواند منجر به عدم انصاف در نمره‌گذاری و اعلام ناصحیح نتایج شود (Talerico, McCallum, Whitman, Tarescavage, Corey & Ben-Porath, 2023). پژوهش‌های دیگر در مورد تفاوت ضداجتماعی زن و مرد در حالت و صفت انجام شده است، بررسی‌ها در مورد ابراز خشم مردان و زنان حاکی از آن است که در مردان حمله فیزیکی و کلامی بیشتر از زنان‌هاست و ابراز خشم در زنان بیشتر به صورت گریه کردن ظاهر می‌شود ولی در هر دو حالت ضداجتماعی و خشم مشاهده می‌شود (Björkqvist, 2018). نتایج کلی بررسی‌ها نشان می‌دهد که مردان و زنانی که چند ویژگی مردانه دارند، تمایل بیشتری به برون‌ریزی خشم دارند و مردان و زنانی که ویژگی‌های زنانگی در آنها بارزتر است، تمایل به درون‌ریزی خشم دارند و این وجود ضداجتماعی و ضعف‌روانی یکسان با حالت‌های بروز متفاوت می‌تواند سبب پاسخگویی به سؤالات به صورت متفاوت گردد که در پی آن نتایج به دور از واقعیت خواهد بود و موجب سوءگیری آزمون می‌شود. از

درست نمایی (likelihood ratio test) است (Embretson & Reiss, 2000; Karami, Gramipour & Minaei, 2021). روش‌های بررسی DIF مبتنی بر نمرات مشاهده شده و متغیرهای مکنون می‌توانند در چهارچوب مدل‌های مختلط خطی و غیرخطی تعمیم داده‌شده ترکیب شود و مدل‌های سلسله‌مراتبی IRT و رگرسیون لوجستیک (logistic regression) را به وجود بیاورد. در این مدل‌ها اندازه اثر DIF و جایگزینی داده‌های گمشده امکان‌پذیر است (Skronidal, Gori and Marin, 2004). (Rabe-Hesketh and Boca Raton, 2004) در پژوهشی به بررسی کنش افتراقی سوالات اسکیزوفرنی، ضعف روانی و ضداجتماعی پرسشنامه MMPI-2 در بین دانشجویان تازه وارد دانشگاه با مدل راش (Rasch Model) پرداخت و نشان داد برخی تفاوت‌ها در نحوه برازش و تشخیص سوالات وجود داشت. همچنین Card, (2008) Stucky, Sawalani & Little در یک بررسی فرا تحلیلی از 148 مطالعه در مورد تفاوت‌های جنسیتی در ضداجتماعی مستقیم و غیرمستقیم نشان داد که یک تفاوت جنسیتی به جهت مردان در مورد ضداجتماعی مستقیم است، اما فقط یک تفاوت جنسی ناچیز به جهت زنان در مورد ضداجتماعی غیرمستقیم وجود دارد. به همین دلیل با توجه به اهمیت رعایت انصاف در گزینش نیروی انسانی سالم و مناسب در آزمون‌های بدو ورود و انتخاب بدون سوءگیری افراد، سوال اول به تعیین ضرایب مطلوبیت سوالات پرداخته است؟ سوال دوم این پژوهش این است که آیا سوالات مقیاس ضداجتماعی و ضعف روانی پرسشنامه MMPI-2 با مدل راش برازش دارد؟ سوال سوم اینکه کدام سوالات مقیاس ضداجتماعی و ضعف روانی با مدل راش و کدام با روش متل‌هنزل دارای کنش افتراقی است؟ و سوال چهارم اینکه آیا تفاوتی در تشخیص سوالات دارای سوءگیری بین دو روش متل‌هنزل و مدل راش وجود دارد یا خیر؟

#### روش پژوهش

طرح این پژوهش توصیفی و از نوع پیمایشی است. جامعه آماری کلیه داوطلبان استخدامی شهر اصفهان بودند که در آزمون‌های بدو ورود به مراکز طب‌کار شهر اصفهان در زمستان

طرفی مرحله گزینش و استخدام اغلب همراه با سوءگیری و عدم انصاف در گزینش به ویژه تورش در جنسیت بوده است (Zahid, nasrollahi & Mahinizadeh, 2020). هرچند که چندین قانون جهت از بین بردن این تبعیض‌ها در نظر گرفته شده است اما همانطور که بیان شد محققان هنوز به خوبی به تاثیر تفاوت‌های جنسیتی در نحوه پاسخگویی افراد به سوالات آزمون‌های روانشناختی نپرداخته‌اند (Baker et al, 2023).

یکی از راه‌ها برای پی‌بردن به اینکه آزمون غیرمنصفانه است کشف پرسش‌های دارای کنش افتراقی (Differential Item Functioning) است. (Dorans & Holland 1993) کنش افتراقی را چنین تعریف کردند: "منظور از DIF، تفاوت‌های موجود در عملکرد هر پرسش از گروه‌ها با توجه به توانایی یا ویژگی است که پرسش‌ها از افراد گروه بطور مفهومی می‌خواهد به دست آورد. DIF یک تفاوت ناخواسته بین گروه‌های امتحان دهنده است که قرار است براساس تفاوت در توانایی‌هاشان در هر پرسش و هر امتحان قیاس شوند" (Wiberg, 2007). (Camilli & Congdon 1999) DIF را در قالب یک مفهوم بی‌طرفی آزمون (Test fairness) توضیح داده است. روش‌های آماری آشکارسازی DIF را می‌توان دو دسته کلی تقسیم نمود. دسته اول براساس نظریه کلاسیک آزمون (Classic Test Theory) است. نمرات مشاهده شده در این روش به عنوان متغیر هم‌تا برای پیش‌بینی نمرات واقعی در نظر گرفته خواهد شد (Lord & Novick, 1968). از جمله روش‌هایی که در این دسته قرار می‌گیرند عبارت از روش متل‌هنزل (Mantel- Haenszel)، روش متل‌هنزل تعمیم داده‌شده (Generalized Mantel- Haenszel) (Penfield & Algina, 2003 ; Su & Wang, 2005)، روش اختلاف میانگین استاندارد شده (Standadized Mean Difference) و روش رگرسیون لوجستیک است (Swaminathan & Rogers, 1990; Zumbo, 2003). دوم شامل روش‌های مبتنی بر نظریه سوال- پاسخ (Item Response Theory) است، زیرا نمره متغیر مکنون به عنوان متغیر هم‌تا برای برآورد نمرات متغیر مکنون مورد استفاده قرار می‌گیرد. روش‌های این دسته شامل روش‌های مجذور کای- دو لرد (Lord's chi-square)، راجو و روش‌های آزمون نسبت

چیزهائی را دزدیده‌ام) و ضعف‌روانی (Psychasthenia) با 48 سوال (برای مثال سوال 31 با عبارت: برایم مشکل است که حواسم را روی یک کار متمرکز کنم) پرداخته شد. برای مثال یک سوال از مقیاس ضد اجتماعی پاسخ‌های به این ماده‌ها به صورت دو گزینه ای بلی/ درست (1) و خیر/ غلط (0) است. روائی این مقیاس در مطالعات فراوانی بررسی و مورد تأیید قرار گرفته است و نشان داده شده است که مقیاس انحراف روانی و اجتماعی با افرادی که دارای نشانگان بالینی ضد اجتماعی دارند همخوان است و دارای روائی قابل قبولی است (Marnat & Wright, 2016). ضریب اعتبار مقیاس ضداجتماعی و ضعف‌روانی پرسشنامه اصلی با روش آلفای کرونباخ برابر با 0/80 و 0/90 و اعتبار مقیاس ضداجتماعی و ضعف‌روانی از راه باز آزمایشی برابر 0/88 و 0/93 بود (Valianpour, Morteza Modarres Gharavi & Behrouz, 2020). اعتبار مقیاس ضداجتماعی و ضعف‌روانی در این تحقیق با روش آلفای کرونباخ به ترتیب برابر با 0/76 و 0/81 بدست آمد.

#### روش اجراء و تحلیل داده‌ها

هدف این پژوهش بررسی و مقایسه کنش‌افتراقی سوالات مولفه ضداجتماعی و ضعف‌روانی پرسشنامه چندوجهی شخصیتی مینه‌سوتا-2 با استفاده از نظریه روش مثل‌هنزل و مدل راش برای داوطلبان استخدامی بود. بدین منظور ابتدا فهرست مراکز طب کار شهر اصفهان استخراج و از بین 17 مرکز 9 مرکز اعلام آمادگی برای همکاری کردند که از این بین 3 مرکز به تصادف و با توجه به حجم داوطلبان استخدامی انتخاب و پس از صحبت با مسئول مرکز مقرر شد ضمن حفظ محرمانه بودن اسامی داوطلبان شرکت داده‌شده در طرح پژوهش و محرمانه ماندن نتایج به دست آمده از آزمون‌های هر فرد از مراجعه کنندگان برای معاینات استخدامی، آزمون MMPI-2 در محل مرکز گرفته شود و نتایج آزمون‌ها به صورت هفتگی در اختیار دستیار پژوهشگر قرار گیرد. دستیار پژوهشگر نیز پاسخ‌های افراد را وارد نرم افزار کرد و پس از جمع‌آوری در اختیار پژوهشگر قرار داده شد. برای محاسبه آماره‌های توصیفی میانگین و انحراف

1401 و بهار 1402 مراجعه نمودند. نمونه مورد مطالعه به روش خوشه‌ای غیر تصادفی از مراکز طب کار پژوهش انتخاب شد. مراکز طب‌کار بنا به جدیدترین تعریف از کالج سلطنتی پزشکان و جراحان کانادا (2012) به مراکز گفته می‌شود که با ادغام جنبه‌های بالینی و اجرایی و با در نظر داشتن محیط شغلی افراد و گروه‌ها به نیازهای مربوط به سلامت ایشان می‌پردازد و شامل فعالیت‌هایی از جمله شناسایی، ارزیابی، درمان، کنترل و پیگیری درمان و پیشگیری از بیماری‌ها جسمی و روانشناختی و صدمات مرتبط با کار است. تعداد کل داوطلبان استخدامی در سال 1401 و بهار 1402 شهر اصفهان بیش از 10 هزار نفر بوده است که با تحصیلات دیپلم به بالا متقاضی شغل معلمی (یا دانشجو معلمی) بودند. حجم نمونه در این پژوهش با توجه به شیوه تجزیه و تحلیل داده‌ها و استفاده از مدل‌های برازش راش و پس از حذف تعداد افرادی که ناقص به سوالات پاسخ داده‌اند (طبق دستور العمل آزمون افرادی که به تعداد 35 سوال و بیشتر پاسخ نداده‌اند، پاسخ‌نامه آن‌ها قابل استفاده نیست) برابر با 5997 نفر در نظر گرفته شد که پس از تحلیل‌های اولیه نشان داد که در داده‌ها، داده پرت تک‌متغیری (univariate outlier) وجود نداشت، اما بر اساس آماره ماهالانوبیس (Mahalanobis) تعداد 733 داده پرت چندمتغیری (multivariate outlier) وجود داشت که تمام این 733 نفر از فایل داده‌ها حذف و تحلیل بر روی حجم 5264 نفر انجام گرفت.

#### ابزار سنجش

ابزارهای استفاده شده در این پژوهش عبارت از پرسشنامه چندوجهی شخصیتی مینه‌سوتا-2 (Minnesota Multiphasic Personality Inventory-2) بود. این پرسشنامه توسط هاتاوی و مک کینلی در 1943 ساخته شده و در سال 1970 و در نهایت سال 2008 مورد آخرین بازنگری قرار گرفت و باز انتشار شد. این نسخه از مقیاس با 370 ماده دارای 3 شاخص روائی (K و F, L) و 10 شاخص بالینی است. در این پژوهش به مولفه منحرف ضداجتماعی (Psychopathic Deviate) با 50 سوال (برای مثال سوال 35 با عبارت: در گذشته گاهی

از 0.5 برای برازش مقیاس اندازه‌گیری بهره پائینی دارد و میزان اعتبار را به صورت غیر واقعی بالا می‌برد. مقادیر ZSTD نیز بهتر است بین -2 تا +2 گزارش شود، البته با توجه به اینکه این آماره به حجم نمونه بسیار حساس است (Lincare, 2012)، و حجم نمونه این پژوهش بالاست، در این پژوهش جهت بررسی برازش سوال با مدل از آماره MnSq که به دو شکل outfit و infit آمده است، استفاده شد.

#### یافته‌ها

با توجه به اینکه هدف این پژوهش بررسی کنش‌افترافی سوالات مولفه ضداجتماعی و ضعف‌روانی پرسشنامه MMPI-2 در بین داطلبین استخدامی با مدل راش در کنار روش کلاسیک اندازه‌گیری (متل‌هنزل) بود، ابتدا به تبیین تناسب مدل و داده‌ها و مفروضات نظریه راش پرداخته شد و سپس شاخص‌های کمی سنجش سوالات گزارش شد. شاخص‌های کمی در جدول 1 آمده است.

استاندارد و ضریب آلفای کرونباخ و اعتبار مقیاس از طریق نظریه کلاسیک از نرم افزار SPSS نسخه 27 استفاده شد و برای بررسی برازش سوال از آماره MnSq outfit و infit MnSq و با استفاده از نرم افزار Winsteps نسخه 3.64 بهره گرفته شد، همچنین ضریب مطلوبیت سوالات نیز با شاخص Measure بیان شد. در بررسی تناسب شاخص‌ها ابتدا به شاخص MNSQ مراجعه می‌شود و در صورت عدم تائید به شاخص ZSTD باید رجوع کرد. مقادیر MNSQ از 0 تا مثبت بی‌نهایت قرار دارد و مقدار مطلوب برابر 1 است. هر چه این مقدار به 0 نزدیک شود نشان از بیش برازش الگوهای پاسخ با مدل است که به معنای وجود وابستگی در میان پاسخ‌ها یا سوالات است و مقادیر بالاتر از 1 بیانگر انحراف از تک بعدی بودن است. در کل مقادیر بیشتر از 2 غیر مفید و حاکی از اعتبار و برازش پائین است. مقادیر بین 1.5 تا 2 با اینکه بالاست اما گمراه کننده نیست و قابل اغماض است. مقادیر بین 0.5 تا 1.5 برای برازش مقیاس اندازه‌گیری مناسب و میزان اعتبار مقیاس را افزایش می‌دهد و در نهایت مقادیر کمتر

جدول 1. پیش‌فرض‌های مدل‌راش و بررسی تناسب مدل با داده‌ها

مقیاس ضعف‌روانی					مقیاس ضداجتماعی				
Outfit ZSTD	Outfit MNSQ	Infit ZSTD	Infit MNSQ	شماره سوال Measure	Outfit ZSTD	Outfit MNSQ	Infit ZSTD	Infit MNSQ	شماره سوال Measure
9,9	1,51	9,9	1,34	33 3,59-	4,9	1,35	1,3	1,06	3 0,26
9,9	1,46	9,7	1,24	293 3,46-	3,8	1,29	1,0	1,05	214 0,40
9,9	1,42	9,9	1,17	309 2,58-	9,9	1,26	9,9	1,16	158 -1,54
9,9	1,39	9,9	1,27	242 1,69-	7,1	1,24	6,0	1,13	171 -0,94
9,6	1,36	9,9	1,27	321 0,96-	7,8	1,15	7,7	1,10	143 -1,75
2,3	1,33	1,0	1,08	174 1,92	6,8	1,15	6,5	1,09	113 -1,60
4,5	1,19	6,2	1,14	9 -0,71	6,9	1,13	7,0	1,09	157 -1,82
1,9	1,13	2,8	1,09	170 0,22	3,1	1,12	3,0	1,07	34 -0,81
0,2-	0,99	4,1	1,10	23 -0,40	3,2	1,11	3,2	1,07	261 -0,93
0,8	1,07	1,9	1,08	165 0,78	6,8	1,10	6,1	1,06	263 -2,65
0,3	1,01	3,6	1,07	285 -0,90	0,7	1,09	0,0	0,99	264 1,64
0,9	1,07	1,8	1,06	140 0,35	5,1	1,09	5,7	1,06	79 -1,94
0,7	1,04	1,8	1,05	313 0,06	1,2	1,08	0,8	1,03	12 0,24
-0,6	0,94	0,3	1,01	109 0,90	1,0	1,07	0,0	1,00	167 0,28

-2,1	0,82	0,2	1,01	0,89	<b>317</b>	1,2	1,05	1,8	1,05	-0,57	<b>266</b>
-3,1	0,89	0,4	1,01	-0,87	<b>304</b>	1,4	1,04	2,1	1,04	-1,13	<b>209</b>
-2,1	0,85	0,0	1,00	0,37	<b>3</b>	0,5	1,02	0,7	1,02	-0,64	<b>9</b>
0,0	0,99	0,0	1,00	1,18	<b>308</b>	0,6	0,93	0,0	1,00	1,23	<b>160</b>
-2,1	0,91	0,1-	1,00	-0,63	<b>331</b>	-2,9	0,82	0,0	1,00	0,27	<b>105</b>
-2,0	0,87	-0,3	0,99	0,32	<b>320</b>	-0,3	0,98	-0,1	1,00	0,02-	<b>225</b>
-3,3	0,92	-1,1	0,98	-1,47	<b>326</b>	0,0	0,99	-0,1	0,98	2,31	<b>54</b>
-1,0	0,89	0,3-	0,98	1,33	<b>11</b>	-0,9	0,98	-0,8	0,99	-1,75	<b>267</b>
-2,9	0,82	0,6-	0,98	0,18	<b>275</b>	-0,5	0,99	-0,8	0,99	-1,44	<b>217</b>
-2,3	0,82	0,5-	0,98	0,59	<b>147</b>	-1,0	0,91	-0,2	0,99	0,77	<b>125</b>
-3,9	0,89	-2,6	0,96	-1,28	<b>38</b>	-2,1	0,86	-0,3	0,98	0,27	<b>219</b>
-3,7	0,76	-1,4	0,95	0,35	<b>16</b>	-0,4	0,96	-0,2	0,98	0,95	<b>70</b>
-2,6	0,78	-1,2	0,95	0,89	<b>310</b>	-1,4	0,86	-0,2	0,98	1,08	<b>35</b>
-5,0	0,79	-2,6	0,94	-0,46	<b>4</b>	-3,1	0,91	-0,9	0,98	-1,04	<b>226</b>
-5,1	0,83	-3,5	0,94	-0,95	<b>89</b>	-3,6	0,81	-1,0	0,96	-0,05	<b>42</b>
-6,3	0,85	-4,8	0,93	-1,56	<b>218</b>	-2,9	0,81	-0,8	0,96	0,38	<b>243</b>
-2,0	0,83	-1,6	0,93	0,89	<b>31</b>	-3,7	0,94	-3,6	0,96	-1,84	<b>122</b>
-4,8	0,62	-2,2	0,90	0,88	<b>130</b>	-2,5	0,79	-0,7	0,96	0,83	<b>95</b>
-2,9	0,74	-2,1	0,90	1,03	<b>175</b>	-2,7	0,69	-0,4	0,96	1,67	<b>202</b>
-3,5	0,62	-1,6	0,90	1,66	<b>329</b>	-1,0	0,89	-0,6	0,95	1,37	<b>195</b>
-6,0	0,45	-1,8	0,90	1,45	<b>327</b>	-3,6	0,81	-1,4	0,95	-0,07	<b>185</b>
-6,1	0,65	-3,3	0,90	0,20	<b>82</b>	-4,0	0,54	-0,5	0,95	1,79	<b>288</b>
-4,8	0,65	-2,8	0,89	0,65	<b>73</b>	-2,4	0,76	-0,7	0,94	1,26	<b>31</b>
-6,1	0,63	-3,3	0,89	0,29	<b>56</b>	-4,5	0,84	-2,4	0,94	-0,73	<b>89</b>
-6,3	0,51	-2,7	0,88	0,99	<b>316</b>	-5,0	0,81	-2,3	0,94	-0,62	<b>99</b>
-7,4	0,54	3,8-	0,87	0,48	<b>302</b>	-5,2	0,84	3,0-	0,94	-0,98	<b>32</b>
-8,1	0,58	-5,0	0,86	0,06	<b>328</b>	-4,2	0,89	-3,8	0,94	-1,28	<b>129</b>
-7,0	0,47	3,3-	0,86	0,99	<b>325</b>	-4,3	0,81	-2,2	0,93	-0,30	<b>94</b>
-7,4	0,47	3,5-	0,86	0,85	<b>273</b>	-4,5	0,60	-1,1	0,92	1,11	<b>22</b>
-7,6	0,43	3,3-	0,85	1,01	<b>289</b>	-3,3	0,73	-1,3	0,92	0,84	<b>17</b>
-8,0	0,63	-6,0	0,85	-0,19	<b>301</b>	-6,1	0,80	-3,6	0,92	-0,81	<b>259</b>
-8,6	0,59	-6,6	0,83	-0,09	<b>277</b>	5,0-	0,66	-1,6	0,92	0,54	<b>82</b>
-8,5	0,46	-5,1	0,82	0,58	<b>65</b>	-4,3	0,41	-0,5	0,92	2,55	<b>21</b>
-9,4	0,45	-6,5	0,79	0,44	<b>196</b>	-4,3	0,66	-1,5	0,91	0,82	<b>56</b>
--	--	--	--	--	--	-2,9	0,63	-0,8	0,91	1,93	<b>71</b>

--	--	--	--	--	--	-3,8	0,48	-0,7	0,90	2,46	52
----	----	----	----	----	----	------	------	------	------	------	----

سوالات در سطح متوسطی از درجه سختی و یا مطلوبیت قرار داشتند. ستون سوم و چهارم مربوط به شاخص‌های تناسب Infit و ستون پنجم و ششم متعلق به شاخص‌های تناسب Outfit است. در پاسخ به سوال دوم می‌توان بیان داشت که با توجه به نتایج جدول، شاخص‌های برازش تمامی سوالات در دامنه مطلوبی قرار دارد و سوالات برازش خوبی با مدل راش دارد. پس از تأیید برازش داده‌ها با مدل، در ادامه برای تشخیص و مقایسه سوالات مقیاس ضداجتماعی و ضعف‌روانی دارای کنش افتراقی بر حسب جنسیت از دو شیوه کلاسیک (متل‌هنزل) و نظریه سوال- پاسخ (مدل‌راش) استفاده شد و نتایج آن در جدول 2 آورده شد.

در جدول شماره 1 ستون اول از سمت راست به شماره سوالات مولفه ضداجتماعی و ضعف‌روانی در پرسشنامه MMPI-2 تعلق دارد و اینکه ترتیب هر سوال در پرسشنامه برای شرکت‌کنندگان به چه صورت بوده است. در ستون دوم (Measure) به میزان و درجه سختی هر پرسش در واحد لوجیت (Logit) اشاره دارد که در پرسشنامه‌های شخصیتی به عنوان ضریب دشواری یا مطلوبیت معرفی گردید (Lincare, 2010). سخت‌ترین سوالات مقیاس ضداجتماعی متعلق به سوال 21، 52 و 54 بود و سخت‌ترین سوالات مقیاس ضعف‌روانی که کمترین میزان پاسخ‌دهی مثبت را داشتند سوالات 174 و 329 بوده است و در این مقیاس بالاترین میزان مقبولیت متعلق به سوالات 33، 293 و 309 بود. مابقی

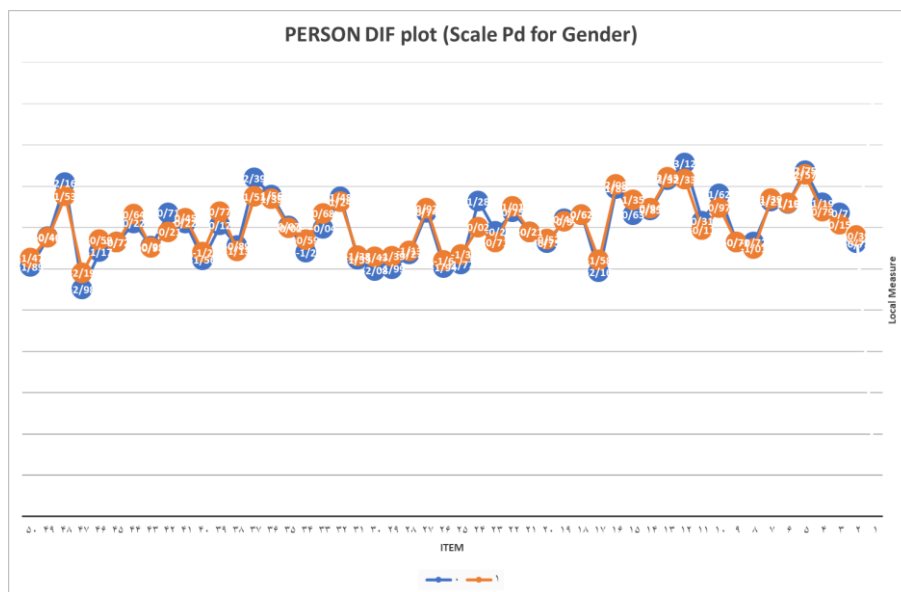
جدول 2. مقایسه توانش مدل‌راش و متل‌هنزل در تشخیص کنش افتراقی پرسش‌ها بر حسب جنسیت

جهت تورش	مدل راش			متل‌هنزل			شماره سوال	نتیجه
	نتیجه	Welch Prob.	DIF Size	DIF Contrast	نتیجه	Prob.		
F	B	0,0000	0,37	0,55	B	0,0000	0,60	12
F	B	0,0004	0,45	0,64	A	0,0045	0,22	35
M	C	0,0000	-0,41	-0,72	C	0,0000	-0,82	70
--	A	0,4221	0,12-	-0,19	C	0,0035	-0,75	71
M	B	0,0000	-0,32	-0,57	B	0,0000	-0,48	79
F	B	0,0000	0,32	0,51	B	0,0000	0,40	99
F	C	0,0000	0,92	1,26	C	0,0000	1,08	105
M	B	0,0000	-0,26	-0,46	A	0,0000	-0,30	113
--	A	0,0000	-0,19	-0,34	B	0,0000	-0,45	122
M	B	0,0000	-0,33	-0,60	B	0,0000	-0,44	143
M	B	0,0000	-0,36	-0,64	B	0,0000	-0,51	157
M	C	0,0000	-0,40	-0,72	C	0,0000	-0,76	167
M	B	0,0000	-0,34	-0,61	A	0,0000	-0,38	171
F	C	0,0006	0,64	0,88	A	0,0103	0,10	202
M	B	0,0000	-0,37	-0,64	B	0,0002	-0,43	214
F	C	0,0000	0,64	0,93	C	0,0000	0,86	225
M	B	0,0003	-0,25	-0,43	B	0,0000	-0,58	243
M	B	0,0000	-0,33	-0,57	B	0,0000	-0,47	261

سوالات ضداجتماعی

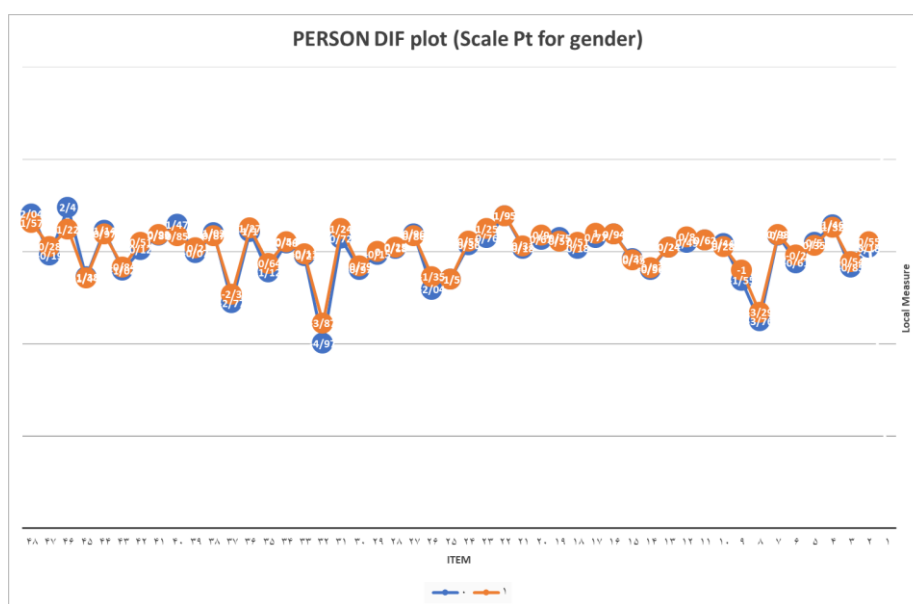
M	C	0,0000	-0,43	-0,79	C	0,0000	-0,68	263
F	B	0,0097	0,43	0,63	A	0,0086	0,23	264
M	B	0,0000	-0,23	-0,43	B	0,0000	-0,49	267
M	B	0,0000	-0,33	-0,55	C	0,0000	-0,68	38
M	B	0,0001	-0,31	-0,49	C	0,0000	-0,66	175
--	A	0,0657	-0,12	-0,19	B	0,0000	-0,57	196
M	C	0,0000	-0,41	-0,69	B	0,0000	-0,47	242
M	B	0,0000	-0,33	-0,52	C	0,0000	-0,98	289
M	C	0,0000	-0,61	-1,11	B	0,0000	-0,53	293
M	B	0,0000	-0,29	-0,47	B	0,0000	-0,57	304
M	B	0,0000	-0,27	-0,47	A	0,0000	-0,36	309
F	B	0,0000	0,45	0,62	B	0,0004	0,43	316
F	C	0,0000	0,92	1,18	C	0,0000	0,83	327
M	B	0,0000	-0,29	-0,47	C	0,0000	-0,76	328
F	B	0,0104	0,34	0,47	A	0,1426	0,10	329

= زنان F = مردان M



شکل 1. نمودار کنش افتراقی سوالات ضداجتماعی بر حسب ویژگی جنسیت (0 = زن و 1 = مرد)





شکل 2. نمودار کنش افتراقی سوالات ضعف روانی بر حسب ویژگی جنسیت (0= زن و 1= مرد)

مورد) کنش افتراقی بین زنان و مردان نداشته است، از بین سوالات دارای کنش افتراقی تعداد 13 سوال دارای کنش افتراقی متوسط و 6 سوال کنش افتراقی بزرگی را بر حسب ویژگی جنسیت گزارش دادند. از بین 48 سوال مقیاس ضعف روانی بر اساس مدل راش تعداد 11 سوال دارای کنش افتراقی تشخیص داده شد که 8 سوال دارای کنش افتراقی در سطح متوسط و 3 سوال دارای کنش افتراقی بزرگ بود. در نهایت در پاسخ به سوال چهارم در ارتباط با مقایسه بین دو رویکرد کلاسیک (متل هنزل) و مدل راش در بررسی کنش افتراقی نیز می‌توان بیان داشت که بیشتر سوالاتی که در مدل راش دارای کنش افتراقی تشخیص داده شدند نیز در روش متل هنزل نیز دارای کنش افتراقی تشخیص داده شدند به جز 5 مورد (سوالات 35، 113، 171، 202 و 264) از مقیاس ضداجتماعی و 2 مورد (سوالات 309 و 329) از مقیاس ضعف روانی که در مدل راش دارای کنش افتراقی بود اما در روش متل هنزل شناسایی نشد، که این نشان دهنده دقت بیشتر مدل راش نسبت به مدل متل هنزل است که توانایی تشخیص بیشتر سوالات دارای کنش افتراقی را داراست. با توجه به حجم سوالات و تعداد نسبتاً یکسان در شناسایی سوالات دارای کنش افتراقی توسط دو روش می‌توان بیان داشت که دو روش متل هنزل و مدل راش با وجود تقریباً عملکرد یکسان در

در ستون اول این جدول شماره سوالات مقیاس ضداجتماعی و ضعف روانی پرسشنامه MMPI-2 آورده شد. سپس نتایج به دست آمده از رویکرد متل هنزل ارائه شد. در این روش مقادیر کنش افتراقی با Size و سطح معناداری آن با Prob. آورده شد. پس از آن نتایج به دست آمده از کنش افتراقی مدل راش گزارش شد که با شاخص‌های DIF Contrast و DIF Size مشخص شد و معناداری کنش افتراقی با Prob. نشان داده شد. برای بررسی وجود کنش افتراقی باید سطح Prob. کمتر از 0/05 گزارش شود. اگر مقادیر Size کمتر از 0/43 باشد نشان از عدم کنش افتراقی (A) است، اگر این شاخص بین 0/43 تا 0/64 باشد دارای کنش افتراقی متوسط (B) است و اگر بالاتر از 0/64 باشد دارای کنش افتراقی بزرگ (C) است. کنش افتراقی بزرگ گواهی بر ناعادلانه بودن سوال است، این بدین معناست که اگر پاسخ به سؤال‌های آزمون با کنترل توانایی آزمودنی‌ها تابعی از ویژگی‌های گروهی که آزمودنی‌ها به آن تعلق دارند، مانند جنسیت و سن باشد، آن سؤال‌ها نسبت به آن افراد دارای کنش افتراقی و سوءگیری است.

همچنین با توجه به نتایج ارائه شده در جدول 2 در پاسخ به سوال سوم می‌توان بیان داشت که از بین 50 سوال مقیاس ضداجتماعی تکمیل شده توسط داوطلبان استخدامی، تعداد 19 سوال دارای کنش افتراقی است و مابقی سوالات (31

کرده است و تنها در 68 سوال از نظر تبعیض جنسیتی انصاف در تشخیص برای مولفه ضداجتماعی و ضعف‌روانی رعایت شده بود. لیکن باید در نظر داشت که مفهوم انصاف آزمون مفهومی چند بعدی است و وجود یا فقدان آن متأثر از عوامل مختلفی است که از زمان طراحی آزمون باید آنان را کنترل کرد و سرانجام در زمان تفسیر نتایج، اثرگذاری آن عوامل را در نظر گرفت.

### بحث و نتیجه‌گیری

هدف از این پژوهش تعیین و تبیین قدرت هر یک از شیوه‌های سنجش و تشخیص کنش‌افتراقی سوالات مقیاس انحراف ضداجتماعی و ضعف‌روانی پرسشنامه چندوجهی مینه‌سوتا-2 بود که بر اساس مدل راش و روش متل‌هنزل بررسی و مقایسه شد. هدف دیگر ارائه شده در این پژوهش، اثرگذاری این شیوه‌ها بر صحت تحلیل‌های آماری انجام شده بر نحوه پاسخگویی داوطلبان استخدای با توجه به خصیصه جنسیت بود. بر اساس نتایج گزارش شده در جدول 1 می‌توان بیان داشت که تمامی سوالات مولفه ضعف‌روانی و انحراف ضداجتماعی با مدل راش یک پارامتری دارای برازش خوبی است و داده‌ها با مدل برازش داشته است. بنابراین نتایج به دست آمده از کنش‌های افتراقی، مورد اعتماد و اتکاء و قابل استفاده است.

بر اساس آنچه در جدول شماره 2 در بحث تشخیص سوالات دارای کنش افتراقی، مشخص شد که از بین سوالات مقیاس ضداجتماعی 38 درصد و از بین سوالات مقیاس ضعف‌روانی 22/9 درصد سوالات دارای کنش‌افتراقی بر اساس جنسیت بین زن و مرد گزارش شد و بیشتر سوالات به جهت (ضرر) مردان دارای تورش بود. این بدین معناست که مردان داوطلب استخدام به طور معمول بیشتر از زنان در سوالات مقیاس ضداجتماعی و ضعف‌روانی نمره بالاتری گرفته‌اند در صورتی که در واقع این صفت در مردان بیشتر از زنان نبوده است. این تورش سبب شده است تا مردان در گزارش‌های نهائی نمرات بالاتری را در مقیاس ضداجتماعی و ضعف‌روانی به دست آورده و احتمال عدم تأیید سلامت روان آنان در مقایسه با زنان در آزمون‌های سلامت روان بدون ورود

شناسائی سوالات دارای کنش افتراقی، اما عملکرد مدل راش 0/08 درصد بهتر از روش متل‌هنزل بوده است. این درصد از تقسیم سوالات اختلافی در تشخیص کنش افتراقی بر تعداد کل سوالات مورد بررسی بدست آمده است (لینکر، 2010). در شکل شماره 1 و 2 نتایج کنش افتراقی سوالات ضداجتماعی و ضعف‌روانی برحسب ویژگی جنسیت به صورت نمودار نشان داده شد. در ستون آخر جدول جهت تورش سوالات دارای کنش افتراقی نشان داده شده است. اگر در سوال تورش دار، مردان بیشتر جواب مثبت داده باشند، آن سوال با M نشان داده شده است و نشان از این است که این سوال برای مردان دارای کنش افتراقی بیشتری است و اگر در سوال تورش دار، زنان جواب مثبت بیشتر داده بودند، آن سوال با F نشان داده شده است و نشان از این است که زنان به این سوال پاسخ مثبت بیشتری داده‌اند. با توجه به ستون آخر جدول شماره 2 نشان داده شد که در مقیاس ضداجتماعی 12 سوال دارای تورش به جهت مردان و 7 سوال دارای تورش به جهت زنان بوده است و از بین سوالات دارای کنش افتراقی مقیاس ضعف‌روانی 8 سوال به جهت مردان و تنها 3 سوال به ضرر زنان تورش داشته است. در مجموع می‌توان بیان داشت که بیشتر سوالات این دو مقیاس به ضرر مردان دارای سوءگیری بود و مردان به طور معمول نمره بالاتری را نسبت به زنان در دو مقیاس ضداجتماعی و ضعف‌روانی خواهند گرفت.

تأثیر وجود کنش افتراقی بر انصاف آزمون را می‌توان از منظر تفاوت‌های عملکردی شرکت‌کنندگان در پرسشنامه مشخص نمود. وجود پرسش‌هایی با ویژگی کنش افتراقی بدان مفهوم است که سازه‌های غیرمرتبطی به پاسخ افراد تأثیر می‌گذارد و مالک توانایی شرکت‌کنندگان در آن آزمون نیست. در مقیاس ضداجتماعی و ضعف‌روانی پرسشنامه MMPI-2، به دلیل وجود تعداد پرسش‌ها با ویژگی کنش افتراقی می‌توان بیان کرد که 38 درصد سوالات مقیاس ضداجتماعی و 23 درصد از مقیاس ضعف‌روانی متأثر از جنسیت شرکت‌کنندگان بود، که در بیشتر موارد به ضرر مردان و در برخی اوقات به نفع آنان بوده است. پر واضح است که به صورت کلی بخشی از سوالات پرسشنامه حاضر دارای تبعیضی برای شرکت‌کنندگان بوده است و انصاف در آزمون را تا حد زیادی نقض

(Wilson, 2005). به نظر Lee, Zhang, Yin (2010) دقت همبستگی‌های بین خرده مقیاس‌ها و اعتبار آزمون با استفاده از مدل راش بهبود می‌یابد زیرا این مدل خطای اندازه‌گیری را در برآورد به حساب می‌آورد. مدل راش با فراهم آوردن یک روش مناسب برای تبدیل غیرخطی داده‌های خام رتبه‌ای به اندازه‌های فاصله‌ای بر یکی از نگرانی‌ها و دغدغه‌های رایج در ساخت ابزارهای اندازه‌گیری‌های روانی و تربیتی، یعنی دستیابی به عینیت خاص یا تغییر ناپذیری مقایسه‌ها غلبه می‌کند (Minai, 2015; Tennant & Conaghan, 2007). البته در تعمیم‌پذیری نتایج به دست آمده از مدل راش باید دقت لازم را در نظر داشت چرا که مقیاس ضداجتماعی و ضعف روانی پرسشنامه چندوجهی شخصیتی مینه‌سوتا-2 به کار رفته در این پژوهش دارای سازه تک بعدی (unidimensional) است و مفروضات لازم نظریه سوال-پاسخ را رعایت کرده است. اما اگر سازه‌ای مفروضه عدم وابستگی مکانی (local Independence) در آن رعایت نشده باشد، ممکن است در سنجش با شیوه‌های ذکر شده نتایج متفاوتی را بدست آورد. همچنین باید در نظر گرفت اگر در مدل‌های نظریه سوال-پاسخ فراتر از مدل راش رفته و مدل‌های دو یا سه پارامتری را استفاده کنیم، محتمل است که نتایج تغییر یافته و حساسیت نظریه سوال-پاسخ افزایش یابد. همچنین با توجه به نتایج حاصل از این پژوهش به روان‌سنج و روان‌شناسانی که با مراکز طب کار جهت بررسی سلامت روان داوطلبان استخدامی همکاری دارند و به ویژه مجریان آزمون MMPI-2 توصیه می‌شود که نگاهی به کنش‌های افتراقی سوالات بر اساس جنسیت کرده و در تفسیر و نمره‌گذار به این ویژگی‌های دموگرافیک داوطلبان نیز توجه بیشتری نمایند، چرا که مردان به طور طبیعی دارای نمره بالاتری در مولفه ضداجتماعی و ضعف روانی است و این نشانه‌ای بر داشتن میزان بالائی از این صفات در مردان نمی‌باشد. در نهایت به پژوهشگران پیشنهاد می‌شود تا کنش افتراقی مولفه‌های دیگر پرسشنامه چندوجهی مینه‌سوتا مانند پارانوئیا، اسکیزوفرنی و هیپومانیا را در بین داوطلبان استخدامی بررسی نموده و همچنین از روش‌های دیگر بررسی کنش‌های افتراقی نیز در کنار مدل راش بهره ببرند. همچنین به

به سازمان بیشتر بوده است و تا حدی انصاف در گزینش رعایت نشده است. این نتایج با یافته‌های Gori & Marin (2008), Card, Stucky, Sawalani & Little (2015) و McNulty, Forbey, Graham & et al (2015) همسو بود. البته در تبیین این نکته که چرا سوالات زیادی دارای کنش افتراقی تشخیص داده شده است را می‌توان چنین بیان داشت که نمونه مورد هدف این پژوهش، داوطلبان استخدامی بودند و این گروه نمونه با توجه به اینکه در حیطه غیر بالینی و جامعه نرمال قرار گرفتند، شاخص‌های خصیصه‌ای بیشتر از شاخص‌های بالینی نمایان می‌گردد و این یکی از دلایلی است که می‌توان برای بالا بودن تعداد سوالات دارای کنش افتراقی در این مقیاس بیان داشت. البته سوالاتی نیز که دارای کنش افتراقی نبودند نیز می‌توان به عنوان سوالات منصفانه و سوالاتی معرفی کرد که گروه‌های جنسیتی تأثیری در نتایج آنان ندارد و این سوالات به دور از هر گونه سوءگیری ناشی از گروه‌بندی‌های جنسیتی نتایج را گزارش می‌کنند.

در مقایسه شیوه‌های سنجش کنش افتراقی مطرح شد در تشخیص پرسش‌هایی با سطح کنش افتراقی متوسط و یا بالا (37 سوال از 57 سوال) مدل متل‌هنزل تاحدی کارآمدی مناسبی از خود ارئه داد. با این حال نتایج نشان داد که نظریه کلاسیک اندازه‌گیری (متل‌هنزل) نسبت به مدل راش در تشخیص کنش افتراقی توانمند نبود و این روش در تشخیص کنش افتراقی سوالات در سطح متوسط نسبت به سوالات بسیار ساده و بسیار سخت ضعیف تر عمل کرده است و این یافته با مفروضه‌ای مبنی بر اینکه روش متل‌هنزل برای آزمون‌های بسیار ساده و یا بسیار سخت کارآمدتر است، اما برای آزمون‌ها و سوالات با میزان سختی متوسط از کارآمدی آن کاسته می‌شود (Karami & Khodi, 2021). همسو بوده است. همچنین در بحث اینکه سوالات در تشخیص کنش افتراقی غیریکنواخت (Non-uniform) چگونه عمل می‌کنند نیز نیازمند بحث و بررسی بیشتر است.

یافته‌های این پژوهش برای روان‌سنج و متخصصان طراحی آزمون و تحلیل‌گران نتایج آزمون، رویکردهای جدید سنجش کنش افتراقی را توصیه می‌نماید، چرا که مدل‌های راش برتری نسبی خوبی نسبت به نظریه کلاسیک اندازه‌گیری دارد

- July. volume 186, pages347–36
- Card N. A, Stucky. B. D, Sawalani. G. M & Little. T. D. (2008). Direct and indirect aggression during childhood and adolescence: *a meta-analytic review of gender differences, intercorrelations, and relations to maladjustment*. Sep-Oct;79(5): 1185-229. doi: 10.1111/j.1467-8624.2008.01184.x.
- Chapelle, C. A. (2020). Validity in language assessment. *The Routledge Handbook of Second Language Acquisition and Language Testing*, 11.
- Chen, M. Y., Liu, Y., & Zumbo, B. D. (2020). A propensity score method for investigating differential item functioning in performance assessment. *Educational and Psychological Measurement*, 80(3), 476-498.
- Cherry Kendra. (2021) *The Minnesota Multiphasic Personality Inventory (MMPI)*. Updated on September 02 Medically reviewed by Amy Morin, LCSW.
- Embretson, S. E., and Reise, S. P. (2000). Item Response Theory for Psychologists. Mahwah, NJ: Lawrence Erlbaum Associates. 11, 57- 74.
- Gori. E & Marin. R. F. (2015). Rasch Analysis of some MMPI-2 scales in a sample of university freshmen. *International Journal of Arts & Sciences*. 1944-6934 :: 08(03):107–150.
- Hambleton, R. K., & Rogers, H. J. (1989). Detecting potentially biased test items: Comparison of IRT area and Mantel-Haenszel methods. *Applied Measurement in Education*, 2(4), 313-334.
- Hambleton, Ronald K.; Saminathan, H.; Rogers, H. Jane (1991). *Fundamentals of Question-Answer Theory*. Translated by Mohammad Reza Filsafinejad (2010). Tehran: Allameh Tabatabai University. [Persian]
- Karami, H. (2013) The quest for fairness in language testing. *Educational Research and Evaluation*, 19(2&3), 158-169. [Persian]
- Karami, Hossein & Khodi, Ali (2021). Differential Item Functioning and Test Performance: a Comparison Between the Rasch Model, Logistic Regression and Mantel-Haenszel. *Journal of Foreign Language Research*, 10 (4), 842-853. [Persian]
- پژوهشگران توصیه می‌گردد که به بررسی نقاط برش مولفه‌های ضداجتماعی و ضعف‌روانی و مولفه‌های دیگر پرسشنامه MMPI-2 به تفکیک جنسیت پرداخته شود تا نهایت انصاف در آزمون صورت پذیرد.
- منابع
- Antikchi. E, Bigdeli. I. A. & Sabahi. P. (2017). The comparison of neuropsychological index related to executive functions in antisocial personality disorder, obsessive-compulsive personality disorder and normal people. *Advances in Cognitive Science*, Vol. 19, No. 1. [Persian]
- Baker. C. A, Baum. L. J & Francis. J. C. (2023). Assessment of test bias on the MMPI-2-RF higher order and restructured clinical scales as a function of gender and race. *Professional Psychology Research and Practice*. 54 (4) Follow journal. DOI: 10.1037/pro0000517.
- Bethune M. M. (2011). Predictors of Performance in a Professional Counselor Masters Program, *Doctoral Thesis*  
<http://gradworks.umi.com/34/93/3493607.html>.
- Björkqvist, K. (2018). Gender differences in aggression. *Current Opinion in Psychology*. Volume 19, February, Pages 39-42.
- Buribayev. Y. A and Khamzina Z. A. (2019). Gender equality in employment: The experience of Kazakhstan. *International Journal of Discrimination and the Law*, Vol. 19(2) 110–124.
- Camilli, G. (2006). *Test fairness*. In R. Brennan (Ed.), *Educational measurement* (4th ed.) (pp. 221-256). New York: American Council on Education & Praeger series on higher education.
- Camilli, G., & Congdon, P. (1999). Application of a method of estimating DIF for polytomous test items. *Journal of Educational and Behavioral Statistics*, 24(4), 323-341.
- Camilli, G., & Shepard, L. (1994). *Methods for identifying biased test items*. Newbury Park, CA Sage.
- Cao. W, Li. P. Vander Wal. R.C & Taris. T.W. (2022). Leadership and Workplace Aggression: A Meta-analysis. *Journal of Business Ethics*. Published: 15

- Multilevel, Longitudinal and Structural Equation Models*: Chapman & Hall/ CRC Press.
- Soltani Shal R, Saadatbin Javaheri F, Zebardast F. (2020). Survey the level of well-being and Psychometric characteristics of hospital nurses' well-being at work scale. *Occupational Medicine Quarterly Journal*:12(1): 55-68. [Persian]
- Su. Y & Wang N. (2005). Use of the Rasch IRT model in standard setting: An item- mapping method. *Journal of Educational Measurement*. Sep; 40 (3): 231- 53.
- Talerico G. M, McCallum. J. J, Whitman M. R, Tarescavage. M, Corey D. M & Ben-Porath. Y. B. (2023): Comparing the Validity of MMPI-3 Scores in Prehire Psychological Screenings of Male and Female Police Officer Candidates, *Journal of Personality Assessment*, DOI: 10.1080/00223891.2023.2191278
- Tennant, A., & Conaghan, P.G. (2007). The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatism journal*. 57, 1358– 1362.
- Valianpour. Z, Modarres Gharavi. M & Mahram. B. (2020). Validation of the Minnesota Multiphasic Personality Inventory (MMPI 2) in psychiatric patients and non-patient individuals in Mashhad city, Iran. *Journal of Fundamentals of Mental Health*, Nov-Dec. [Persian]
- Wiberg, M. (2007). Measuring and detecting differential item functioning in criterion-referenced licensing test: A theoretic comparison of methods. *Educational Measurement*, technical report No. 2.
- Zahid. Z, nasrollahi. Z & Mahinizadeh. M. (2020). From Gender Discrimination to Equality and Economic Growth (Study of Developing Countries). *Journal of Woman in Development and Politics*. 662- 643. [Persian]
- Zumbo, B. D. (2003). Does item-level DIF manifest itself in scale-level analyses? Implications for translating language tests. *Language testing*, 20(2), 136-147.
- Karami. H. R, Gramipour. M & Minaei. A. (2021). Differential Item Functioning (DIF) Detection Rate Using Rasch Trees Model: A Simulated and Real Data Study of the NAJA High Stakes Tests. *Educational Measurement*. 11(44), 1-30. [Persian]
- Lee JC, Zhang Z, Yin H. (2010). Using multidimensional Rasch analysis to validate the Chinese version of the Motivated Strategies for Learning Questionnaire (MSLQ-CV). *Eur J Psychol Educ*; 25.
- Linacre, J. M. (2010). *A User's Guide to WINSTEPS®*. Retrieved May 2, from <http://www.winsteps.com/>.
- Lord, F. M.y. & Novick, MR. (1968). *Statistical Theory of Mental Test Scores*. Addison-Wesley Publishing Company.
- Marnat, G. G. & Wright, A. J. (2016). *Handbook of Psychological Assessment*. 6th ed.
- McNulty J. L., Forbey J. D., Graham J. R., Ben-Porath Y. S., Black M. S., Anderson S. V. & Burlew A. K. (2015). MMPI-2 Validity Scale Characteristics in a Correctional Sample. *Sage Journals*, Volume 10, No. 3, September. 288-298.
- Minaei. A. (2015). Application of Rasch measurement model to evaluate measurement properties of the Test of Visual- Motor Skills-Revised. *Journal of Educational Measurement*. Volume 5, Issue 18 - Serial Number 18. January. 77-114. [Persian]
- Penfield, R. D. & Algina, J. (2003). Applying the Liu–Agresti estimator of the cumulative common odds ratio to DIF detection in polytomous items. *Journal of Educational Measurement*, 40: 353–370.
- Queirolo L, Bacci C, Roccon A, Zanette G and Mucignat C. (2023). Anxiety in a regular day of work: A 24 hour psychophysiological investigation in young dentists with gender comparison. *Front Psychol*. 14:1045974. doi: 10.3389/fpsyg.1045974
- Siegert, R. J., Tennant, A., & Turner-Stokes, L. (2010). Rasch analysis of the Beck Depression Inventory-II in a neurological rehabilitation sample. *Disability and Rehabilitation*, 32(1), 8–17.
- Skronidal, A., Rabe-Hesketh, S., and Boca Raton, F. (2004). *Generalized Latent Variable Modeling*: