

ویژگی‌های روانسنجی مقیاس افسردگی نوجوانان براساس نظریه سوال- پاسخ و مقایسه نتایج با نظریه کلاسیک آزمون

قاسم کشاورز گرامی: دانشجوی دوره دکتری، گروه سنجش و اندازه گیری، دانشکده روانشناسی و علوم تربیتی، دانشگاه علامه طباطبائی، تهران، ایران-نویسنده
رابط: ghaseem_keshavarz@yahoo.com

افضل اکبری بلوطبگان: دانشجوی دوره دکتری، گروه روانشناسی و علوم تربیتی، دانشکده علوم تربیتی و روانشناسی، دانشگاه سمنان، سمنان، ایران
خدیجه باباخانی: دانشجوی دوره دکتری، گروه روانشناسی سلامت، دانشکده روانشناسی و علوم تربیتی، دانشگاه آزاد اسلامی کرج، کرج، ایران
تاریخ دریافت: ۱۳۹۵/۱۲/۲۱ تاریخ پذیرش: ۱۳۹۶/۴/۱۱

چکیده

زمینه و هدف: هدف از مقاله حاضر بررسی ویژگی‌های روانسنجی مقیاس افسردگی نوجوانان براساس نظریه سوال- پاسخ و مقایسه نتایج به دست آمده با نظریه کلاسیک آزمون‌سازی می‌باشد.

روش بررسی: مقیاس محقق‌ساخته شامل ۴۶ سؤال در بین ۷۵۰ دانش‌آموز راهنمایی و دبیرستان شهر زنجان (۳۶۴ پسر و ۳۸۶ دختر) که به صورت خوشه‌ای چندمرحله‌ای (طبقه‌ای متناسب با حجم) و تصادفی برگزیده شده بودند، اجرا شد.

نتایج: نتایج تحلیل عاملی و ساختار سؤالات نشان داد پیش‌فرض‌های مهم نظریه سوال- پاسخ یعنی تک بعدی بودن و استقلال موضعی در مقیاس سنجش افسردگی نوجوانان برقرار است. در مقایسه ویژگی‌های روانسنجی سؤالات براساس دو نظریه نتایج مشابهی مشاهده شد با این تفاوت که در نظریه سوال- پاسخ توانایی برآورد پارامترهای معتبرتری برای ارزیابی سؤالات وجود داشت. ویژگی‌های روانسنجی آزمون حاصل از دو نظریه نیز مقادیر نزدیکی برای همسانی درونی نشان دادند، به طوری که مقدار آلفای کرونباخ ۰/۹۳۳ و اعتبار حاشیه‌ای ۰/۹۴۷ به دست آمد. نمرات هنجاری براساس تتا در نظریه سؤال-پاسخ و نمرات تی در نظریه کلاسیک محاسبه شد. در رتبه‌بندی افراد توسط دو نظریه روی پیوستار نمرات افسردگی تفاوت‌های قابل توجه‌ای مشاهده شد که می‌توان به تفاوت‌های ۹۳، ۹۲ و ۸۶ نمری در رتبه‌ها اشاره کرد.

نتیجه‌گیری: بررسی و مقایسه نتایج نشان داد مدل‌های نظریه سوال- پاسخ پارامترها و شاخص‌های بیشتر و دقیقتری برای قضاوت و تصمیم‌گیری در مورد سؤالات، آزمون و افراد ارائه می‌دهند. بر این اساس کاربرد بیشتر نظریه سوال-پاسخ در سنجش دیگر سازه‌های روانشناسی پیشنهاد می‌شود.

واژگان کلیدی: نظریه سوال-پاسخ، افسردگی، نوجوانی، نظریه کلاسیک آزمون

مقدمه

قرن پیش افسردگی را نتیجه سیلان صفرا در مغز دانست در حقیقت این اصطلاح ترجمه واژه ملانکولی (Melankoli) است و در حدود ۱۰۰ سال قبل از میلاد نیز کورنیوس سلسوس (Cornelius Celsus) افسردگی را ناشی از

افسردگی (Depression) در ادبیات بشر تاریخچه‌ای طولانی دارد. در نوشته‌های به‌جامانده از فرهنگ‌های مختلف، تاریخچه‌ای از افسردگی مشاهده می‌شود. بقراط حدود ۲۵

روانی - اجتماعی و هنجاری نوجوان را به تأخیر اندازد. به علاوه، این افسردگی اغلب در صورت عدم درمان به دوران بزرگسالی تداوم می‌یابد و نوجوانان افسرده در خطر بالاتری برای سوءاستفاده از مواد و بیماری‌های روان‌پزشکی دیگر هستند (Paus et al. 2008). پس افسردگی افت فراگیر خلق است که همراه با احساس غمگینی و عدم توانایی تجربه لذت می‌باشد (Abramowitz and Mahaffey 2012).

با توجه به اهمیت و شیوع افسردگی و پیامدهای ناگوارش که گوشه‌ای از آن بیان شد، پژوهش‌های بیشتری در جهت شناخت، کنترل و درمان آن، لازم و ضروری می‌باشد. انجام پژوهش در این حوزه نیازمند ابزارها و پرسشنامه‌های معتبر است که با نظریه‌های قوی و کارآمد سنجش و اندازه‌گیری ساخت و هنجاریابی شده باشند و اغلب پرسشنامه‌های موجود بر اساس تئوری کلاسیک اندازه‌گیری تجزیه-تحلیل و هنجاریابی شده‌اند، درحالی‌که تئوری‌های جدید سنجش و اندازه‌گیری مانند مدل‌های سوال-پاسخ به شکل و شیوه اساسی‌تری می‌توانند به کمی‌سازی صفت‌های مکنون بپردازند. "نظریه سوال-پاسخ" (IRT)، شامل خانواده‌ای از مدل‌های ریاضی است که روابط تابعی بین متغیرهای مشاهده‌پذیر و سازه‌های صفات زیربنایی این متغیرها را نمایش می‌دهد. Lord در سال ۱۹۵۲ مقاله‌ای منتشر کرد و در آن نظریه منحنی ویژه سؤال را به عنوان یک مدل یا نظریه مطرح کرد. وی در همین سال، پژوهشی را در این زمینه ترتیب داد و نشان داد که استفاده از تئوری سؤال-پاسخ برای سؤالات چندگزینه‌ای مناسب است. در سال ۱۹۶۰ Rasch مدل تک پارامتری را مطرح ساخت که انگیزه و علاقه وافر را برانگیخت. در کل چارچوب IRT گروهی از مدل‌ها را شامل می‌شود و کاربرد هر مدل به موقعیت‌های خاص و به آیت‌های آزمون مربوط است و همچنین فرض‌های نظری متفاوتی درباره آیت‌های آزمون دارد (Birnbaum 1968). در طول این سال‌ها آزمون‌های بااهمیتی بر اساس نظریه سوال-پاسخ تحلیل و هنجاریابی شده‌اند. یکی از این

صفرای سیاه معرفی نمود. در قرون وسطی نیز طبابت در سرزمین‌های اسلامی رونق داشت، به طوری که رازی، ابن سینا و پزشکان دیگر ملانکولی را بیماری مشخصی می‌دانستند. سرانجام کریپلین (Cripelin) بود که آن را "روان‌گسستگی آشفتگی-افسردگی" (Manic-Depressive Psychosis) نامید. اما این حالت افسردگی روان‌گسستگی، انواع دیگر افسردگی‌ها را پوشش نمی‌دهد و فقط می‌توان آن را به عنوان یک زیرگروه افسردگی‌ها در نظر گرفت (Akbari 2013). بنابراین اختلالات افسردگی و اضطرابی طبق تخمین سازمان جهانی بهداشت، در رأس بیماری‌های روانی قرار می‌گیرند و حدود ۲۵ درصد مراجعین به مراکز بهداشتی در جهان را به خود اختصاص می‌دهند و پیش‌بینی می‌شود این اختلال از لحاظ فراوانی در سال ۲۰۲۰ در مرتبه دوم بیماری‌های وخیم قرار گیرد (Lambert 2006). افسردگی یکی از پرهزینه‌ترین و ناتوان‌کننده‌ترین مشکلات پزشکی در افراد جامعه ما است و می‌تواند به مرگ نابهنگام به وسیله خودکشی منجر شود (Moussavi et al. 2007). از سویی در مورد افسردگی نوجوانی، تا سال‌های ۱۹۶۰ پزشکان بر این تصور بودند که امکان بروز بیماری افسردگی در کودکان و نوجوانان وجود ندارد اما بعضی مطالعات نشان داده است که بین ۳ تا ۶ میلیون آمریکایی زیر ۱۸ سال به بیماری افسردگی مبتلا هستند و افسردگی درمان نشده نوجوانان، به اندازه افسردگی سالمندان خطرناک است (Sadock and Sadock 2010). در واقع افسردگی کودکی و نوجوانی به طور ناگهانی بعد از ۴۰ تا ۵۰ سال افزایش یافته است و یکی از رایج‌ترین ناراحتی‌های روانی در بین دختران و پسران نوجوان می‌باشد. نوجوانان افسرده احساس ناامیدی و احساسات منفی شدید دارند. همه را طرد می‌کنند، دوست ندارند با دیگران آشنا شوند و در موارد شدید میل دارند خود را از بین ببرند (Reynolds 2010). افسردگی در نوجوانی به طور معنی‌داری بر کارکردهای اجتماعی، آموزشی، خانواده و حوزه‌های شغلی آسیب وارد ساخته و امکان دارد رشد

پاسخ‌های خوب یا بد او در دیگر سؤالات باشد (Junker and Sijtsma 2001).

نظریه سؤال- پاسخ به جای تأکید بر نمرات کل آزمون، بر پاسخ‌های آزمودنی‌ها به تک‌تک سؤالات آزمون تکیه می‌کند. در نظریه سؤال- پاسخ با استفاده از مدل‌های ریاضی می‌توان احتمال پاسخ درست به یک سؤال آزمون را به‌عنوان تابعی از توانایی آزمودنشونده به‌حساب آورد و همچنین برخی ویژگی‌های سؤال را پیش‌بینی کرد. ویژگی‌هایی که برای سؤالات یا ماده‌های آزمون بدست می‌آیند به نوع مدل یا الگوی نظریه سؤال-پاسخ وابسته‌اند. مدل‌های نظریه سوال-پاسخ می‌توانند نمره‌های دو مقوله‌ای و چند مقوله‌ای را تحلیل کنند به‌شکلی که طبقه‌بندی امتیازات می‌تواند منظم یا نامنظم باشند (Gierl and Lai 2013).

نظریه کلاسیک آزمون و نظریه سؤال-پاسخ چارچوب‌های متفاوتی برای سنجش ارائه می‌دهند. اگرچه در گذشته نظریه کلاسیک بیشتر از نظریه سؤال-پاسخ استفاده می‌شد، اما در دهه‌های اخیر نظریه سؤال-پاسخ جایگاه خود را به دست آورده است. نظریه سوال-پاسخ برای توصیف سؤالات و آزمون، مقایسه آزمون‌ها، سنجش توانایی‌ها، نگرش نگرش‌ها و سازه‌های دیگر به‌کاربرده می‌شود. از طرف دیگر "نظریه کلاسیک آزمون" (CTT)، الگوی ساده و کاملاً مفیدی است که می‌تواند نحوه تأثیر خطاهای اندازه‌گیری را بر نمره‌های مشاهده‌شده توصیف کند. می‌توان گفت که مهم‌ترین امتیاز نظریه کلاسیک اندازه‌گیری این است که بر فرض‌های نظری کمتری استوار است که همین ویژگی کاربرد آن را در بسیاری از موقعیت‌های اندازه‌گیری آسان‌تر می‌کند. از مهم‌ترین ایرادات وارد بر نظریه کلاسیک اندازه‌گیری وابسته بودن ویژگی‌های روان‌سنجی آزمون بر گروه نمونه وابسته است، یعنی مقادیر آن‌ها به ویژگی‌های نمونه وابسته است. درحالی‌که در نظریه سوال-پاسخ علاوه بر وابسته نبودن ویژگی‌های روان‌سنجی آزمون به گروه نمونه،

آزمون‌ها، آزمون HEXACO بوده که توسط Ashton and Lee در سال ۲۰۰۷ در زمینه شخصیت ارائه کرده بودند. در این آزمون شش بعد شخصیتی به دست آوردند و این ابعاد شخصیتی را هگزاکو نامیدند. آزمون Navy Computer Adaptive Personality Scales یک ابزار جدید برای ارزیابی ویژگی‌های شخصیتی است که برای طبقه‌بندی و غربالگری افراد در مشاغل نیروی دریایی ارتش آمریکا بکار برده می‌شود. تحلیل مشاغل بر اساس داده‌های دودویی در آزمون سازمان اونت (O*NET) نیز یکی دیگر از کاربرد مدل‌های نظریه سوال-پاسخ بوده است که توسط هاروی (Harvey 2003) صورت گرفته است. آزمون افسردگی بک (Beck Depression Inventory) یکی دیگر از آزمون‌هایی است که با استفاده از این روش موردبررسی قرار گرفت. به‌منظور تأیید ویژگی‌های روان‌سنجی آزمون افسردگی بک از مدل نظریه سوال-پاسخ استفاده شد که این موضوع در سال ۲۰۱۵ توسط کاروالو و همکاران صورت پذیرفت. پرسشنامه کیفیت زندگی از دیدگاه سازمان بهداشت جهانی (World health organization quality of life) و آزمون شخصیت استخدامی طراحی‌شده توسط انجمن روانشناسان بریتانیا (BPS) موارد دیگری از کارنامه نظریه- سؤال پاسخ می‌باشند.

در منابع گوناگون چندین مفروضه برای مدل‌های نظریه سوال-پاسخ مطرح‌شده که در تمامی آن‌ها تک‌بعدی بودن (Unidimensionality) و استقلال موضعی (Local indepedence) مهم‌ترین و اساسی‌ترین مفروضه‌ها معرفی شده‌اند. از دیگر مفروضه‌ها می‌توان به برازش داده‌ها با مدل اشاره نمود. تک‌بعدی بودن به این معنا است که همه سؤالات آزمون فقط یک حوزه توانایی یا دانش را اندازه بگیرد. مفروضه استقلال موضعی بیان می‌کند که پاسخ‌های آمودنی‌ها به سؤالات آزمون از لحاظ آماری مستقل از یکدیگرند و فقط سطح توانایی آزمودنی به‌عنوان عامل اصلی به‌حساب آید. اگر این پیش‌فرض برقرار باشد عملکرد آزمودنی نباید تحت تأثیر

امکان اجرای تجزیه و تحلیل های بیشتر و پیشرفته تر نیز امکان پذیر است (Embretson and Reise 2013).

بررسی سوگیری سؤالات و برازش سؤالات ازجمله امکانات نظریه سؤال-پاسخ می باشند. بحث نحوه برخورد با ارزیابی پایایی و روایی امری ضروری در این نظریه هست. پایایی، توانایی فرد یا گروه را به صورت پیوسته ارزیابی می کند و روایی می گوید آزمون باید آن چیزی را بسنجد که برای اندازه گیری آن ساخته شده است. نظریه کلاسیک اندازه گیری روش های به نسبت ساده ای را برای تعیین پایایی و روایی آزمون دارا است. در تحلیل کلاسیک از طریق بررسی نمرات کل آزمون و روش های آماری ساده ای، روایی و اعتبار تعیین می شود. اما نظریه سؤال-پاسخ روش های پیچیده تری برای تعیین روایی و پایایی آزمون ارائه داده است. اگرچه نظریه کلاسیک اندازه گیری بر کل آزمون تمرکز کرده، ولی نظریه سؤال-پاسخ بر هر ماده و هر فرد آزمودنی تمرکز می کند (Morales 2009).

در نظریه سوال-پاسخ به جای خطای معیار اندازه گیری، از مفهوم شاخص دقت اندازه گیری استفاده می شود. یک تفاوت مهم بین این دو مفهوم این است که شاخص دقت اندازه گیری در طول مقیاس نمرات دارای اندازه های مختلف است، در حالی که خطای معیار اندازه گیری در طول مقیاس نمرات همواره دارای اندازه های یکسان می باشد. به بیان دیگر شاخص دقت اندازه گیری در نظریه سوال-پاسخ در طول پیوستار تا متغیر می باشد. شاخص دقت اندازه گیری تنها به روش های پایایی مربوط به همسانی درونی آزمون ارتباط دارد. اگر هدف از تعیین پایایی ثبات نمرات در طول زمان باشد، در آن صورت روش باز آزمایی یا روش فرم های موازی به کار برده می شود که کاربرد آن در هر دو نظریه سوال-پاسخ و نظریه کلاسیک اندازه گیری یکسان است (Sayf 2016).

نظریه سؤال-پاسخ شامل خانواده ای از مدل ها است و کاربرد هر مدل به موقعیت های خاص و به آیت های آزمون

بستگی دارد و همچنین برخی از مدل ها، فرض های نظری متفاوتی درباره آیت های آزمون دارند. Hambleton اظهار می دارد که هیچ گونه محدودیتی در تعداد مدل هایی که می توان در قالب نظریه سؤال-پاسخ ارائه نمود وجود ندارد. مدل سؤال-پاسخی که در این پژوهش به کار رفته، مدل پاسخ مدرج (GRM) نام دارد. مدل پاسخ مدرج در سال ۱۹۶۹ توسط Samejima ارائه گردید و توانایی تجزیه و تحلیل پاسخ های طبقه ای مانند طیف لیکرت را دارد. این مدل در موقعیت هایی قابل کاربرد است که میزان تمایل یا نگرش آزمودنی ها با یک عبارت و چند مقوله بررسی می شود. در این رابطه میزان نگرش و نظر آزمودنی ها در طول یک پیوستار با مقوله های گسسته بیان می شود (Samejima et al. 2005).

بنابراین، هدف از انجام پژوهش حاضر بررسی خصوصیات روان سنجی مقیاس افسردگی در نوجوانان بر اساس نظریات کلاسیک آزمون و نظریه سؤال-پاسخ بوده است.

روش کار

طرح پژوهش حاضر توصیفی و از نوع مطالعات روان سنجی است. جامعه آماری پژوهش شامل کلیه دانش آموزان دختر و پسر شهر زنجان در دوره متوسطه اول و دوم که در سال تحصیلی ۹۱-۱۳۹۰ مشغول به تحصیل بوده اند، تشکیل شده است. برای مدل پاسخ مدرج نظریه سوال-پاسخ یافته های Jiang and Jweiss همراه با خطایی که حتی برای نمونه های بزرگ تر مدام کاهش پیدا می کند حداقل ۵۰۰ آزمودنی پیشنهاد می کند (Jiang et al. 2016).

آزمودنی ها ۷۵۰ نفر در پایه های تحصیلی هفتم تا دوازدهم (۳۶۴ پسر و ۳۸۶ دختر) بودند که از ۸ مدرسه متوسطه اول و دوم با روش نمونه گیری خوشه ای چند مرحله مرحله ای (طبقه ای و متناسب با حجم) انتخاب شدند. ۱۹۱ نفر از آزمودنی ها پسر در دوره اول متوسطه و ۱۷۳ نفر از

محورهای اصلی استفاده شد. پیش از اجرای تحلیل عاملی شاخص کفایت نمونه برداری (KMO) و مقدار آزمون کرویت بارتلت محاسبه گردید. نتایج نشان داد مقدار KMO برابر با ۰/۹۴۹ و مقدار مجذور کای کرویت بارتلت برابر با ۱۰۱۱۲ بدست آمد که با درجه آزادی ۹۹۰ معنی دار ($p \leq ۰/۰۰۰۱$) بود. بدین ترتیب علاوه بر کفایت نمونه برداری، اجرای تحلیل عاملی بر پایه ماتریس مورد مطالعه نیز قابل توجیه بود. برای تعیین اینکه مقیاس افسردگی نوجوانان از چند عامل اشباع شده، شاخص های ارزش ویژه، نسبت واریانس تبیین شده و نمودار اسکری (شکل ۱) مورد بررسی قرار گرفت. بر این اساس عامل استخراج شد که این عامل توانسته است ۲۶/۲۵۱ درصد از واریانس کل آزمون را تبیین نماید. ارزش ویژه این عامل ۱۱/۸۱۳ بوده است. Reckase عنوان می کند در صورتی که عامل اول حداقل ۲۰٪ از کل واریانس تست را تبیین کند و یا مقدار ویژه عامل اول حداقل ۳ برابر عامل بعد از خود باشد پیش فرض تک بعدی بودن در حد قابل قبولی برقرار است (Reckase 2006).

مقیاسه نظریه کلاسیک آزمون و نظریه سؤال-پاسخ در مقیاس افسردگی نوجوانان: در ادامه پژوهش خصوصیات روان سنجی سؤالات بر اساس نظریه سؤال-پاسخ و نظریه کلاسیک محاسبه شده است که خلاصه ای از نتایج آن در جدول ۲ گزارش شده است. در این جدول، ۱۰ سؤال قوی و ۱۰ سؤال ضعیف از پرسشنامه ۴۵ سؤالی افسردگی نوجوانان گزارش شده است. همان گونه که مشاهده می شود در قسمت محاسبات کلاسیک توان افتراقی (ضریب تمیز) و ضریب دشواری و در قسمت محاسبات سؤال-پاسخ ضریب تمیز (شیب منحنی سؤال) و ضرایب آستانه ای گزارش شده است. در جدول ۲ مشاهده می شود که سؤالات ۶، ۲۸ و ۳۱ با مقادیر ۰/۷۵، ۰/۷۴ و ۰/۷۹ دارای کمترین توان های افتراقی در بین سؤالات هستند. سؤالات ۴۱، ۱۸ و ۱۳ با مقادیر ۰/۲۰۳، ۰/۲۰۶ و ۰/۲۱۱ دارای بیشترین توان افتراقی در بین

آزمودنی های پسر در دوره دوم متوسطه قرار داشتند. همچنین ۱۸۴ نفر از آزمودنی های دختر در دوره متوسطه اول و ۲۰۲ نفر از آزمودنی های دختر در دوره دوم متوسطه قرار داشتند. در کل نمونه بیشترین فراوانی به دانش آموزان پایه یازدهم با تعداد ۱۳۱ نفر (۱۷/۴۷) و کمترین فراوانی به دانش آموزان پایه دوازدهم (سوم دبیرستان) با تعداد ۱۱۷ نفر (۱۵/۶۰) اختصاص دارد. در جدول ۱ گروه نمونه به تفکیک جنسیت و مقطع تحصیلی توصیف شده است.

ابزارهای اندازه گیری ۱- پرسشنامه افسردگی نوجوانان: این ابزار به منظور سنجش افسردگی در نوجوانان توسط پژوهشگران ساخته شده است. دارای ۴۶ آیتم بوده که افسردگی را به صورت سازه ای تک عاملی اندازه گیری می کند. آیتم ها بر روی یک طیف لیکرت ۵ گزینه ای از ۱=هرگز تا ۵=تقریباً هرروز قرار دارد. پس از ساخت ابزار به منظور بررسی ویژگی های روان سنجی آن، بر روی یک نمونه ۳۰ نفری از افراد اجرا شد. نتایج نشان داد ابزار از قابلیت مطلوبی برخوردار است و آیتم هایی که قابل درک نبودند با واژگان مناسب تر معادل سازی شدند. در نهایت نسخه نهایی ابزار اجرا و مورد تجزیه و تحلیل قرار گرفت و نتایج نشان داد این ابزار از ساختار یک عاملی برخوردار است. برای بررسی اعتبار (پایایی) ابزار از روش های آلفای کرونباخ ۰/۹۳۲ و باز آزمایی ۰/۷۵ استفاده شده که نتایج رضایت بخشی بدست آمده و نشان از قابلیت بالای ابزار می باشد.

برای تجزیه و تحلیل اطلاعات از روش های تحلیل اعتبار، همبستگی پیرسون، تحلیل عاملی اکتشافی و نیز تجزیه و تحلیل داده ها به روش سؤال- پاسخ به کمک نرم افزارهای SPSS19 و Multilog7 استفاده شده است.

نتایج

تحلیل عاملی اکتشافی: به منظور تعیین عاملی و مطالعه ویژگی های ۴۶ آیتم پرسشنامه افسردگی نوجوانان از تحلیل عاملی اکتشافی استفاده شد. برای استخراج عامل ها از تحلیل

نظریه سوال-پاسخ و کلاسیک اندازه‌گیری در سنجش صفت افسردگی نوجوانان صورت گرفت. همان‌گونه که در جدول ۵ مشاهده می‌شود رتبه آزمودنی ۴۰ در نمره هنجار کلاسیک ۲۷۲ است و در تتای نظریه سوال-پاسخ ۱۷۹ است. بین دو رتبه ۹۳ نفر اختلاف وجود دارد. شرکت‌کننده‌های ۱، ۳۸ و ۳۹ نیز دارای تفاوت‌های قابل توجهی هستند.

همان‌طور که مشخص است در نظریه کلاسیک آزمون (CTT) دشواری سؤالات با میانگین نمرات در هر سؤال مشخص می‌شود، اما در نظریه سوال-پاسخ (IRT) برای هر سؤال به تعداد گزاره‌ها یک آستانه تعریف می‌شود. زمانی که سؤالات یک پرسشنامه تهیه‌شده و اجرای مقدماتی انجام شد می‌توان از دو نظریه موجود استفاده کرد.

در شکل ۲ ماتریسی از ۴۵ سؤال آزمون افسردگی نوجوانان گزارش‌شده است که از خصوصیات ویژه برنامه مالتی لوگ می‌باشد. بر این اساس می‌توان یک قضاوت کلی از پرسشنامه ارائه داد. مثلاً سؤالات ۲۲ و ۳۵ سؤالات نامناسب و سؤالات ۲۰ و ۴۱ سؤالات مناسبی می‌باشند. افزون بر آن در شکل ۳ نمودار آگاهی‌دهندگی و خطای استاندارد اندازه‌گیری کل پرسشنامه در سطح تتای ۳- تا ۳ گزارش شده است. همان‌طور که در این شکل مشاهده می‌شود، این آزمون در سطح تتای ۰ تا ۱ بیشترین میزان آگاهی‌دهندگی را دارد.

روایی همگرا و واگرایی مقیاس افسردگی نوجوانان: به‌منظور بررسی روایی مقیاس افسردگی نوجوانان از آزمون‌های افسردگی بک، اضطراب زانک و عزت‌نفس کوپر اسمیت استفاده‌شده است. نتایج ارائه‌شده در جدول ۶ نشان می‌دهد که بین مقیاس افسردگی نوجوانان با افسردگی بک و اضطراب رابطه مثبت و معنی‌داری وجود دارد. این در حالی بود که بین افسردگی نوجوانان با مقیاس عزت‌نفس کوپر اسمیت رابطه منفی و معنی‌داری مشاهده شد.

سؤالات هستند. سؤال ۱۰ با مقدار ۲/۴۶ دارای بیشترین ضریب دشواری و سؤال ۲۸ با مقدار ۰/۵ دارای کمترین ضریب دشواری است. همچنین در این جدول مشاهده می‌شود که سؤالات ۶، ۲۲ و ۳۴ با مقادیر ۰/۵۱، ۰/۵۲ و ۰/۵۶ دارای کمترین ضرایب تمیز در بین سؤالات هستند. سؤالات ۴۱، ۱۸ و ۱۹ با مقادیر ۱/۶۳، ۱/۶۸، ۱/۷۵ دارای بیشترین ضرایب تمیز در بین سؤالات هستند. همچنین سؤالات ۴۱، ۱۸ و ۱۹ دارای مناسب‌ترین ضرایب آستانه‌ای می‌باشند.

افزون بر آن در جدول ۳ محاسبات سوال-پاسخ مقادیر آگاهی‌دهندگی سؤالات گزارش‌شده است. در این جدول ۵ سؤال ضعیف و ۵ سؤال قوی بر اساس میزان آگاهی‌دهندگی بر روی ۷ نقطه از پیوستار تتا گزارش‌شده‌اند. همان‌گونه که مشاهده می‌شود، سؤال ۱۹ و ۱۸ بیشترین آگاهی‌دهندگی را داشته است. سؤال ۶ نیز در بین تمامی سؤالات پرسشنامه نامناسب‌ترین آگاهی‌دهندگی را نشان داده است.

در مرحله بعد نمرات هنجار بر اساس محاسبات کلاسیک و نظریه سوال-پاسخ محاسبه شد که خلاصه‌ای از این نمرات در جدول ۴ گزارش‌شده است. همان‌گونه که در این جدول مشاهده می‌شود تمامی نمرات به تفکیک جنسیت نیز گزارش‌شده است. میانگین کل نمرات خام ۱۱۳/۶ بدست آمد که معادل نمره هنجار آن صفر و میانگین تتای آن ۰/۵۴- است. میانگین نمرات افسردگی دختران و پسران تفاوتی ۶ نمره‌ای دارند. کمترین نمره خام ۴۵ بدست آمده که معادل نمره هنجار آن ۰/۳۱۴- و معادل تتای آن ۳/۷۸- است. بیشترین مقدار نمره خام ۲۲۵ بدست آمد که معادل نمره هنجار آن ۳۱۴ و معادل تتای آن ۳/۷۸ بدست آمد. دامنه تغییرات نمرات خام ۱۸۰، دامنه تغییرات نمرات نرمال ۶/۲۹ و دامنه تغییرات نمرات تتا ۷/۵۶ بدست آمد.

در آخرین مرحله از تحلیل رتبه‌های افراد بر اساس نمرات هنجار Z و نمرات توانایی تتا محاسبه شد و بخشی از نتایج آن که مربوط به ۴۰ آزمودنی می‌باشد در جدول ۵ گزارش‌شده است. این تحلیل به‌منظور مقایسه تفاوت دو

بحث

افسردگی یکی از اختلالاتی است که در بین نوجوانان شایع است. از این رو سنجش این سازه با روش های معتبر حائز اهمیتی فراوان است. بر این اساس هدف از پژوهش حاضر سنجش میزان افسردگی نوجوانان بر اساس نظریه سوال-پاسخ و مقایسه نتایج آن با نظریه کلاسیک آزمون می باشد. برای دستیابی به این هدف، پرسشنامه محقق ساخته افسردگی نوجوانان انتخاب شد تا براساس داده های تجربی و دارای ارزش کاربردی به اعتبار یافته های پژوهش افزوده شود. Galton عنوان می کند هر پدیده ای در هر رشته علمی در صورتی می تواند از اعتبار علمی برخوردار باشد که قابل اندازه گیری باشد، و به مقدار تبدیل شود. در واقع سال-هاست که روانشناسان می کوشند مشاهدات خود را بر اساس اندازه گیری و کمی سازی اعتبار ببخشند (Everett 2013).

در این راستا متخصصان روان سنجی تلاش های قابل توجهی نمودند که نظریه سوال-پاسخ نقطه عطف این تلاش ها می باشد. همان گونه که ذکر شد تک بعدی بودن و استقلال موضعی پیش فرض های مهم نظریه سوال-پاسخ هستند. تک بعدی بودن برحسب وابستگی آماری در میان سوال ها تعریف می شود. لازمه تک بعدی بودن آزمون آن است که وابستگی آماری در میان سوال ها بتواند تنها به وسیله یک صفت مکنون توضیح داده شود. پیش فرض استقلال موضعی به این معناست که در صورت ثابت نگه داشتن سطح توانایی سوال های آزمون از یکدیگر مستقل هستند به عبارت دیگر احتمال پاسخ درست به یک سوال بر اساس سطح توانایی آزمودنی در صفت مکنون مشخص می شود (Mc Donald 2013).

بر این اساس در اولین مرحله از تحلیل ها مفروضه های اصلی نظریه سوال-پاسخ مورد بررسی قرار گرفتند. برای بررسی تک بعدی بودن از تحلیل عاملی خطی با روش محورهای اصلی استفاده شد. قابل ذکر است؛ مبنای محاسبات تحلیل عاملی در برنامه اس.پی.اس.اس ماتریس ضریب

همبستگی پیرسون است که این ضریب همبستگی برای سوالات با طیف لیکرت که مقیاس اندازه گیری آنها رتبه ای می باشد مناسب نیست. به همین دلیل ماتریس ضرایب همبستگی بر پایه ضریب پلی کوریک با کمک برنامه لیزرل محاسبه شد و با استفاده از قابلیت دستورنویسی برنامه اس-پی.اس.اس وارد محاسبات تحلیل عاملی گردید.

در روش کلاسیک اندازه گیری سوال ها یا ماده های آزمون با توجه به ضرایب دشواری و تمیز ارزیابی می شوند. مقدار ضریب دشواری از طریق محاسبه درصد پاسخ دهندگان به یک سوال به دست می آید و مقدار ضریب تمیز از راه محاسبه ضریب همبستگی بین سوال و کل آزمون تعیین می شود. پارامتر شیب سوال معادل ضریب تمیز در نظریه کلاسیک اندازه گیری می باشد. ضرایب آستانه بدین معنا که اندازه آن ها معرف سطح صفتی است که شخص باید داشته باشد تا پاسخ او با احتمال ۰/۵ ولی در نظریه سوال-پاسخ (مدل پاسخ مدرج) هر سوال با یک پارامتر شیب سوال و پارامترهای آستانه توصیف می شود. پارامتر شیب سوال معادل ضریب تمیز در نظریه کلاسیک اندازه گیری می باشد. پارامترهای آستانه را با فضای بین طبقات توصیف می کنیم بنابراین اگر آزمونی پنج گزینه ای باشد دارای چهار آستانه خواهد بود. در مدل پاسخ مدرج برای هر آستانه یک منحنی ویژه عملیاتی محاسبه می شود. پارامترهای آستانه تعریف ساده ای دارند، بدین معنا که اندازه آن ها معرف سطح صفتی است که شخص باید داشته باشد تا پاسخ او با احتمال ۰/۵ بالای آستانه آن مقوله قرار بگیرد. براساس نتایج آزمون مدل (IRT)، سوالات ۲۲ و ۳۵ سوالات نامناسب و سوالات ۲۰ و ۴۱ سوالات مناسبی در این آزمون بودند.

در شکل های ۴ و ۵ نمونه هایی از منحنی های طبقه پاسخ (CRC) برای یک سوال ضعیف و یک سوال قوی گزارش شده است. منحنی های طبقه پاسخ نشان دهنده احتمال پاسخ دادن آزمودنی در یک طبقه خاص به شرط سطح معینی از صفت مکنون است. این منحنی ها احتمال پاسخ آزمودنی

نمره سوال شونده برابر با مجموع پاسخ‌های درست او نیست، نمره آزمودنی که با تنا نمایش داده می‌شود هم تابع تعداد پاسخ‌های آزمون شونده و هم تابع ویژگی‌های سوال و آزمون است. مورد دیگر اختلاف بین این دو تئوری نا متغیر بودن ویژگی‌های روانسنجی سوالات نسبت به نمونه است (Hogan 2007).

همچنین با بررسی نمره‌گذاری در دو نظریه مشخص می‌شود که نمره‌گذاری در نظریه سوال- پاسخ بسیار معتبرتر از نمرات نظریه کلاسیک است. زیرا در نظریه سوال- پاسخ هر سوالی که ضریب تمیز بیشتری دارد و می‌تواند سهم بیشتری در سنجش صفت مکنون داشته باشد وزن بیشتری بر نمره توانایی اعمال می‌کند. در حالی که در نظریه کلاسیک در حالت کلی تاثیر سوالات قوی و ضعیف در نمره برآورد شده فرد یکسان است.

از سوی دیگر می‌توان گفت IRT ارزان‌ترین روش برای ارزیابی ویژگی‌های روان‌سنجی برای هر دو مقیاس‌های جدید و قدیمی و همین‌طور برای اصلاح پرسشنامه، مورد استفاده قرار می‌گیرد. بر اساس نظریه (IRT) منحنی‌های ویژگی‌های ماده‌ها نشان داده شد که به سازندگان پرسشنامه کمکی برای طبقه‌بندی پاسخ‌ها می‌کند و همچنین اطلاعات منحنی‌ها شرح داده شد که سازندگان آزمون را قادر می‌کند تا ماده و کارکرد مقیاس را بیش از دامنه سازه زیربنایی ارزیابی کنند تا ابزار آنها وسیله‌ای مطمئن برای بررسی جمعیت‌ها باشد (Ramsay and Houston 1997).

اما کاربرد روش IRT دارای محدودیت‌هایی است: اول اینکه بیشتر محققان برطبق مدل نظریه کلاسیک آزمون (CTT) آموزش دیده‌اند و نیز به دلیل تعبیر آسان این مدل بسیار کاربرد دارد و همین‌طور این مدل به راحتی نرم‌افزاری را برای روشی از قبیل آلفای کرانباخ ایجاد می‌کند که به آسانی در دسترس می‌باشد.

در مقابل، مدل‌های IRT دانش مورد نیاز از اندازه‌گیری تئوری تا فهم پیچیده ریاضی را توسعه می‌دهند تا بر طبق

در هر طبقه را به شرط سطح صفت نشان می‌دهند. پارامترهای سوال شکل و جایگاه منحنی‌های طبقه پاسخ را تعیین می‌کنند. هر چه پارامتر شیب سوال بیشتر باشد منحنی‌های طبقه پاسخ به هم فشرده‌تر و ارتفاع آن‌ها بیشتر می‌شود مانند شکل ۵ که به سوال ۱۹ مقیاس اختصاص دارد و هرچه شیب سوال کمتر باشد نمودارها در هم تنیده‌تر و ارتفاع آن‌ها کمتر می‌شود مانند شکل ۴ که مربوط به سوال ۱۶ مقیاس افسردگی است.

علاوه بر این مطالب می‌توان کارکرد هر طبقه پاسخ را با طبقه دیگر مقایسه نمود. همچنین براساس میزان برازش طبقه سوال می‌توان تعداد طبقه مورد نیاز برای هر سوال تصمیم‌گیری نمود (Reeve and Fayers 2005) در شکل ۴ مشاهده می‌شود. طبقه ۳ بصورت کامل تحت پوشش طبقات ۲ و ۴ قرار گرفته و کارکرد مناسبی از خود نشان نداده است. طبقه پاسخ ۵ این سوال نیز در زیر طبقه ۴ قرار گرفته است که آن نیز نشانه‌ای از ناتوانی این سوال در تفکیک افراد به طبقات با توانایی‌های متفاوت است. به همین دلایل است که سوال ۲۶ جزء سوالات ضعیف قرار گرفته است. از طرف دیگر در سوال ۱۹ (شکل ۵) مشاهده می‌شود هر طبقه شکل خود را حفظ نموده و از طبقه پاسخ‌های دیگر مجزا است و این سوال توانسته است افراد با توانایی‌های مختلف را در طبقات مجزا جای دهد.

در جدول ۵ مشاهده می‌شود که بین رتبه‌های افراد بر اساس نمرات هنجار دو نظریه تفاوت‌های قابل توجه‌ای وجود دارد. به عنوان مثال آزمودنی شماره ۴۰ بر اساس نمرات کلاسیک ۲۷۲ مین نفر است و بر اساس نمرات سوال-پاسخ صد و هفتاد و نهمین (۱۷۹) نفر است که نشان دهنده اختلاف رتبه ۹۲ نفری بین ۷۵۰ نفر است. در گزینش ها و تصمیم‌گیری‌ها چنین تفاوت رتبه‌ای بسیار اهمیت دارد.

در تبیین این تفاوت‌ها می‌توان گفت؛ برخلاف روش کلاسیک که نمره آزمون شونده از مجموع پاسخ‌های درست او به سوال‌های آزمون به دست می‌آید در روش سوال-پاسخ

نتیجه گیری

بر اساس یافته‌های بدست آمده از این پژوهش می‌توان گفت هر دو روش کلاسیک آزمون و نظریه سؤال-پاسخ از ویژگی‌های منحصر به فرد خود برخوردار بوده و می‌توانند مسیر راه پژوهشگران را در مطالعات آتی طرح‌ریزی کنند. مقایسه نتایج نظریه سؤال-پاسخ با کلاسیک نشان داد که این نظریه پارامترها و ملاک‌های بیشتر و معتبرتری برای ویژگی‌های روان‌سنجی سوالات و آزمون ارائه می‌دهد. نمرات توانایی‌ها تنها که در نظریه سؤال-پاسخ برای سنجش صفت مکنون ارائه می‌شود بر اساس پایه‌های اساسی‌تری نسبت به نظریه کلاسیک محاسبه می‌شود. در این پژوهش سعی شد تا بخشی از قابلیت‌های نظریه سؤال-پاسخ مانند شیب منحنی سوال، ضرایب آستانه‌ای، توابع آگاهی سوال و آزمون، نمرات تنا و انواع نمودارهای گرافیکی در سنجش صفت مکنون معرفی شود. با توجه به مزایا و توانی که در نظریه سؤال-پاسخ در سنجش افسردگی نوجوانان مشاهده شد پیشنهاد می‌شود در سنجش ویژگی‌ها و سازه‌های روانشناختی از مدل‌های نظریه سؤال-پاسخ به شکل گسترده‌تری استفاده شود.

تشکر و قدردانی

امید است این پژوهش توانسته باشد گام مفیدی در معرفی قابلیت‌ها و کاربردهای نظریه سؤال-پاسخ برداشته باشد. در نهایت پژوهشگران بدین وسیله از کلیه نوجوانان شرکت‌کننده در این پژوهش به خاطر اعتماد و همکاری صمیمانه سپاسگزاری می‌کنند.

فرضیات مدل مناسب را انتخاب کنند. به علاوه، حمایت از این نرم افزارها برای محققانی خارج از زمینه ارزیابی آموزشی مورد قبول واقع نشده است. علی رغم چالش‌های مفهومی و محاسبه‌ای، امتیازهای بسیاری برای مدل‌های IRT وجود دارد که نباید نادیده گرفته شود. همچنین کاربرد این مدل برای رشته‌های مختلف دانشگاهی، آموزشی، روانشناسی گسترده شده و تحقیق در مورد مدل‌های IRT غالباً نتایج خوبی را نشان داده است. در حقیقت فهم بهتری از مدل‌ها و کاربردهای IRT بوجود آمده که از طریق این مدل می‌توان ابزارهای پایاتر و مناسبتر ایجاد کرد (Stochl et al. 2012).

یافته جانبی این پژوهش مربوط به بررسی روایی پرسشنامه افسردگی نوجوانان بود که نتایج نشان بین پرسشنامه افسردگی محقق‌ساخته با پرسشنامه افسردگی بک و اضطراب رابطه مثبت و معنی‌داری وجود دارد. همچنین افسردگی نوجوانان با میزان عزت نفس آنها رابطه منفی و معنی‌داری داشته است. این نتایج با پژوهش (akbari 2015) (Balootbangan and Rezaei 2015) همسو بوده است. یادآوری این مطلب ضروری است که این پژوهش در ارتباط با افسردگی نوجوانان انجام شده است لذا در تعمیم نتایج به سایر گروه‌های سنی و سازه‌های دیگر دقت لازم مبذول شود زیرا ممکن است عملکرد الگوهای نمره‌گذاری مختلفی که در این پژوهش مورد توجه قرار گرفته در مورد سازه‌های دیگر روانی و توانایی متفاوت عمل کند.

جدول ۱- مقطع تحصیلی به تفکیک جنسیت دانش آموزان: ویژگی های روانسنجی مقیاس افسردگی

مقطع تحصیلی	پسر		دختر		کل	
	فراوانی	درصد	فراوانی	درصد	فراوانی	درصد
هفتم	۶۴	۸/۵۳	۶۴	۸/۵۳	۱۲۸	۱۷/۰۷
هشتم	۶۲	۸/۲۷	۵۷	۷/۶۰	۱۱۹	۱۵/۸۷
نهم	۶۵	۸/۶۷	۶۳	۸/۴۰	۱۲۸	۱۷/۰۷
دهم	۵۹	۸/۸۷	۶۸	۹/۰۷	۱۲۷	۱۶/۹۳
یازدهم	۵۹	۷/۸۷	۷۲	۹/۶۰	۱۳۱	۱۷/۴۷
دوازدهم	۵۵	۷/۳۳	۶۲	۸/۲۷	۱۱۷	۱۵/۶۰
کل	۳۶۴	۴۸/۵۳	۳۸۶	۵۱/۴۷	۷۵۰	۱۰۰

جدول ۲- خصوصیات روانسنجی نمونه سوالات ضعیف و سوالات قوی: ویژگی های روانسنجی مقیاس افسردگی نوجوانان

مدل کلاسیک					مدل سوال-پاسخ			
سوال	توان افتراقی	ضریب دشواری	ضریب تمیز	آستانه ۱	آستانه ۲	آستانه ۳	آستانه ۴	
سوال ۰۶	۰/۷۴	۰/۸۱	۰/۵۱	۰/۴۳	۲/۲۶	۴/۰۱	۶/۷۳	
سوال ۲۸	۰/۷۵	۰/۵۰	۰/۶۴	۱/۶۳	۲/۶۷	۴/۰۸	۵/۸۱	
سوال ۳۱	۰/۷۹	۰/۶۰	۰/۵۸	۱	۲/۸	۴/۱۵	۷/۱۷	
سوال ۳۴	۰/۹۱	۲/۰۶	۰/۵۲	-۴/۳۹	-۱/۴۹	۰/۸۹	۳/۵۴	
سوال ۰۴	۰/۹۲	۱/۴۰	۰/۶۴	-۱/۸۸	۰/۲۲	۲/۳۹	۴/۵۹	
سوال ۲۷	۱/۹۴	۱/۳۱	۱/۴۷	-۰/۷۱	۰/۱۱	۱/۱۱	۲/۱۴	
سوال ۲۰	۲/۰۱	۱/۷۰	۱/۵۱	-۱/۲	-۰/۳۲	۰/۵۴	۱/۶۱	
سوال ۴۱	۲/۰۳	۱/۲۶	۱/۶۳	-۰/۵	۰/۲۵	۰/۸	۱/۷۶	
سوال ۱۸	۲/۰۶	۱/۵۳	۱/۶۸	-۱/۰۸	-۰/۱۴	۰/۸۶	۱/۷۵	
سوال ۱۳	۲/۲۱	۱/۶۳	۱/۵۵	-۱/۰۱	-۰/۱۲	۰/۶۳	۱/۴۵	

جدول ۳- میزان آگاهی دهنده گی سؤالات ضعیف و قوی بر روی پیوستار تتا: ویژگی های روانسنجی مقیاس افسردگی نوجوانان

سوال	سطوح پیوستار تتا						
	۳	۲	۱	۰	-۱	-۲	-۳
سوال ۶	۰/۰۸۳	۰/۰۸۲	۰/۰۷۹	۰/۰۷۱	۰/۰۶	۰/۰۴۶	۰/۰۳۳
سوال ۳۱	۰/۱۰۷	۰/۱۰۴	۰/۰۹۷	۰/۰۸۲	۰/۰۶۳	۰/۰۴۳	۰/۰۲۸
سوال ۲۸	۰/۱۳	۰/۱۲۴	۰/۱۰۷	۰/۰۸۱	۰/۰۵۵	۰/۰۳۳	۰/۰۱۹
سوال ۳۴	۰/۰۷۸	۰/۰۸۲	۰/۰۸۳	۰/۰۸۴	۰/۰۸۳	۰/۰۸۲	۰/۰۸۱
سوال ۲۲	۰/۰۹۸	۰/۱	۰/۱۰۱	۰/۱	۰/۰۹۸	۰/۰۹۳	۰/۰۸۳
سوال ۱۴	۰/۳۴۳	۰/۶۱۴	۰/۶۶	۰/۶۹۵	۰/۶۷۶	۰/۴۷۱	۰/۱۶۹
سوال ۳۵	۰/۳۸	۰/۶۵۶	۰/۷۰۵	۰/۷۱۱	۰/۶۷۶	۰/۵۲۴	۰/۱۹۸
سوال ۱۸	۰/۲۷۵	۰/۷۱۷	۰/۸۶۷	۰/۸۶۵	۰/۸۰۷	۰/۴۱۳	۰/۱۰۴
سوال ۰۳	۰/۱۷۹	۰/۵۳۹	۰/۷۵۷	۰/۷۴۸	۰/۷۴۸	۰/۷۲۸	۰/۵۳۴
سوال ۱۹	۰/۲۰۱	۰/۶۵۹	۰/۸۹۹	۰/۹۱	۰/۸۸	۰/۷۷۹	۰/۳۰۶

جدول ۴- نمره هنجار Z و تتا برای میزان افسردگی نوجوانان: ویژگی های روانسنجی مقیاس افسردگی نوجوانان

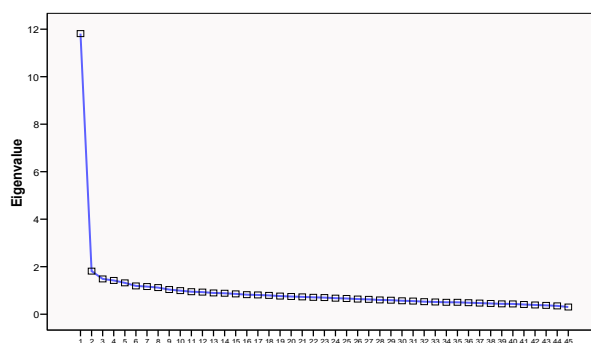
نمرات	دختر			پسر		
	میانگین	انحراف معیار	حجم نمونه	میانگین	انحراف معیار	حجم نمونه
نمرات خام	۱۱۹/۲۲	۳۰/۴۰	۳۸۶	۱۰۷/۸	۲۲/۸۵	۳۶۴
نمرات استاندارد	۰/۱۸	۱/۰۸	۳۸۶	-۰/۲۰	۰/۸۶	۳۶۴
تتا	-۰/۴۰	۰/۸۴	۳۸۶	-۰/۶۹	۰/۶۴	۳۶۴

جدول ۵- مقایسه رتبه افراد بر اساس نمرات هنجار و تتای مقیاس افسردگی نوجوانان

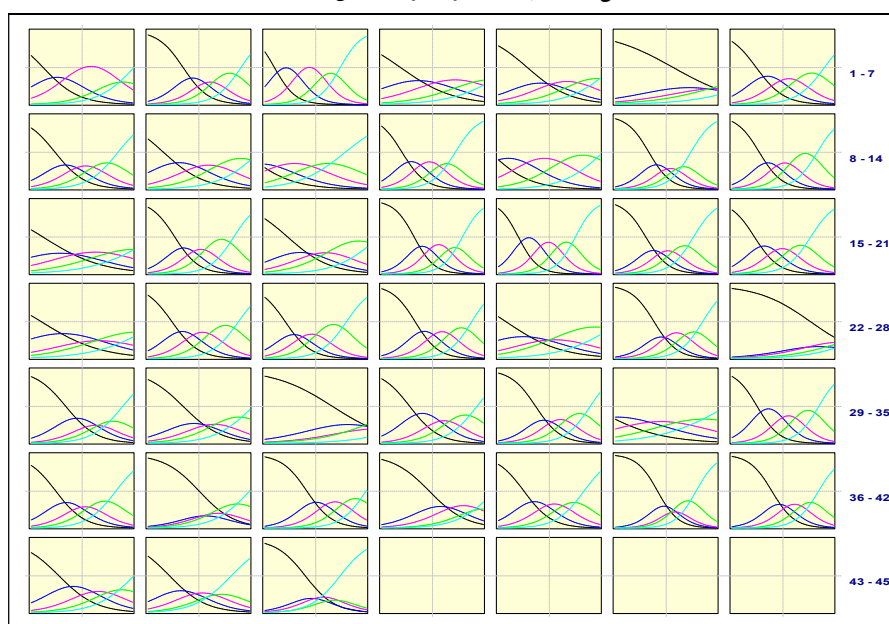
شرکت کننده	رتبه در Z	رتبه در θ	اختلاف رتبه	شرکت کننده	رتبه در Z	رتبه در θ	اختلاف رتبه
شرکت کننده ۱	۱۶۵	۲۳۶	۷۱	شرکت کننده ۳۱	۲۹۹	۲۳۸	۶۱
شرکت کننده ۲	۲۸۶	۳۴۰	۵۴	شرکت کننده ۳۲	۲۵۴	۱۹۲	۶۲
شرکت کننده ۳	۲۴۲	۲۹۳	۵۱	شرکت کننده ۳۳	۲۸۷	۲۲۵	۶۲
شرکت کننده ۴	۴۸	۹۳	۴۵	شرکت کننده ۳۴	۳۵۲	۲۹۰	۶۲
شرکت کننده ۵	۲۵۳	۲۹۸	۴۵	شرکت کننده ۳۵	۱۱۲	۴۴	۶۸
شرکت کننده ۶	۲۳۰	۲۷۴	۴۴	شرکت کننده ۳۶	۱۹۲	۱۲۴	۶۸
شرکت کننده ۷	۲۴۷	۲۹۱	۴۴	شرکت کننده ۳۷	۳۱۱	۲۳۱	۸۰
شرکت کننده ۸	۲۸۵	۳۲۶	۴۱	شرکت کننده ۳۸	۳۵۱	۲۶۵	۸۶
شرکت کننده ۹	۶۲	۱۰۲	۴۰	شرکت کننده ۳۹	۳۱۰	۲۱۸	۹۲
شرکت کننده ۱۰	۲۷۱	۳۱۰	۳۹	شرکت کننده ۴۰	۲۷۲	۱۷۹	۹۳

جدول ۶- همبستگی مقیاس افسردگی نوجوانان با افسردگی، اضطراب و عزت نفس

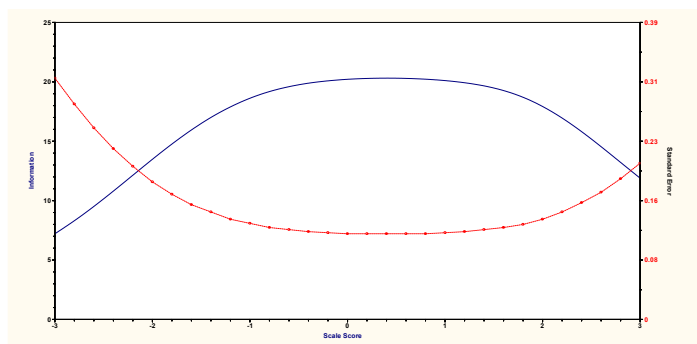
متغیرها		افسردگی بک		عزت نفس		اضطراب	
		p	r	p	r	p	r
افسردگی نوجوانان		۰/۰۰۰۵	۰/۷۳	۰/۰۰۰۵	-۰/۷۱	۰/۰۰۰۶	۰/۳۸



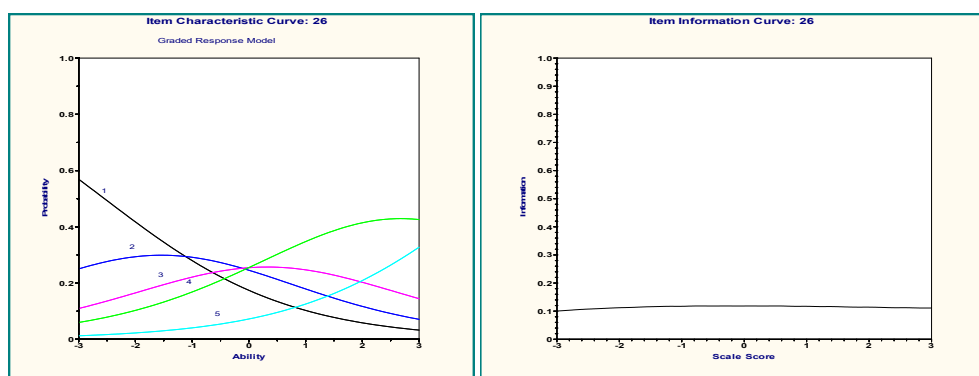
شکل ۱- نمودار اسکری برای تعیین بعدیت



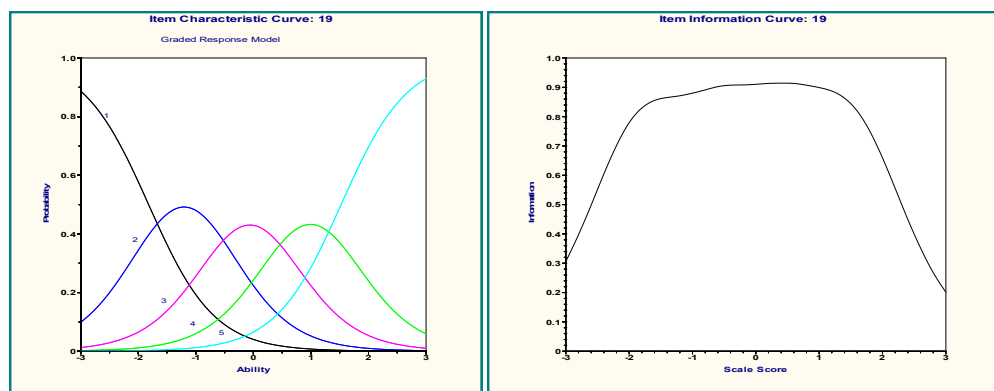
شکل ۲- منحنی پاسخ-طبقه ۴۵ آیتم آزمون افسردگی نوجوانان



شکل ۳- تابع آگاهی آزمون افسردگی نوجوانان



شکل ۴- منحنی های طبقه پاسخ و منحنی تابع آگاهی یکی از سوالات ضعیف مقیاس افسردگی نوجوانان (سوال ۲۶)



شکل ۵- منحنی های طبقه پاسخ و منحنی تابع آگاهی یکی از سوالات خوب مقیاس افسردگی نوجوانان (سوال ۱۹)

Reference

- Abramowitz, J.S. and Mahaffey, B., 2012. *The Oxford Handbook of Clinical Psychology*. In Oxford University Press.
- Akbari Balootbangan, A. and Rezaei, A.M., 2015. Construction and Validation to Self-Discovery Scale (SDS) Based on Different Theories of Personality in Students,

Journal of Clinical Psychology. 3(23), pp. 55-68.

- Akbari, A., 2013. *Youth and young Problems*. Theran: Savalan Publisher.
- Birmaher, B., Brent, D. and AACAP Work Group on Quality Issues., 2007. Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. *Journal of the American Academy*

- of *Child and Adolescent Psychiatry*, 46(11), pp. 1503-1526.
- Ashton, M.C. and Lee, K., 2007. *Empirical, theoretical, and practical advantages of the HEXACO model of personality structure*. *Personality and Social Psychology Review*, 11, pp. 150-166.
- Birnbaum, A., 1968. *Some latent trait models and their use in inferring an examinee's ability*. *Statistical theories of mental test scores*.
- Embretson, S.E. and Reise, S.P., 2013. *Item response theory*. Psychology Press.
- Everett, B. 2013. *An introduction to latent variable models*. Springer Science and Business Media.
- Gierl, M.J. and Lai, H., 2013. Instructional topics in educational measurement (ITEMS) module: Using automated processes to generate test items. *Educational Measurement: Issues and Practice*, 32(3), pp. 36-50.
- Harvey, R.J., 2003. *Applicability of binary IRT models to job analysis data*. In A. Meade (Chair), *Applications of IRT for measurement in organizations*. Symposium conducted at the Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Hogan, T.P., 2013. *Psychological testing: A practical introduction*. Wiley Global Education.
- Jiang, S., Wang, C. and JWeiss, D., 2016. *Sample Size Requirements for Estimation of Item Parameters in the Multidimensional Graded Response Model*. *Frontiers in Psychology*. 7, pp. 1-10.
- Junker, B.W. and Sijtsma, K., 2001. Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25(3), pp. 258-272.
- Lambert, K.G., 2006. Rising rates of depression in today's society: Consideration of the roles of effort-based rewards and enhanced resilience in day-to-day functioning. *Neuroscience and Biobehavioral Reviews*, 30(4), pp. 497-510.
- McDonald, R.P., 2013. *Test theory: A unified treatment*. Psychology Press.
- Morales, R.A., 2009. Evaluation of mathematics achievement test: A comparison between CTT and IRT. *The International Journal of Educational and Psychological Assessment*, 1(1), pp. 19-26.
- Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V. and Ustun, B., 2007. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *The Lancet*, 370(9590), pp. 851-858.
- Paus, T., Keshavan, M. and Giedd, J.N., 2008. Why do many psychiatric disorders emerge during adolescence? *Nature Reviews Neuroscience*, 9(12), pp. 947-957.
- Ramsay, S. L. and Houston, D.C., 1998. The effect of dietary amino acid composition on egg production in blue tits. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1404), pp. 1401-1405.
- Reynolds, W.M., 2010. *Reynolds adolescent depression scale*. John Wiley and Sons, Inc.
- Sadock, B. J. and Sadock, V.A., (Eds.) 2010. *Kaplan and Sadock's pocket handbook of clinical psychiatry*. Lippincott Williams and Wilkins.
- Reckase, M., 2006. *Multidimensional item response theory*. *Handbook of statistics*, 26.
- Reeve, B. B. and Fayers, P., 2005. *Applying item response theory modeling for evaluating questionnaire item and scale properties*. In P. Fayers and R.D. Hays (Eds.), *Assessing Quality of Life in Clinical*

- Trials: Methods of Practice. 2nd Edition. Oxford University Press. pp. 55-73.
- Samejima, K., Ueda, Y., Doya, K. and Kimura, M., 2005. *Representation of action-specific reward values in the striatum*. *Science*, 310(5752), pp. 1337-1340.
- Sayf, A.A., 2016. *Measurement, assessment and training evaluation*. Tehran: Dowran Publisher.
- Stochl, J., Jones, P.B. and Croudace, T.J., 2012. Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers. *BMC Medical Research Methodology*, 12(1), pp. 74-83.

Comparison of the Psychometric Characteristics of the Adolescents Depression Scale Based on the Item Response Theory with Those Based on the Classic Test Theory

Keshavarz, GH., Ph.D. Student, Department of assessment and Measurement, School of Psychology and Educational Sciences, Allame tabataba'i University, Tehran, Iran-Corresponding Author: ghaseem_keshavarz@yahoo.com

Akbari Balootbangan, A., Ph.D. Student, Department of Psychology and Educational Sciences, School of Psychology and Educational Sciences, Semnan University, Semnan, Iran

Babakhani, KH., Ph.D. Student, Department of Health Psychology, School of Psychology and Educational Sciences, Karaj Azad University, Karaj, Iran

Received: Mar 12, 2017

Accepted: Jul 2, 2017

ABSTRACT

Background and Aim: The objective of this study was to assess the psychometric properties of the Adolescent Depression Scale (ADS) based on the item-response theory and compare the results with those based on the classic test theory.

Materials and Methods: A total of 750 students (364 males and 386 females) were selected through multistage random clustering (levels proportional to size) and completed a questionnaire comprising 46 questions developed by the investigators.

Results: Results of factor analysis and questions structure showed that the important item-response theory presumptions, namely, unidimensionality and local independence, are contained in ADS. As regards the psychometric characteristics of the questions, results from the two theories were similar, except that the item-response theory provided more valid indices for assessing the questions. Further analysis of the data indicated that the two theories showed close values for internal consistency with regard to psychometric characteristics, such that values of 0.933 and 0.947 were obtained for Alpha Cronbach and marginal validity, respectively. Finally, the normality scores were calculated based on the Item Response Theory with those based on the Classic Test Theory. Ranking the individuals by the two theories showed considerable differences as regards the depression construct continuum, such as a difference of 86, 92, or 93 individuals.

Conclusions: Based on the findings, it can be concluded that the item-response theory models provides a higher number of more exact indices for judgment and making decisions about questions, tests and individual ranking. Therefore, application of the item-response theory for assessment of other educational and psychological constructs is recommended.

Keywords: Item-Response Theory, Depression, Adolescence, Classic Test Theory