ضرورت تک‌بعدی بودن نشانگرها در اندازه‌گیری‌های علوم رفتاری و پزشکی: کاربرد مدل‌سازی معادلات ساختاری

بلال ایزانلو [۱]، مجتبی حبیبی [۲]، رضا باقریان سرارودی [۳]

مقاله پژوهشی

چکیده

زمینه و هدف: ویژگی تک‌بعدی بودن نشانگرهای یک سازهدر نظریه کلاسیک اندازه‌گیری ریشه داشته و در بافت این نظریه اصطلاح متجانس بودن سؤال‌هایی که یک سازه را اندازه‌گیری می‌کنند به طور غیرمستقیم به این موضوع اشاره دارد. برداشت‌های نادرست از این اصطلاح باعث شد، بسیاری از پژوهشگران برای بررسی و تعیین این ویژگی، از روش‌هایی مثل آلفای کرونباخ یا تحلیل عاملی اکتشافی استفاده کنند. این در حالی است که هیچ‌یک از آن‌ها برای این منظور طراحی نشده‌اند.

مواد و روش‌ها: در قالب یک طرح زمینه‌یابی کاربردی، از بین جامعه مادران کودکان عادی به روش نمونه‌گیری خوشه‌ای چند مرحله‌ای ۴۶۸ مادر پرسش‌نامه PSI-SF را تکمیل نمودند. داده‌ها با روش مدل معادلات ساختاری و نرم افزار LISREL تحلیل شدند.

یافته‌ها: با استفاده از رویکرد مدل‌سازی معادلات ساختاری و نرم‌افزارهای مربوطه که امکان بررسی توزیع مانده‌ها را فراهم می‌کنند، می‌توان وجود یا عدم وجود ویژگی تک بعدی بودن را در نشانگرهای مربوط به یک سازه بررسی کرد.

نتیجه‌گیری: پژوهش حاضر علاوه بر مباحث نظری با یک نمونه داده تجربی این موضوع را از نقطه نظر مدل‌سازی معادلات ساختاری مورد بررسی قرار است.

واژه‌های کلیدی: معادلات ساختاری، تک‌بعدی بودن، فضای مکنون

ارجاع: ایزانلو بلال،حبیبی مجتبی،باقریان سرارودی رضا.**ضرورت تک‌بعدی بودن نشانگرها در اندازه‌گیری‌های علوم رفتاری و پزشکی: کاربرد مدل‌سازی معادلات ساختاری**. مجله تحقیقات علوم رفتاری ۱۳۹۲؛ ۱۱(۶): ۶۸۴-۶۷۱

دریافت مقاله:۱۳۹۲/۱۲/۱۲ پذیرش مقاله: ۱۳۹۲/۱۲/۱۳

۱- دانشجوی دکتری سنجش آموزش، دانشکده علوم تربیتی و روان‌شناسی دانشگاه تهران، تهران، ایران

۲- استادیار، روان‌شناسی سلامت، پژوهشکده خانواده، دانشگاه شهید بهشتی، تهران، ایران

۳- دانشیار، مرکز تحقیقات علوم رفتاری، گروه روان‌پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران(نویسنده مسؤول) Email:bagherian@med.mui.ac.ir

[Downloaded from rbs.mui.ac.ir on 2024-04-19]

[DOR: 20.1001.1.17352029.1392.11.6.15.6]

مقدمه

شاید بهترین شیوه برای ورود به بحث تک‌بعدی بودن (Unidimensionality) نشانگرها اشاره‌ای کوتاه به رویکردهای ساخت مقیاس باشد. در ساخت مقیاس‌ها سه رویکرد تقریباً متفاوت وجود دارد. اولین رویکرد که نظری- منطقی (Rational-theoretical approaches) نام دارد، ساده‌ترین روش ساخت مقیاس محسوب می‌شود به‌خصوص زمانی که در ضعیف‌ترین شکل، که در آن ملاحظات رسمی ویژگی‌های روان‌سنجی مقیاس در نظر گرفته نمی‌شود.با استفاده از این رویکرد، سازنده مقیاس به سادگی سؤال‌هایی را که با درک نظری وی از سازه مورد نظر هماهنگ هستند، می‌نویسد (یعنی سؤال‌هایی که روایی صوری خوبی دارند) (۱). در دومین رویکرد که در اصل یک رویکرد تجربی است سؤال‌ها صرفا بر اساس سهم آن‌ها در روایی تجربی کلی با توجه به ملاک‌های خاصی انتخاب می‌شوند. در این رویکرد ملاحظات مربوط به روایی سؤال و به میزان کمتری همبستگی‌های درون سوالی (inter-item) بسیار مهم هستند. بر اساس این رویکرد، اگر قرار باشد آزمونی ساخته شود سؤال‌ها طوری انتخاب می‌شوند که با ملاک مورد نظر همبستگی‌های بالایی داشته و تا جایی که امکان دارد همبستگی درونی آن‌ها کم باشد. رویکرد سوم در ساخت آزمون کمتر به مباحث آماری می‌پردازد و بیشتر جنبه روان‌شناختی دارد. این رویکرد برای نشان دادن خصیصه‌های واحد خاصی که در مطالعه رفتار انسان مهم هستند از نظریه‌های روان‌شناسی استفاده می‌کند؛ بنابراین آزمون‌ها برای اندازه‌گیری این خصیصه‌های واحد ساخته می‌شوند. در این رویکرد، تأکید زیادی بر انتخاب سؤال‌هایی است که چیز مشابهی را اندازه‌گیری می‌کنند و به اصطلاح متجانس (Homogenous) هستند. بر خلاف رویکرد قبلی، همبستگی‌های دو رشته‌ای (biserial correlation) وچهارخانه‌ای (Tetrachoric correlation) ارزش بیشتری دارند.

مقیاس‌سازی به روش سوم که مقیاس‌سازی متجانس نام دارد، به دلایل بسیار زیادی مطلوب‌تر است؛ چرا که تحت این شرایط سؤال‌هایی که رابطه بالایی با سایر سؤال‌ها دارند همه یک خصیصه واحد را اندازه‌گیری می‌کنند. پس همگی در یک مقیاس قرار می‌گیرند. به‌علاوه، چنین خزانه‌ای از سؤال‌های متجانس دارای مزیت‌های ریاضی و آماری نیز می‌باشد. در عین حال، به‌نظر می‌رسد که مزیت‌های این نوع مقیاس‌ها در ساخت نظریه و بررسی‌های نظری توسط اکثر روان‌شناسان علاقه‌مند به این نوع سؤال‌ها پذیرفته شده است (۲-۱).

وجود مزیت‌های ریاضی و آماری در مقیاس‌های تک‌بعدی باعث شده است که مبانی نظری مدل‌های مختلف اندازه‌گیری نیز بر اساس ابزارهایی که خصیصه واحدی را اندازه‌گیری می‌کنند پایه‌گذاری گردد. نظریه‌های کلاسیک اندازه‌گیری، سؤال- پاسخ (۳) و مدل‌سازی معادلات ساختاری هر یک به نوعی مبتنی بر این فرض هستند که مقیاس مورد نظر خصیصه واحدی را اندازه‌گیری می‌کند و از روش‌هایی برای بررسی درستی این موضوع استفاده می‌کنند (۵-۴).

در نظریه معادلات ساختاری متغیرهای آشکار یا نشانگرها تابع خصیصه مکنون قلمداد می‌شوند. نشانگرهای مربوط به هر خصیصه مکنون باید دارای ویژگی‌هایی باشند که عبارتند از: الف) هماهنگی نشانگرها با هم که توسط ضرایب اعتبار به خصوص آلفای کرونباخ ارزیابی می‌شود. ب) همه نشانگرهای مربوط به یک خصیصه همان خصیصه مکنون را منعکس کنند (روایی همگرا) در عین حال ج) بتوانند خصیصه مورد نظر را به لحاظ آماری از سایر خصیصه‌های مکنون متمایز سازند (روایی واگرا) د) واریانس مشترک هر نشانگر با سایر نشانگرها در همان خصیصه به متغیر مکنون نامشخصی مربوط نباشد (تک بعدی بودن) (۶). هر یک از این موارد ویژگی‌های متفاوتی هستند که باید با روش‌های مناسب خاصی آن‌ها را بررسی کرد، نادیده گرفتن هر یک از این موارد در ساخت و ارزیابی مقیاس‌ها می‌تواند به نتایج گمراه‌کننده‌ای منجر گردد. بسیاری از سردرگمی‌های موجود در این خصوص ناشی از عدم تفکیک بین این ویژگی‌ها به خصوص اعتبار ثبات درونی (Internal consistency reliability) و همگنی یا تک بعدی بودن (Homogeneity or unidimensionality) است (۷).

اگر چه تفکیک بین هر یک از موارد به لحاظ نظری روشن است. ولی آن‌چه در عمل بیشتر از همه دیده می‌شود این است

[Downloaded from rbs.mui.ac.ir on 2024-04-19]

[DOR: 20.1001.1.17352029.1392.11.6.15.6]

که پژوهشگران بیشتر و در بهترین حالت برای ابزار مورد نظر فقط اعتبار، روایی همگرا و روایی واگرا را بررسی کرده و اغلب بدون بررسی ویژگی تک‌بعدی بودن نتایج خود را گزارش می‌کنند. گذشته از این پژوهشگرانی هم که این ویژگی را بررسی کرده‌اند اغلب از روش‌های سنتی نادرست نظیر ضریب آلفای کرونباخ و تحلیل عاملی اکتشافی که برای بررسی اعتبار، روایی همگرا و واگرا مورد استفاده قرار می‌گیرند، استفاده کرده‌اند، این در حالی است که هیچ‌یک از این موارد نمی‌توانند وجود یا عدم وجود تک‌بعدی بودن را مشخص سازند. این موضوع توسط پژوهشگران زیادی مورد اشاره و تأکید قرار گرفته است (۱۰-۸).

به لحاظ نظری ریشه‌های ویژگی تک‌بعدی بودن در نظریه توانایی ذهنی اسپیرمن نیز مشهود است. در واقع، فرض عدم وجود کوواریانس بین خطاهای دو اندازه (measures) می‌تواند به نوعی به این فرض اشاره داشته باشد (۱۲-۱۱). بنابراین، وقتی از تک‌بعدی بودن یک مقیاس صحبت می‌شود، در واقع این موضوع مورد نظر است که با حذف اثر عامل زیربنایی که نشانگرهای مربوطه برای بازنمایی آن به کار رفته‌اند، هیچ عامل دیگری نباید کوواریانس باقی مانده بین نشانگرها را تبیین کند (۶-۴).

در ادامه ابتدا مفهوم بعد در نظریه خصیصه مکنون مورد بحث قرار گرفته، سپس ناکارامدی روش‌های سنتی نظیر آلفای کرونباخ و تحلیل عاملی اکتشافی برای تعیین تک‌بعدی بودن مورد اشاره قرار می‌گیرد و در نهایت با استفاده از یک سری داده فرایند بررسی آن بر اساس مدل‌سازی معادلات ساختاری بررسی خواهد شد.

مفهوم بعد در نظریه خصیصه مکنون

یک مجموعه n سؤالی و مجموعه‌ای از k خصیصه مکنون را در نظر بگیرید. فرض کنیم که هر یک از این خصیصه‌ها حداقل عملکرد فرد را در یکی از سؤال‌های این مجموعه تحت تأثیر قرار می‌دهد. این خصیصه‌ها به عنوان ابعاد روان‌شناختی لازم برای توصیف افراد از نظر روان‌شناختی در نظر گرفته می‌شوند. این خصایص، خصیصه‌های مکنون یا متغیرهای مکنون نامیده می‌شوند که آن‌ها را به صورت بردار (θ_K، ... ،θ_2، θ_1) =Θ نشان می‌دهند (Θ یک بردار است). اکنون می‌توان هر فرد را در یک فضای k بعدی که فضای مکنون یا مکنون نامیده می‌شود، با یک نقطه نشان داد. سپس فرض کنید که عملکرد همه افراد در این n سؤال مورد توجه باشد. فرض کنید که هر سؤال برای هر فرد یک بار اجرا شده است (یعنی همه n سؤال یک بار برای همه افراد اجرا شده‌اند). اکنون توزیع فراوانی توأم شرطی (conditional Joint frequency) (در بین افراد) نمره‌های همه سؤال‌ها را برای هر مقدار ثابت θ در نظر بگیرید. اگر این توزیع (غیر قابل مشاهده) برای گروهی از آزمودنی‌ها جامعه که θ آن‌ها مقدار ثابتی است یکسان نباشد پس علاوه بر (θ_K، ...، θ_2،θ_1) یک یا چند بعد روان‌شناختی دیگر نیز وجود دارند که در بین افراد جامعه مورد نظر تفکیک ایجاد می‌کنند. بنابراین باید این ابعاد را نیز به تعریف فضای مکنون کامل (Complete latent space) اضافه کنیم. پس با تعریف فضای کامل، توزیع شرطی نمره سؤال در یک θ ثابت برای همه افراد جامعه مورد نظر یکسان است. تعداد ابعاد فضای مکنون کامل را با k نشان می‌دهیم و در صورتی که تعداد k کم باشد نظریه خصیصه مکنون بسیار مفید است. ویژگی‌ها و ابعاد فضای مکنون کامل طوری تعریف شده که هر دوی آن‌ها به صفات و جوامع خاص مورد توجه در روان‌سنجی وابسته‌اند.

روان‌سنجی تمایل دارد که فضای نهفته کاملش را طوری تعریف کند که همه ابعاد روان‌شناختی مهم که عملکرد افراد را در یک مجموعه معین از سؤال‌ها تحت تأثیر قرار می‌دهند را شامل شده و متغیرهایی که خطای اندازه‌گیری محسوب می‌شوند را نادیده بگیرد. متأسفانه به لحاظ منطقی غیر ممکن است متغیرهایی که خطای اندازه‌گیری محض هستند را از متغیرهایی که خطا نیستند، به طور عینی تفکیک کرد. مدل‌هایی که برای توضیح عملکرد فرد در آزمون به بیش از یک خصیصه نیاز دارند به مدل‌های چندبعدی معروف هستند. سازندگان آزمون‌ها به دنبال ساخت آزمون‌هایی هستند که یک خصیصه را اندازه‌گیری کنند این کار چندین مزیت دارد که عبارت هستند از: تفسیر توانایی مجموعه نمرات یک آزمون بهتر و راحت‌تر خواهد شد. به علاوه کاربران آزمون نیز به طور معمول، صرف نظر از تفسیر توانایی‌شان، خواهان یک نمره

[Downloaded from rbs.mui.ac.ir on 2024-04-19]
[DOR: 20.1001.1.17352029.1392.11.6.15.6]

واحد هستند. البته مفروضه تک‌بعدی بودن نمی‌تواند به طور دقیق برقرار باشد چرا که همیشه سایر عوامل شناختی، شخصیتی، عوامل ناشی از موقعیت و اجرای آزمون عملکرد فرد در آزمون را تا حدودی تحت تأثیر قرار می‌دهند. این عوامل عبارتند از انگیزه آزمون، اضطراب آزمون، سرعت عملکرد، تجربه آزمون و سایر مهارت‌های شناختی (۲).

نظریه دو عاملی اسپیرمن نیز در واقع یک توصیف دقیق ریاضی از تک بودن ارایه می‌دهد که در ادامه توصیف شده است. پیش فرض‌های مدل کلاسیک عامل مشترک (Common factor) اسپیرمن عبارت هستند از: (۱) رگرسیون نمرات آزمون بر روی عامل‌ها خطی است (یعنی مدل برای همه افراد از پایین‌ترین نمره تا بالاترین نمره کارایی دارد و تلویحاً یعنی به ازای هر مقدار در متغیر مستقل پراکندگی در متغیر وابسته یکسان است– منطقی بودن یا نبودن این پیش فرض جای تأمل دارد)– (۲) به ازای هر مقدار از نمره عامل‌ها، واریانس باقی مانده نمرات خام یکسان است– با شرط بستن روی نمرات هر عامل، نمرات متغیرها به لحاظ احتمالاتی از هم مستقل خواهند بود. در واقع در این‌جا از اصل استقلال شرطی در شکل قوی آن برای تعیین تعداد ابعاد آزمون‌ها استفاده می‌شود. به‌طوری که وقتی اثر نمرات عامل‌ها حذف شود، نه تنها همبستگی خالص متغیرها باید صفر باشد بلکه گشتاورهای توأم مرتبه بالاتر هر یک از دو متغیر نیز برابر با حاصل ضرب گشتاورهای تک متغیری آن‌ها باشد (این همان تعریف دقیق ریاضی استقلال موضعی و به تبع از آن بعدیت است که برآورد شدن آن در عمل بسیار دشوار است). این که چرا پس از حذف اثر عامل‌ها برای بررسی استقلال آماری فقط گشتاور توأم مرتبه دوم (صفر بودن کوواریانس باقی‌مانده‌ها) متغیرها را بررسی می‌کنیم و گشتاورهای توأم مرتبه بالاتر، مثل کشیدگی یا کجی همزمان (joint Coskewness or joint Cokurtosis) را نادیده می‌گیریم، فقط از روی سهل‌انگاری و راحتی است. بنابراین، می‌توان گفت ما بعدیت آزمون‌ها را عملاً فقط در شکل ضعیف آن بررسی می‌کنیم ولی تمایل داریم تفسیرهای خود از نمرات عاملی و ابعاد آزمون را بر اساس شکل قوی این اصل ارایه دهیم. تنها در حالتی که توزیع متغیرها نرمال چند متغیره (Multivariate normality) باشد صفر بودن کوواریانس متغیرها دلیل بر استقلال آماری متغیرها خواهد بود و نیازی به بررسی گشتاورهای توأم مرتبه سوم و چهارم نخواهد بود در غیر این صورت برای بررسی بعدیت در شکل قوی آن باید تمام گشتاورهای توأم مرتبه بالاتر را نیز بررسی کرد. مدل اسپیرمن در شکل اولیه خود برای متغیرهای پیوسته ارایه شد (۱۳). اما آیا می‌توان این مدل را برای متغیرهای دو ارزشی نیز استفاده کرد؟ وجود متغیرهای دو ارزشی در حوزه اندازه‌گیری باعث شد پژوهشگران برای توجیه آن‌چه که سؤال‌ها اندازه‌گیری می‌کنند همانند مدل کلاسیک به دنبال مبانی ریاضی باشند. در پیشینه نظری و پژوهشی این حوزه برای مفهوم‌سازی بعدیت و تک‌بعدی بودن n متغیر دو ارزشی سه مدل ارایه شده است: ۱) اولین مدل از کارهای افرادی مثل والکر در سال ۱۹۳۱ شروع شد و توسط افرادی مثل فرگوسن، لوینگر، گاتمن و لومسدن ادامه یافت. در این رویکرد، یک مجموعه سؤال دو ارزشی در صورتی تک‌بعدی خواهند بود که اگر سؤال‌ها به ترتیب دشواری مرتب شده باشند، پاسخ درست به یک سؤال به این معنی است که فرد به تمام سؤال‌ها پیش از آن نیز پاسخ درست داده است. در این رویکرد برای توصیف تک بعدی بودن از اصطلاحاتی نظیر، الگوی پاسخ ایده‌آل (ideal response pattern)، مقیاس کامل (perfect scale) و همجنس بودن (homogenous) استفاده شده است. ۲) دومین مدل بر اساس نظریه خصیصه مکنون شکل گرفته است. این نظریه که معادل تحلیل عامل مشترک برای متغیرهای دو ارزشی است بیان می‌کند که به ازای مقادیر خصیصه مکنون پاسخ افراد به سؤال‌ها دو ارزشی متقابلاً از نظر آماری از هم مستقل خواهد بود. تحت این شرایط اگر یک خصیصه مکنون برای تبیین توزیع الگوی پاسخ افراد به n سؤال دو ارزشی کافی باشد آن مجموعه سؤال تک‌بعدی خواهند بود. ۳) سومین مدل برای تعیین بعدیت و بررسی تک‌بعدی بودن بیشتر برای توجیه عباراتی مثل همجنس بودن (homogeneity) و هماهنگی درونی (internal consistency) شکل گرفته است؛ بنابراین، کمتر به‌عنوان یک الگو محسوب می‌شود. این رویکرد واژه‌های همجنس بودن یا هماهنگی درونی را معادل‌های واژه تک‌بعدی

[Downloaded from rbs.mui.ac.ir on 2024-04-19]

[DOR: 20.1001.1.17352029.1392.11.6.15.6]

بودن می‌داند و برای اندازه‌گیری میزان آن‌ها از شاخص‌های آماری خاصی مثل آلفا استفاده می‌کند. چون تک‌بعدی بودن یک مجموعه سؤال کمیتی است که می‌توان بود یا نبود آن را به شکلی قطع مشخص کرد، ولی کمیت‌هایی مثل هماهنگی درونی یا همجنس بودن یک مجموعه سؤال طوری تفسیر می‌شوند که بر اساس آن می‌توان گفت یک آزمون نسبت به دیگری هماهنگی درونی کمتر یا بیشتری دارد، پس در استفاده از این عبارات به جای تک‌بعدی بودن باید احتیاط کرد. زیرا اگر از این دو اصطلاح به عنوان معادل‌های واژه تک‌بعدی بودن استفاده کنیم دچار تناقضی می‌شویم که در فوق ذکر گردید (اگر دو اصطلاح ذکر شده به معنای تک‌بعدی بودن نباشند، در این صورت ویژگی یا ویژگی‌هایی که برای توصیف آن‌ها از این دو اصطلاح استفاده می‌شود کدام ویژگی یا ویژگی‌ها هستند؟ البته این ویژگی اعتبار یک مجموعه سؤال نیست چرا که اعتبار در بافت نظریه کلاسیک برابر است با نسبت واریانس واقعی به واریانس مشاهده شده، در حالی که هیچ‌یک از این دو اصلاح و شاخص آلفا که برای بررسی میزان آن‌ها استفاده می‌شود چنین تعریفی ندارند) (۱۳).

تا این‌جا مشخص شد که ویژگی استقلال موضعی یا شرطی در تعریف ابعاد یک سؤال یا مقیاس نقش مهمی دارد (۱۴). این ویژگی که در واقع زیر بنای تعریف بعد در نظریه خصیصه مکنون است، دارای تعاریف مختلفی است. این اصل در شکل اولیه خود (فرم قوی استقلال موضعی) به صورت زیر تعریف شده است.

استقلال موضعی به این معنی است که در داخل هر گروه از آزمون‌دهندگان که همگی بر اساس مقادیر یکسان $\theta_1,\theta_2,...\theta_k$ توصیف شده‌اند، توزیع شرطی نمرات سؤال‌ها همگی از یکدیگر مستقل هستند. این تعریف به این معنی نیست که نمرات سؤال‌ها در کل گروه آزمودنی‌ها با یکدیگر رابطه ندارند بلکه به این معنی است که نمرات سؤال‌ها فقط از طریق متغیرهای نهفته $\theta_1, \theta_2, ... \theta_k$ با یکدیگر رابطه دارند (۲).

با گذشت زمان در بررسی تعداد ابعاد یک آزمون بر اساس استقلال موضعی قوی مشکلاتی پیش آمد. این موضوع باعث گردید که تعاریف جدیدی از اصل استقلال موضعی ارایه شود. چرا که بر اساس تعریف استقلال موضعی قوی باید پس از استخراج یک عامل، که همان میزان خصیصه هر یک از افراد است و فرض می‌شود که تنها عامل مؤثر در پاسخ‌گویی به سؤال‌ها می‌باشد، کوواریانس بین سؤال‌ها در ماتریس باقی مانده‌ها صفر یا نزدیک به صفر باشد، در حالی که شواهد عملی به دست آمده این موضوع را تایید نمی‌کنند (۱۶-۱۵). اولین تلاش رسمی در این خصوص، تعریف اصل استقلال موضعی در شکل ضعیف آن بود که به وسیله مک‌دونالد صورت گرفت. وی تعریفی از استقلال موضعی ارایه داد که به تعریف ضعیف از استقلال موضعی (Weak Local Independency) مشهور است. این تعریف تنها نیازمند آن است که کوواریانس شرطی (غیر خطی یا خطی) هر جفت از سؤال‌ها موجود در یک مقیاس در همه مقادیر θ برابر صفر باشد، یعنی $P=(X_i=x_i,X_j=x_j|\theta)=P(X_i=x_i|\theta).P(X_j=x_j|\theta)$. مک دونالد (۱۷، ۳) بر اساس این تعریف، تعریفی از بعد ارایه داد که این مفهوم را به طور مستقیم به مفهوم استقلال موضعی سؤال ربط می‌دهد. وی پیشنهاد کرد که تعداد ابعاد یک آزمون بر اساس تعداد صفاتی که باید به حساب آیند تا استقلال موضعی ضعیف بین سؤال‌های موجود در آزمون به دست آید، تعریف شود. این همان تعریف پذیرفته شده از استقلال موضعی در مدل‌سازی معادلات ساختاری است و اصل تک‌بعدی بودن در مدل‌سازی معادلات ساختاری در واقع ناشی از عملیاتی شدن این اصل است. توجه کنیم که در شکل قوی استقلال موضعی علاوه بر این که پس از استخراج عامل اول کوواریانس بین متغیرها باید برابر با حاصل ضرب توزیع هر یک از متغیرها باشد، کجی و کشیدگی توأم (Coskewness and Cokurtosis) متغیرها نیز باید برابر با حاصل ضرب توزیع هر یک از متغیرها باشد.

در مدل‌سازی معادلات ساختاری هر یک از سؤال‌ها موجود در مقیاس با معادله $X_i=\lambda_i\xi+\delta_i$ تعریف شده‌اند (۱۸). با توجه به این معادله، هر یک از سؤال‌ها در صورتی تک‌بعدی خواهند بود که:

الف) سوال مورد نظر (X_i)با سایر سؤال‌ها مربوط به سازه مورد نظر (X_j) سازگاری (consistency) داشته باشد. یعنی

[Downloaded from rbs.mui.ac.ir on 2024-04-19]

[DOR: 20.1001.1.17352029.1392.11.6.15.6]

همبستگی بین هر زوج از سؤال‌ها (ρ_{ij})مربوط به سازه (ξ) برابر با حاصل ضرب همبستگی هر یک از آن‌ها با سازه($\rho_{i\xi}$, $\rho_{j\xi}$)مورد نظر باشد ($\rho_{ij}=\rho_{i\xi}*\rho_{j\xi}$).

ب)سؤال مورد نظر باید دارای سازگاری بیرونی (external consistency) باشد. این ویژگی زمانی به‌دست می‌آید که همبستگی سؤال (X_i) مربوط به سازه ξ ($\rho_{i\xi}$)، با سؤال (X'_j) مربوط به سازه ξ' ($\rho_{j'\xi'}$)برابر است با حاصل ضرب همبستگی هر یک از سؤال‌ها با سازه خودشان ضرب در همبستگی بین دو سازه ($\rho_{ij'}=\rho_{i\xi}*\rho_{j'\xi'}*\rho_{\xi\xi'}$).

بر اساس هر دو قاعده واریانس غیر مشترک سؤال‌ها نباید سهمی در همبستگی بین سؤال‌ها داشته باشد. با در نظر گرفتن این مطلب به ناکارامدی روش‌هایی مثل آلفا یا تحلیل عاملی که برخی پژوهشگران برای بررسی تک‌بعدی بودن مورد استفاده قرار می‌دهند، پی می‌بریم. زیرا این روش‌ها واریانس باقی مانده را نادیده می‌گیرند پس نمی‌توانند وجود یا عدم وجود تک‌بعدی بودن در سؤال‌ها را تعیین کنند.

شاخص اعتبار برای هر آزمون برابر با یک منهای نسبت مجموع واریانس هر سؤال بر واریانس کل آزمون ضرب در تعداد سؤال‌های آزمون است، در واقع نسبت واریانس واقعی بر واریانس کل مشاهده شده است (۲۰-۱۹). باگسترش رویکرد معادلات ساختاری شاخص دیگری تحت عنوان اعتبار عاملی مرکب (composite factor reliability) به عنوان شاخص اعتبار مورداستفاده قرارگرفت. برعکس آلفای کرونباخ که فرض می‌کند همه سؤال‌های مربوط به سازه بار یکسانی روی سازه مربوط به خود دارند (آلفا در شکل اولیه خود براین اساس استخراج شد)، اعتبار مرکب، برای هر سؤال مربوط به سازه وزن خاصی که برآورد بار آن بر روی سازه است را در نظر می‌گیرد و به همین دلیل اکثر مواقع مقدار آن از آلفا بیشتر است. روشن است که اگر نشانگرهای مربوط به یک سازه با همسازگاری درونی داشته باشند فردی که در یکی از آن‌ها نمره بالایی بگیرد در سایر نشانگرها نیز نمره بالایی خواهدداشت. در صورتی که این شرایط برای افرا زیادی برقرار باشد پس در آن صورت مقیاس مورد نظر دارای ثبات درونی خوبی است (۲۱).

چون اعتبار بر اساس این پیش فرض که سؤال‌ها از قبل تک بعدی هستند، مجموع کوواریانس بین سؤال‌ها را بر واریانس مشاهده شده تقسیم می‌کند، پس واریانس غیر مشترک هر سؤال را که می‌تواند واریانس خطا را نیز در بر داشته باشد، نادیده می‌گیرد و فقط به واریانس مشترک توجه دارد. بنابراین، چون قادر نیست وجود یا عدم وجود همبستگی بین واریانس باقی‌مانده را بررسی کند، پس نمی‌تواند تک‌بعدی بودن را به طور کامل و دقیق تعیین کند. تحت این شرایط این امکان وجود دارد که برخی از سؤال‌های مربوط به سازه مورد نظر با سازه دیگری نیز ارتباط داشته باشند، چیزی که در واریانس باقی مانده آن‌ها منعکس می‌شود (۲۳-۲۱).

در ارتباط با موضوع واریانس غیر مشترک، لازم به ذکر است که این مسأله جای تأمل دارد که باقی مانده‌های مدل در کل خطا محسوب شود یا به عنوان عوامل اختصاصی (unique) فرض شود. مدل عامل مشترک در شکل اولیه خود فرض می‌کند که عوامل مشترک یا خصیصه مکنون مقادیر ثابتی هستند که باید برآورد شوند (یعنی، آن‌ها را نمره واقعی در نظر می‌گرفت که در طی زمان ثابت هستند و تغییر نمی‌کنند) و مقادیر باقی‌مانده‌ها را در کل به عنوان خطا در نظر می‌گرفت، که در طی زمان نوسان دارند. این نوع نگرش نسبت به تجزیه نمرات مربوط به متغیرهای پیوسته که اولین بار توسط اسپیرمن مطرح شد به سرعت توسط خود وی تغییر یافت. به این صورت که به جای نمره واقعی از عامل مشترک صحبت به میان آمد و به جای واریانس خطا از واریانس عامل اختصاصی. عامل مشترک و عامل اختصاصی بر خلاف نمره واقعی و خطا ثابت نبوده، بلکه به عنوان یک متغیر تصادفی دارای توزیع خاصی هستند. از ابتدا تفسیر عامل مشترک و عامل اختصاصی به این صورت بود که عامل مشترک را یک ویژگی آزمودنی‌ها می‌دانستند که به طور مشترک توسط تمام n متغیر پیوسته اندازه‌گیری می‌شود و عامل‌های اختصاصی که تعداد آن‌ها برابر با تعداد متغیرها است ویژگی‌هایی از افراد هستند که هر یک از متغیرها به تنهایی اندازه‌گیری می‌کنند. در این تفسیر از عامل اختصاصی باقی مانده‌ها از دو مؤلفه تشکیل شده است، بخشی از آن که ویژگی خاص مورد اندازه‌گیری به وسیله سؤال یا

[Downloaded from rbs.mui.ac.ir on 2024-04-19]
[DOR: 20.1001.1.17352029.1392.11.6.15.6]

آزمون خاصی را منعکس می‌کند و ثابت است و بخشی که به خطای اندازه‌گیری مربوط است و در طی زمان تغییر می‌کند. معمولاً ویژگی خاص به عنوان بخش از واریانس معتبر در نظر گرفته می‌شود که خطا مکمل آن بوده و در طی زمان تغییر می‌کند. پس هر دو تشکیل واریانس اختصاصی را می‌دهند؛ بنابراین، با یک بار اجرا، واریانس خطا مشخص نمی‌شود. در واقع، با یک بار اجرا امکان تفکیک واریانس خاص و خطا وجود ندارد. در نتیجه واریانس خطای یک آزمون، مکمل اعتبار آن به شیوه بازآزمایی است. به این ترتیب، همیشه اعتبار به شیوه بازآزمایی از مجذور همبستگی چندگانه آزمون با عامل مشترک بزرگ‌تر است، زیرا در اعتبار به شیوه بازآزمایی واریانس خاص آزمون در هر دو بار به واریانس مشترک اضافه شده و بخش مکمل، واریانس خطا محسوب می‌شود که در طی زمان نوسان دارد. در این رویکرد، نمره واقعی یک فرد از طریق پایداری آن در طی زمان مشخص می‌شود و نمره خطا با ناپایداری آن در طی زمان. به عبارتی، دست‌یابی به نمره واقعی و نمره خطا نیازمند اجرای مجدد آزمون است. یعنی خطا یک متغیر تصادفی تعریف شده در طول تکرارها در طی زمان تعریف می‌شود. این در حالی است که نمره عامل اختصاصی به وسیله اصل استقلال موضعی یا شرطی روی آزمودنی‌هایی تعریف می‌شود که نمره عامل مشترک آن‌ها یکسان است و ربطی به تکرار ندارد. با توجه به این مطالب می‌توان گفت: در مدل عامل مشترک هر یک از آزمودنی‌ها دارای یک نمره عامل مشترک است، به طوری که در هر زیر جامعه‌ای که نمره عاملی یکسانی دارند در هر اجرا (یک اجرای معین) نمرات افراد در متغیرها به لحاظ آماری مستقل از هم خواهند بود. به علاوه در هر یک از اجراها، هر آزمودنی دارای n نمره اختصاصی است (برابر با تعداد متغیرها) که در طول اجراهای مختلف ثابت هستند. در حالی که n نمره خطای آزمودنی‌ها در طول اجراهای مختلف متفاوت هستند (یعنی خود خطاهای مربوط به n آزمون در طول تکرارها به صورت تصادفی تغییر می‌کنند). متأسفانه در حالت نظری تفکیک بین نمره واقعی و عامل‌های مشترک روشن نیست. مبانی نظری نمره واقعی و نمره خطا با فرض این ایده شکل گرفتند که می‌توان برای هر آزمون بی نهایت فرم معادل در نظر گرفت. به این ترتیب می‌توان تعداد زیادی نظریه آزمون مطرح کرد که فقط بر اساس وجود فرم‌های معادل شکل می‌گیرند. یعنی می‌توان برای یک آزمون بی‌نهایت فرم معادل در نظر گرفتکه در هر اجرا از یکی از آن‌ها استفاده می‌شود. این فرم‌ها در هر اجرا دارای یک نمره واقعی به اضافه نمره خطای تصادفی مستقل هستند (که از یک اجرا به اجرای بعدی تغییر می‌کند). تحت این شرایط مدل‌های زیر شکل می‌گیرند (هرچند که امکان ارایه مدل‌های دیگری نیز وجود دارد، همان طور که مک دونالد (۱۳) مدل شماره ۴ زیر را ارایه کرده است): (۱) فرم‌های معادل (parallel forms)، در این حالت هم نمره واقعی فرم‌ها برابر با یکدیگر است و هم واریانس‌های بخش تصادفی خطا یکسان هستند؛ (۲) تاو معادل (tau equivalent) ، در این حالت فقط نمره واقعی فرم‌ها یکسان فرض می‌شود در حالی که واریانس خطای آن‌ها با یکدیگر برابر نیست؛ (۳) اساسا تاو معادل (essentially tau equivalent) ، در این حالت نمره واقعی فرم‌ها نیز در حد یک مقدار ثابت با هم متفاوت است و واریانس خطای فرم‌ها نیز همانند قبل نابرابر فرض می‌شوند؛ (۴) مدل اساسا معادل (essentially equivalent) ، یعنی فقط نمره واقعی فرم‌ها در حد یک مقدار ثابت متفاوت است ولی واریانس خطای آن‌ها یکسان است؛ (۵) مدل هم جنس (congeneric) ، در این مدل که آزادترین مدل اندازه‌گیری است، رابطه بین نمره واقعی سؤال‌ها به صورت یک معادله خطی در نظر گرفته شده و متغیرها به صورت معادله $X_{ik}=[\alpha_k+\beta_k(T_i)]+E_{ik}$ توصیف می‌شوند (i نماد سؤال و kنماد فرد است) توجه کنید که در مقاله حاضر ما شکل استاندارد این معادله را که در آن عرض ازمبداء (α) صفر است مورد بررسی قرار دهیم. البته همان‌طور که می‌دانیم کاهش یا افزایش یک مقدار ثابت به مقادیر یک متغیرتأثیری بر واریانس آن متغیر ندارد (۲۴). تمایز بین تجزیه یک نمره به نمره واقعی- خطا و تجزیه آن به عامل مشترک- اختصاصی در این است که نمره واقعی و خطا به اجرای مکرر فرم‌های یک آزمون در طی زمان مربوط هستند، به‌طوری که نمره واقعی کمیتی است که با اجرای فرم‌های معادل، ثبات در طی زمان را نشان می‌دهد و خطا نمایانگر نوسان در طی زمان است (۱۳).

[Downloaded from rbs.mui.ac.ir on 2024-04-19]
[DOR: 20.1001.1.17352029.1392.11.6.15.6]

استفاده از شاخص‌های مربوط به روایی سازه نظیر روایی همگرا و واگرا (convergent validity & discriminant validity) نیز برای تعیین تک‌بعدی بودن کافی نیست.روایی همگرا به این موضوع می‌پردازد که اندازه‌های مربوط به یک سازه که بر اساس چندین روش اندازه‌گیری به دست آمده‌اند تا چه حد با یکدیگر همبسته هستند و روایی واگرا به دنبال آن است که میزان منحصر به فرد بودن سؤال‌ها در تعریف سازه را تعیین کند. بر اساس این دو ویژگی نشانگرهای یک سازه در حالی که با هم همبستگی بالایی دارند (روایی هم‌گرا) باید با نشانگرهای مربوط به سازه‌های دیگر نیز همبستگی پایینی داشته باشند (روایی واگرا). یکی از روش‌های رایج در بررسی این دو ویژگی روایی عاملی است که در عمل از طریق تحلیل عاملی اکتشافی انجام می‌شود. تحت چنین شرایطی سؤالاتی که یک سازه را اندازه‌گیری می‌کنند روی سازه‌ای که آن را اندازه‌گیری می‌کنند بار عاملی بالایی دارند در حالی که روی عامل یا سازه دیگری که به وسیله سؤال‌ها دیگری اندازه‌گیری می‌شود بار عاملی پایینی خواهند داشت. چون در تحلیل عاملی اکتشافی میزان رابطه هر یک از سؤال‌ها با همه عوامل برآورد شودو اختصاص سؤال‌ها به عوامل بر مبانی ملاک آماری است پس امکان این که بر اساس نتایج بتوان میزان روایی همگرا و واگرا سؤال‌های مربوط به یک سازه را تعیین کرد وجود دارد. اگر چه بر اساس نتایج این روش می‌توان برخی از مشکلات موجود در تک بعدی بودن مقیاس نظیر عوامل غیر قابل انتظار یا سؤالاتی که بر روی سازه‌ای غیر از سازه خود بار دارند را تعیین کرد ولی تحلیل عاملی اکتشافی که برای تعیین روایی همگرا و واگرا استفاده می‌شود نمی‌تواند تضمین کنند تک‌بعدی بودن سؤال‌ها و مقیاس باشد. زیرا این روش نیز دارای چندین نقطه ضعف است. از جمله مهم‌ترین آن‌ها این که فقط بخشی از واریانس هر سؤال را که اشتراک نام دارد مورد توجه قرار می‌دهد و سایر واریانس را نادیده می‌گیرد. به علاوه عامل‌های حاصل از تحلیل عاملی اکتشافی مستقیماً با سازه‌هایی که به وسیله مجموعه‌ای از نشانگرها مشخص شده‌اند متناظر نیستند. زیرا این عامل‌ها بر اساس مجموع موزون همه متغیرهای مشاهده شده تعریف گردیده است. در واقع عامل‌های حاصل از تحلیل عاملی اکتشافی یکی از بی‌نهایت راه حل‌های ممکن حاصل هستند. راه حل یگانه در این خصوص زمانی به دست می‌آید که نشانگرها مربوط به هر عامل را از قبل مشخص کنیم. بر این اساس استفاده از روش‌هایی مثل اعتبار و تحلیل عاملی اکتشافی برای بررسی تک بعدی بودن می‌تواند به نتایج گمراه‌کننده‌ای منجر شود (۲۵).

یکی از کارامدترین روش‌های موجود در بررسی ابعاد سؤال‌ها و آزمون‌ها استفاده از مدل‌سازی معادله ساختاری است. با این که روش‌های دیگری مثل تحلیل عاملی غیر خطی که به وسیله مک دونالد ارایه شد و با نرم‌افزار NOHARM انجام می‌شود و روش ناپارامتری استوت که به وسیله نرم افزار DIMTEST اجرا می‌گردد نیز برای این منظور استفاده می‌شوند ولی مدل‌سازی معادله ساختاری در این بین دارای مزایایی است که روش‌های فوق فاقد آن‌ها هستند (۲۶).

رویکرد مدل‌سازی معادله ساختاری به دلیل این‌که میزان رابطه بین واریانس باقی مانده سؤال‌ها را برآورد می‌کند امکان بررسی تک‌بعدی بودن سؤال‌ها را امکان پذیر می‌سازد. توانمندی این رویکرد اجازه می‌دهد روایی واگرا، همگرا و تک بعدی بودن نشانگرها را با هم بررسی کنیم. در این رویکرد با استفاده از شاخص‌هایی مثل باقی مانده‌های استاندارد شده (Standardized Root Mean Residual)، شاخص‌های برازش و شاخص‌های اصلاح امکان بررسی تک‌بعدی بودن سؤال‌ها فراهم است. باقی‌مانده‌های استاندارد شده که مقدار آن‌ها بزرگ‌تر از ۲/۸۵ باشد (معادل با $P < ۰/۰۱$ در توزیع استاندارد) به‌عنوان باقی مانده‌های بزرگ در نظر گرفته می‌شوند. سؤالاتی که مقادیر استاندارد شده آن‌ها بزرگ‌تر از این مقدار باشد در صورت لزوم باید بازنگری یا حذف شوند. شاخص برازش استاندارد شده (Root Mean Residual) که مقادیر کوچک‌تر از ۰/۰۵ آن قابل قبول هستند نیز برای بررسی این موضع کارایی دارند. مقدار کای دو بزرگ نیز دال بر عدم برازش ماتریس حاصل از مدل با ماتریس حاصل از داده‌ها دارد و در حالت مطلوب نباید معنی‌دار باشد (۲۷، ۲۵).

[Downloaded from rbs.mui.ac.ir on 2024-04-19]

[DOR: 20.1001.1.17352029.1392.11.6.15.6]

مواد و روش‌ها

برای بررسی این موضوع به‌صورت تجربی از داده‌های حاصل از اجرای نسخه کوتاه پرسش‌نامه ۳۶ سؤالی استرس فرزندپروری (Parenting stress index)، با مقیاس ۵ درجه‌ای لیکرت از کاملاً موافقم (نمره ۵) تا کاملاً مخالفم (نمره ۱)، که بر روی ۴۶۸ نفر از مادران کودکان مقطع سنی ۷ تا ۱۲ سال که در سال تحصیلی ۸۸-۸۷ در دبستان‌های شهر تهران مشغول به تحصیل بودند اجرا شده استفاده گردید. این پرسش‌نامه دارای سه زیر مقیاس ۱۲ سؤالی است که عبارت هستند از: کودک مشکل‌آفرین، آشفتگی والدین و تعاملات ناکارآمد والد–کودک. در این‌جا فقط از خرده مقیاس اول این پرسش‌نامه ۳۶ سؤالی استفاده شده است.این ابزار در پژوهش قبلی (۲۸) مورد استفاده قرار گرفته و تقریباً به نوعی بر وجود اعتبار و روایی کافی آن تأکید شده است. اما همان‌طور که در ادامه خواهیم دید متأسفانه پژوهشگران به اشکالات موجود در ابزارها اشاره نکرده‌اند. یکی از دلایل انتخاب یکی از خرده‌مقیاس‌های این ابزار این است که توجه افراد را به اهمیت موضوع تک‌بعدی بودن نشانگرها و بررسی آن در مرحله ساخت مقیاس جلب کنیم. زیرا مشابه این پژوهش داخلی با این ابزار، در مجلات معتبر بین‌المللی نیز مقالاتی وجود دارند که صرفا بر اساس بررسی شاخص‌های برازش سرانگشتی در خصوص اعتبار و روایی کافی ابزارها قضاوت کرده و استفاده از آن را توصیه می‌کنند (۲۸). این موضوع توسط برخی پژوهشگران نیز مورد اشاره قرار گرفته است (۲۹، ۸). به این معنی که در مرحله ساخت ابزار یا ارایه مدل حتماً توزیع باقی مانده‌ها را مورد توجه قرار دهیم و صرفا شاخص‌های برازش متوسط ملاک قضاوت قرار ندهیم.

یافته‌ها

نتایج تحلیل اولیه نشان داد برخی از شاخص‌های برازش در سطح مطلوبی نیستند. $P= ۰/۰۰۰۰۱$ ،$\chi^2_{۵۴}= ۴۰۰/۹۳$ ،$=۰/۱۲$ RMSEA، CFI=۰/۹۳، GFI =۰/۹۷، NFI=۰/۹۲ و از همه مرتبط‌تر به بحث حاضر SRMR =۰/۰۸۳ است. اگر چه برخی از این شاخص‌ها برازش نسبی مدل را نشان می‌دهند ولی آن‌چه بیشتر به بحث حاضر مربوط است باقی مانده‌های استاندارد شده و شاخص SRMR است که نشان می‌دهند کوواریانس غیر مشترک برخی از سؤال‌ها از دامنه مطلوب بیشتر است. بنابراین بررسی بیشتر در این خصوص را با استفاده از واریانس باقی‌مانده انجام خواهیم داد.

مقادیر باقی مانده استاندارد شده و شاخص‌های اصلاح نشان می‌دهند که برخی از سؤال‌ها دارای باقی مانده‌های بسیار بزرگی هستند طوری که برآورد واریانس باقی‌مانده غیر مشترک بین آن‌ها افزایش زیادی در بهبود مدل دارد. روشن است که با توجه به مسأله پژوهش حاضر این شاخص‌های باقی مانده هستند که با توجه حساسیت زیادشان در خصوص واریانس باقی مانده میزان تک‌بعدی بودن نشانگرها را بهتر نشان می‌دهند و دلیل تأکید بر استفاده از آن‌ها برای این منظور نیز همین است. نتایج مربوط به باقی مانده‌های استاندارد شده که تا حدودی دارای کجی مثبت است حاکی از کم برآورد شدن اکثر پارامترها است.

بزرگ‌ترین باقی مانده استاندارد شده ۱۱/۲۳ (بین سؤال ۴ و ۵) و کوچک‌ترین باقی مانده‌های استاندارد ۳/۹۷- (بین سؤال ۴ و ۱۱) و ۳/۴۷- (بین سؤال ۵ و ۱۱) است. شاخص‌های اصلاح نیز حاکی از آن هستند که برآورد کردن کوواریانس خطا برای برخی از سؤال‌ها مقدار کای دوی مدل را به میزان قابل توجهی کاهش می‌دهد. مثلاً شاخص اصلاح برای سؤال‌ها ۴ و ۵ نشان می‌دهند که برآورد کردن کوواریانس بین باقی مانده‌های این دو سؤال با سؤال‌ها ۲، ۸، ۹، ۱۰، ۱۱، ۱۲، به ترتیب به میزان قابل توجهی از مقدار کای دو می‌کاهد. مقادیر برخی از بزرگ‌ترین شاخص‌های اصلاح در جدول ۱ ارایه شده است.برای بررسی‌کردن تک‌بعدی بودن نشانگرها باید هر بار یکی از سؤال‌ها مورد نظر را حذف کرد و نتایج را بررسی کرد. حذف کردن بی‌مورد سؤال‌ها با هم کار درستی نیست زیرا گاهی وجود واریانس باقی‌مانده غیر مشترک بین دو سؤال باعث می‌شود بین سایر زوج‌های سؤال نیز واریانس باقی مانده معنی‌دار به وجود آید (۶).

با یک فرایند تکراری پی در پی هر بار یکی از سؤال‌ها حذف شدند و شاخص‌های مورد نظر بررسی شدند ابتدا سؤال ۴ و سپس سؤال ۵ حذف شدند نتایج نشان داد که با حذف سؤال ۴ مقادیر باقی مانده کاهش بیشتری در مقایسه با سؤال ۵ پیدا

[Downloaded from rbs.mui.ac.ir on 2024-04-19]
[DOR: 20.1001.1.17352029.1392.11.6.15.6]

می‌کنند. سپس حذف هر یک از سؤال‌ها ۵، ۱۰ و ۱۱ به طور جداگانه با سؤال ۴ بررسی شدند که کمترین مقادیر باقی مانده و بهترین شاخص‌های برازش در حالتی به وجود آمده که سؤال‌ها ۴، ۱۱، ۵ و ۳ حذف شدند مقادیر شاخص‌های برازش در این شرایط عبارت بودند از: P= ۰/۰۰۲۸ =۴۱/۹۵، $\chi^2_{.}$= ۰/۰۴۸= RMSEA، CFI=۰/۹۸، GFI =۰/۹۸، NFI=۰/۹۸ و از همه مهم‌تر SRMR =۰/۰۴۷ است. به علاوه تحت چنین شرایطی هیچ‌یک از مقادیر باقی مانده استاندارد شده از ۲/۸۵ بزرگ‌تر نبودند.

جدول ۱. شاخص‌های اصلاح برخی از سؤال‌ها با بیشترین مقدار

سؤال / سؤال	۳	۴	۵	۹	۱۰
۲	۲۲/۲۳	۲۴/۴۷	۱۲/۸۵	۸/۸۲	–
۸	–	۱۲/۰۷	–	–	–
۹	–	۱۸/۲۱	۱۱/۹۹	–	–
۱۰	–	۱۶/۴۳	۲۱/۰۴	–	–
۱۱	۱۰/۵	۲۳/۹۴	۱۷/۳	۶۶	۴۴/۶۲
۱۲	–	۹/۳۶	۵/۹۷	–	۱۱/۱۲

یافته‌ها

نتایج تحلیل اولیه نشان داد برخی از شاخص‌های برازش در سطح مطلوبی نیستند. P= ۰/۰۰۰۰۱ =۴۰۰/۹۳، $\chi^2_{۵۴}$ =۰/۱۲= RMSEA، CFI=۰/۹۳، GFI =۰/۹۷، NFI=۰/۹۲ و از همه مرتبط‌تر به بحث حاضر SRMR =۰/۰۸۳ است. اگر چه برخی از این شاخص‌ها برازش نسبی مدل را نشان می‌دهند ولی آن‌چه بیشتر به بحث حاضر مربوط است باقی مانده‌های استاندارد شده و شاخص SRMR است که نشان می‌دهند کوواریانس غیر مشترک برخی از سؤال‌ها از دامنه مطلوب بیشتر است. بنابراین بررسی بیشتر در این خصوص را با استفاده از واریانس باقی‌مانده انجام خواهیم داد.

مقادیر باقی مانده استاندارد شده و شاخص‌های اصلاح نشان می‌دهند که برخی از سؤال‌ها دارای باقی مانده‌های بسیار بزرگی هستند طوری که برآورد واریانس باقی‌مانده غیر مشترک بین آن‌ها افزایش زیادی در بهبود مدل دارد. روشن است که با توجه به مسأله پژوهش حاضر این شاخص‌های باقی مانده هستند که با توجه حساسیت زیادشان در خصوص واریانس باقی مانده میزان تک‌بعدی بودن نشانگرها را بهتر نشان می‌دهند و دلیل تأکید بر استفاده از آن‌ها برای این منظور نیز همین است. نتایج مربوط به باقی مانده‌های استاندارد شده که تا حدودی دارای کجی مثبت است حاکی از کم برآورد شدن اکثر پارامترها است.

بزرگ‌ترین باقی مانده استاندارد شده ۱۱/۲۳ (بین سؤال ۴ و ۵) و کوچک‌ترین باقی مانده‌های استاندارد ۳/۹۷- (بین سؤال ۴ و ۱۱) و ۳/۴۷- (بین سؤال ۵ و ۱۱) است. شاخص‌های اصلاح نیز حاکی از آن هستند که برآورد کردن کوواریانس خطا برای برخی از سؤال‌ها مقدار کای دوی مدل را به میزان قابل توجهی کاهش می‌دهد. مثلاً شاخص اصلاح برای سؤال‌ها ۴ و ۵ نشان می‌دهند که برآورد کردن کوواریانس بین باقی مانده‌های این دو سؤال با سؤال‌ها ۲، ۸، ۹، ۱۰، ۱۱، ۱۲، به ترتیب به میزان قابل توجهی از مقدار کای دو می‌کاهد. مقادیر برخی از بزرگ‌ترین شاخص‌های اصلاح در جدول ۱ ارایه شده است.برای بررسی‌کردن تک‌بعدی بودن نشانگرها باید هر بار یکی از سؤال‌ها مورد نظر را حذف کرد و نتایج را بررسی کرد. حذف کردن بی‌مورد سؤال‌ها با هم کار درستی نیست زیرا گاهی وجود واریانس باقی‌مانده غیر مشترک بین دو سؤال باعث می‌شود بین سایر زوج‌های سؤال نیز واریانس

[Downloaded from rbs.mui.ac.ir on 2024-04-19]
[DOR: 20.1001.1.17352029.1392.11.6.15.6]

باقی مانده معنی‌دار به وجود آید (۶).

بحث و نتیجه‌گیری

بررسی تک‌بعدی بودن نشانگرهای یک سازه یکی از ویژگی‌های است که باید در مرحله ساخت مقیاس مورد توجه قرار بگیرد. متأسفانه بسیاری از سازندگان مقیاس‌ها بدون توجه به این مطلب یا بررسی، این موضوع را نادیده می‌گیرند و یا با استفاده از روش‌های نامناسبی این موضوع را مورد بررسی قرار می‌دهند. در ارتباط با مثال حاضر اگر قرار بود بر اساس ضریب آلفا ۰/۸ برای این خرده مقیاس در مورد این ویژگی قضاوت کنیم مشخص نمی‌شد که واریانس باقی مانده برخی از سؤال‌ها نظیر دو سؤال ۴ و ۵ با بسیاری از سؤال‌ها دیگر بالا است. تنها شاخص‌های برازش مثل SRMR و توزیع باقی مانده‌های استاندارد شده که از مقادیر مطلوب فاصله زیادی داشتند این مطلب را نمایان ساختند. سایر شاخص‌های برازش مثل CFI یا GFI دارای مقادیر مطلوبی بودند طوری که قضاوت صرف بر اساس آن‌ها می‌توانست گمراه کننده باشد.

جالب آن که دو سؤال ۴ و ۵ که با بسیاری از نشانگرها دارای کوواریانس غیر مشترک بودند از نظر محتوایی تقریباً یکسان هستند. محتوای دو سؤال ۴ و ۵ به ترتیب از این قرار بود "از زمان بچه‌دار شدن، نتوانسته‌ام کارهای جدید و متفاوتی انجام دهم" و "احساس می‌کنم از زمان بچه‌دار شدن، هرگز نتوانسته‌ام کارهای مورد علاقه‌ام را انجام دهم".

این که چرا افراد استفاده‌کننده از این ابزار به این موضوع توجه نکرده‌اند جای تأمل دارد. البته امروز در کشور ما ابزارهای زیادی وجود دارند که حتی در آن‌ها اشکالات نگارشی و تایپی وجود دارند ولی افراد بدون توجه به این موضوع از آن‌ها استفاده می‌کنند بدون آن که زحمت ویرایش متن سؤال‌ها را به خود بدهند. از این گذشته روند ترجمه ابزارهای خارجی نیز بر این مشکلات می‌افزاید و گاه مفاهیم به‌طور نادرست ترجمه می‌شوند بدون آن‌که کسی به آن ها پی ببرد.

البته بررسی این‌گونه ویژگی‌ها باید در مرحله ساخت مقیاس مورد توجه قرار گیرد. متأسفانه امروزه برخی از سازندگان مقیاس فقط بر اساس نتایج تحلیل‌های آماری از جمله تحلیل عاملی اکتشافی سؤال‌ها را در عامل‌ها قرار می‌دهند بدون آن که به محتوای سؤال و این که انطباق معنایی عبارت سؤال با عامل مورد نظر را بررسی کنند.به علاوه نتایج نشان داد که وجود سؤالاتی از نوع فوق می‌تواند باعث به وجود آمدن کوواریانس بین بسیاری از زوج‌های سؤال‌ها شود. بنابراین بررسی توزیع مقادیر باقی مانده بر اساس رویکرد مدل‌سازی معادله ساختاری می‌تواند در این خصوص بسیار مفید و نویدبخش باشد.

همان‌طور که توصیف شد روش‌های کلاسیک مثل آلفا یا سایر روش‌های برآورد اعتبار و یا تحلیل عاملی مشترک نمی‌توانند ویژگی تک بعدی بودن سؤال‌ها یک ابزار و یا خرده مقیاس‌ها را تعیین کنند. نکته قابل تأمل آن است که برخی افراد برای ابزارهایی که دارای چندین خرده مقیاس هستند در کل یک آلفا محاسبه می‌کنند بدون آن که ماهیت متفاوت سؤال‌های هر خرده مقیاس را مدنظر قرار دهند. روشن است که ارایه هر نوع تحلیل آماری بر اساس نتایج این‌گونه ابزارها می‌تواند به شدت گمراه کننده باشد. به عبارتی محاسبه آلفا باید به تفکیک هر یک از خرده مقیاس‌های یک ابزار صورت گیرد نه بر اساس کل سؤال‌های ابزار مورد نظر. به‌علاوه یکی از اقدامات لازم در مرحله ساخت ابزار تعیین مدل اندازه‌گیری سؤال‌های مربوط به یک مقیاس (تاو معادل، اساسا تاو معادل، هم‌جنس و ...) و سپس محاسبه اعتبار بر اساس آن است.

در بسیاری از مقالات علمی یکی از راهکارهای افزایش برازش مدل بازگذاشتن کوواریانس بین عبارات خطا است، به‌طوری که شاخص‌های برازش بهبود یافته و می‌توان نتایج را گزارش کرد. اما این موضوع به لحاظ نظری محل تردید است. وجود کوواریانس معنی‌دار بین عبارات باقی مانده دلیل بر وجود عامل مشترک دیگری در بین سؤال‌ها است. یعنی علاوه بر آن چه که سؤال‌ها با هم اندازه‌گیری می‌کنند و ما آن را عامل مشترک می‌نامیم، عوامل دیگری نیز در کوواریانس بین سؤال‌ها نقش داشته‌اند. ممکن است برخی وجود کوواریانس معنی‌دار بین عبارات باقی مانده را به عوامل اختصاصی نسبت دهند، اما چون عامل اختصاصی فقط خاص هر یک از متغیرها است، این مطلب نمی‌تواند وجود کوواریانس بین سؤال‌ها را تبیین کند. به عبارتی

[Downloaded from rbs.mui.ac.ir on 2024-04-19]

[DOR: 20.1001.1.17352029.1392.11.6.15.6]

عامل اختصاصی که فقط توسط هر یک از متغیرها اندازه‌گیری می‌شود باعث وجود کوواریانس بین متغیرها نمی‌گردد. به علاوه نسبت دادن آن به خطا نیز که خود دارای توزیع تصادفی است و در بافت یک بار اندازه‌گیری قابل برآورد نیست نمی‌تواند درست باشد. در صورتی که عامل به وجودآورنده کوواریانس عوامل ناشی از خطای محیطی باشند عامل خطا می‌تواند باعث بروز کوواریانس بین عبارات خطا شود هر چند که معمولاً تلاش می‌شود این گونه موارد در موقع اجرا کنترل شوند. وجود واریانس منفی در برخی عبارات خطا (که البته به لحاظ نظری غیر ممکن است) می‌تواند سرنخ‌هایی در خصوص خطای محیطی باشد، هرچند که همیشه این طور نیست و نباید همه این گونه موارد را به عوامل محیطی نسبت داد. زیرا گاهی مشخص شدن نادرست مدل یا متغیرها به واریانس منفی بین عبارات خطا منجر می‌شود. به علاوه توصیه می‌شود که پژوهشگران توزیع باقی‌مانده‌های استاندارد شده را نیز بررسی کنند و منفی مثبت بودن مقادیر را مورد توجه قرار دهند این موضوع می‌تواند نشان دهد که مدل بیش برآورد داشته (over estimate) یا کم برآورد (under estimate) که خود می‌تواند بسیار روشنگر باشد.

در ارتباط با این موضوع می‌توان به مطلب دیگری نیز اشاره کرد. آیا امکان دارد ارایه سؤال‌ها به افراد مختلف به تداعی‌های متفاوتی در ذهن آن‌ها منجر شود و خود این موضوع باعث ایجاد واریانس‌های مختلف در نمرات یک سؤال گردد؟ به عبارتی آیا می‌توان گفت متفاوت بودن فرایند ذهنی در بین افراد در خصوص یک سؤال منجر به واریانس‌های متفاوت شود؟نتایج حاصل از پژوهش‌ها در حوزه آزمون‌های شناختی این موضوع را تأیید کرده‌اند که فرایندهای ذهنی متفاوت باعث ایجاد واریانس متفاوت می‌شود (۳۰). در حوزه آزمون‌های شخصیتی حداقل ما با پژوهشی در این خصوص روبرو نشدیم. در هر حال خود این موضوع می‌تواندقابل بررسی باشد. با توجه به این مطالب، وجود کوواریانس بین عبارات خطا دلیل بر وجود عامل مشترک دیگری است که منجر به نقض پیش‌فرض تک‌بعدی بودن خواهد شد. نقض پیش فرض تک‌بعدی بودن کل فرایند اندازه‌گیری را مختل خواهد کرد، هر چندکه این موضوع برای ما زیاد محسوس نیست و به آن توجه نمی‌کنیم. مبحث بعدیت آزمون‌ها با مباحث روایی، اعتبار، کارکرد متفاوت سؤال، بانک سؤال، سنجش انطباقی و معادل‌سازی آزمون‌ها ارتباط دارد و همه آن‌ها را تحت تأثیر آن قرار می‌دهد. بنابراین لزوم توجه به آن در مرحله ساخت ابزار بسیار ضروری است.

References

1. Simms LJ. Classical and modern methods of psychological scale construction. Social and Personality Psychology Compass 2008;2(1):414-33.
2. Lord FM, Novick MR, Birnbaum A. Statistical theories of mental test scores. Birnbaum, Allan Oxford, England: Addison-Wesley; 1968.
3. Hambleton RK. Principles and selected applications of item response theory. In: Linn RL, editor. Educational Measurement New York: Macmillan Publishing Company; 1989.
4. Segars AH. Assessing the unidimensionality of measurement: A paradigm and illustration within the context of information systems research. Omega 1997;25(1):107-21.
5. Jöreskog KG, Sörbom D. LISREL 8 user's reference guide. Chicago: Scientific Software International; 1996.
6. Gefen D. Assessing unidimensionality through LISREL: An explanation and an example. Communications of the Association for Information Systems 2003;12(1):23-47.
7. Clark LA, Watson D. Constructing validity: Basic issues in objective scale development. Psychological assessment 1995;7(3):309-19.
8. Raykov T. Bias of Coefficient afor Fixed Congeneric Measures with Correlated Errors. Applied Psychological Measurement 2001;25(1):69-76.
9. Green SB, Hershberger SL. Correlated errors in true score models and their effect on coefficient alpha. Structural equation modeling 2000;7(2):251-70.
10. Schmitt N. Uses and abuses of coefficient alpha. Psychological assessment 1996;8(4):350-3.
11. Zimmerman DW, Williams RH. Is classical test theory" robust" under violation of the assumption of

[Downloaded from rbs.mui.ac.ir on 2024-04-19]
[DOR: 20.1001.1.17352029.1392.11.6.15.6]

uncorrelated errors? Canadian Journal of Psychology/Revue canadienne de psychologie 1980;34(3):227-37.
12. Novick MR. The axioms and principal results of classical test theory. Journal of Mathematical Psychology 1966;3(1):1-18.
13. McDonald RP. The dimensionality of tests and items. British Journal of Mathematical and Statistical Psychology. 2011;34(1):100-17.
14. Henning G. Meanings and implications of the principle of local independence. Language Testing. 1989;6(1):95-107.
15. Nandakumar R. Traditional dimensionality versus essential dimensionality. Journal of Educational Measurement 2005;28(2):99-117.
16. Stout W. A nonparametric approach for assessing latent trait unidimensionality. Psychometrika 1987;52(4):589-617.
17. Yen WM, Fitzpatrick AR. Item response theory. In: Brennan RL, editor. Educational measurement, 4th ed. Westport, CT: American Council on Education/Praeger; 2006:111-53.
18. Gerbing DW, Anderson JC. An updated paradigm for scale development incorporating unidimensionality and its assessment. Journal of marketing research 1988;25(2):186-92.
19. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16(3):297-334.
20. Wainer H, Thissen D. How is reliability related to the quality of test scores? What is the effect of local dependence on reliability? Educational Measurement: Issues and Practice 2005;15(1):22-9.
21. Green SB, Lissitz RW, Mulaik SA. Limitations of Coefficient Alpha as an Index of Test Unidimensionality 1. Educational and Psychological Measurement 1977;37(4):827-38.
22. Huysamen G. Coefficient alpha: unnecessarily ambiguous; unduly ubiquitous. SA Journal of Industrial Psychology 2007;32(4):34-40.
23. Hattie J. Methodology review: Assessing unidimensionality of tests and ltenls. Applied Psychological Measurement. 1985;9(2):139-64.
24. Graham JM. Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability What They Are and How to Use Them. Educational and Psychological Measurement 2006;66(6): 930-44.
25. Segars AH. Scale development in information systems research: A paradigm incorporating unidimensionality and its assessment. Proceedings of the Fifteenth International Conference on Information Systems; 1994 Dec 14-17; Vancouver; Canada; 1994.
26. Jasper F. Applied dimensionality and test structure assessment with the START-M mathematics test. International Journal of Educational and Psychological Assessment 2010;6(1):104-25.
27. Byrne B. Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming: Mahwah, N.J.: Lawrence Erlbaum Associates; 1998.
28. Fadaei Z, Dehghani M, Tahmasian K, Farhadei F. Investigating reliability, validity and factor structure of parenting stress-short form in mothers of 7-12 year-old children. Journal of Research in Behavioural Sciences 2010;8(2):81-91.
29. Raykov T, Marcoulides GA. A first course in structural equation modeling. Mahwah, N.J.: Lawrence Erlbaum Associates; 2006.
30. Ackerman T, Bolt D, editors. How different cognitive strategies produce differential item performance. Proceedings of the conference of the American Educational Research Association, San Francisco, CA; 1995.

[Downloaded from rbs.mui.ac.ir on 2024-04-19]

[DOR: 20.1001.1.17352029.1392.11.6.15.6]

The necessity of unidimensionality of indicators in the behavioral and medicine sciences measurements: Application of structural equation modeling

Balal Izanloo [1], Mojtaba Habibi [2], Reza Bagherian-Sararoudi [3]

Original Article

Abstract

Aim and Background: The unidimensionality feature of indicators of a construct is rooted in the classical test theory (CTT), and in the context of CTT the term of being Homogeneity for itemswhich was developed to measure a construct indirectly refers to this topic. Misunderstanding of the term has led many of researchers use exploratory factor analysis or Cronbach's alpha procedures to investigating and determination of unidimensionality. While, none of these procedureswere not design for this propose.

Methods and Materials: Using the survey design, 468 mothers with multiphase cluster sampling method fill in the PSI-SF. The data was analysis by structural equating model using LISREL software.

Findings: Using structural equation modeling approach and related software's, which provide investigation of distribution of residuals, one could test Presence or absence of the unidimensionality feature of indicators.

Conclusions:This study, in addition to theoretical discussions, examined this issue with a sample of empirical data.

Keywords: Structural Equation Modeling, Unidimensionality, Latent Trait

Citation: Izanloo B,Habibi M, Bagherian-Sararoudi R.**The necessity of unidimensionality of indicators in the behavioral and medicine sciencesmeasurements: Application of structural equation modeling**. J Res Behave Sci 2014; 11(6): 671-684

Received: 03.03.2014 Accepted: 04.03.2014

1- PHD student of Educational Measurement, School of Educational Sciences and Psychology, Tehran University, Tehran, Iran:
2- Assistant Professor, Family Reseach Center, University of ShahidBeheshti, Tehran, Iran
3- Associate Professor, Behavioral Sciences Research Center, Department of Psychiatry, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran(Corresponding Author) Email: bagherian@med.mui.ac.ir

[Downloaded from rbs.mui.ac.ir on 2024-04-19]

[DOR: 20.1001.1.17352029.1392.11.6.15.6]