

ضروریات طراحی و رواسازی مقیاس‌های روانشناختی: راهنمایی برای پژوهشگران

حجت‌اله فراهانی^{۱*}، رسول روشن چسلی^۲

۱. استادیار گروه روانشناسی دانشگاه تربیت مدرس، تهران، ایران. (نویسنده مسئول)

۲. استاد گروه روانشناسی دانشگاه شاهد، تهران، ایران

چکیده

مقیاس‌سازی فعالیتی رایج در پژوهش‌های علوم روانشناختی و رفتاری است، اما بسیاری از مقیاس‌های ساخته‌شده نارسا، نادقیق، کم‌مایه و پایه و به‌ندرت میان‌مایه و مطلوب‌اند. طرح‌اندازی یک مقیاس نو، رواسازی و اعتباریابی آن، نه تنها به دوراز پیچ‌وخم نیست، که فعالیتی ناهموار و نامانوس است که نیازمند دقت و ریزینی است. بنابراین هدف از این مقاله، بررسی گام‌های ضروری و گاه غایب در مقیاس‌سازی‌های کنونی و به‌ویژه برای نوپژوهشگران است. در این مقاله بنیان‌های نظری و گام‌های عملی توانمند مورد نظر بوده‌اند و از کوته‌نگری نظری و وانهادگی عملی دوری شده است.

نشریه علمی

روانشناسی
بالینی
و شخصیت

دوره ۱۷، شماره ۲، پیاپی ۳۳
پاییز و زمستان ۱۳۹۸
صص: ۱۹۷-۲۱۲

تاریخ دریافت: ۱۳۹۷/۰۹/۱۸

تاریخ پذیرش: ۱۳۹۸/۰۳/۲۸

مقاله پژوهشی

Journal of

**Clinical
Psychology
& Personality**

Vol. 17, No. 2, Serial 33

Autumn & Winter
2019 - 2020
pp.: 197-212

کلیدواژه‌ها: مقیاس‌سازی، روان‌سنجی، روایی، اعتبار، تحلیل عاملی

*Email: h.Farahani@modares.ac.ir

مقدمه

در این سالها، شمار فراوانی مقیاس اندازه‌گیری برای به درفکندن سازه‌های روانشناختی به دام کمیت، طراحی و ساخته شده اند [۱]. این مقیاس‌ها در گستره گسترده‌ای از کاربردهای علمی طرح‌اندازی شده‌اند.

یک مقیاس همان رخ‌نمایی آشکار سازه(های) پنهان^۲ است. یک مقیاس، نگرش، رفتار و صفات مفروضی را اندازه می‌گیرد، در نتیجه درکی نظری از جهان پیرامون به دست می‌آید که بی‌واسطه اندازه‌پذیر نیست [۲]. بنابراین مقیاس‌ها، ابعاد کمیت‌پذیر پدیده‌ها را اندازه می‌گیرند، مقیاس‌ها گردآیه^۳‌ای از گویه‌هایی هستند که در جاتی از متغیرهای نظری بدست می‌دهند، متغیرهایی که با ابزارهای مستقیم مشاهده‌شده نیستند [۳]. به دیگر سخن، یک مقیاس برای برون‌آری^۴ یک رفتار، احساس یا یک کنش به کار گرفته می‌شود که نمی‌توان آنها را از دل یک تک‌گویه یا مفهوم یگانه برون‌آورد [۴].

اندازه‌گیری قلب علم و فعالیت است پایه‌ای و بنیادین، چرا که، به پژوهشگران این شانس را می‌دهد تا در مورد انسان، اشیا و فرایندها شناخت و معرفتی بدست آورند. هر چند نزدیک به بی‌شمار، مقیاس ساخته شده‌اند و این مقیاس‌ها پهنه گسترده‌ای از تجربه‌ها، رفتارها و ویژگی‌ها را اندازه می‌گیرند، با بالندگی فزونی گرایانه علم و پرسشگری‌های نوین پژوهشی، نیاز به ساخت مقیاس‌ها و ابزارهای نو نیز شدت می‌گیرد.

از دیگر سو، مقیاس‌سازی، راهی کاملاً روشن و به دور از پیچ و خم نیست، در این راه، افزون بر زمان و هزینه، چیرگی نظری بر قلمرو پژوهشی و به ویژه سازه مورد پژوهش، و زبردستی در تحلیل‌های آماری پیچیده نیز بس ضرورت می‌یابد.

شماری از پژوهشگران [۳، ۴، ۵، ۶] بر این باورند که فرآیند مقیاس‌سازی روندی نظام‌مند و پیچیده است که ریزبینی روش‌شناختی و نظری از استلزامات آن است. اگرچه، پژوهشگران این قلمرو و دیگر قلمروهای وابسته، کوششی رسا کرده‌اند تا مقیاس‌ها و ابزارهایی برای اندازه‌گیری سازه‌های روان‌شناختی طراحی، رواسازی و اعتبار کنند، اما همچنان شمارگان چشمگیری از آن ابزارها زیر پرسش است.

هدف از این مقاله بررسی نارسایی‌های پنهان شده در مقیاس‌های موجود، در گستره نظری و ارائه راهکارهایی برای چیرگی بر نارسایی‌های آنها و ایجاد پایه‌هایی استوار و مایه‌های بسنده برای همه پژوهشگران این حوزه و بویژه برای نوپژوهشگران است. در این مقاله سعی بلیغی شده است که تا می‌توان، این فرآیند در چهارچوبی کم‌ابهام، سراسر و کاربردی چنان توصیف شود که برای همه پژوهشگران حوزه علوم انسانی در ساخت مقیاس‌های نو، روا و معتبر از یک سو و بهسازی مقیاس‌های موجود از دیگر سو، چراغی تابان و تکیه‌گاهی مطمئن باشد.

گذری کوتاه بر مقیاس‌سازی

برای ساخت، رواسازی و اعتباریابی یک مقیاس مرحله‌ها و گام‌های مختلفی وجود دارد. هنکین (۱۹۹۵) سه مرحله اصلی برای آن توصیف می‌کند [۷] که عبارتند از:

(۱) گویه‌سازی^۵

(۲) مقیاس‌سازی^۶

(۳) مقیاس‌آزمایی^۷

که این مرحله‌ها به گام‌های کوچکی شکسته می‌شوند.

۱. گویه‌سازی

ساخت گویه، طرح جمله‌های گزاره‌ای، پرسشی و یا عبارت‌هایی است که در شکل پایانی یک مقیاس برخی از آنها خواهند ماند و برخی دیگر، در مراحل و گام‌هایی از مقیاس اولیه کنار گذاشته می‌شوند.

این مرحله در برگزیده روش‌ها و گام‌هایی است. این مرحله همان سنگ‌سخت و زیرین است که مقیاس بر آن بنا نهاده می‌شود.

در این مرحله، شناسایی حوزه^۸ (ها)ی مورد بررسی، ساخت گویه، و بررسی تحلیل محتوای آنها اهمیت ویژه دارد. یک حوزه یا سازه همان مفهوم، صفت یا رفتار مشاهده‌ناپذیری است که هدف نهایی پژوهش، اندازه‌گیری و سنجش آن است [۸]، بنابراین نخستین گام در گویه‌سازی تصمیم در مورد حوزه و گستره مورد بررسی است [۹]، اگر حوزه‌ی مورد بررسی رساوروشن تعریف شود، معرفت‌کارایی از پدیده مورد نظر را بدست می‌دهد، این دانش مرزهای حوزه را مشخص می‌کند و فرآیند گویه‌سازی را آسان کرده و روایی محتوای مقیاس را استحکام می‌بخشد.

¹ manifestation

² latent constructs

³ set

⁴ capturing

⁵ item generation or item development

⁶ scale development

⁷ scale testing

⁸ domain

جامعه‌ی هدف، بهره‌می‌جوید [۱۴، ۱۵]. مثلاً تحلیل بن‌مایه^۸ [۱۶]، تحلیل محتوا^۹ [۱۶]، گروه‌کانونی^{۱۰} [۱۶] و گروه‌کارشناس^{۱۱} [۱۶].

در روش ترکیبی روش‌قیاسی و استقرایی توأمان به کار بسته‌می‌شوند، در حالیکه بررسی گستره نظری، بنیان‌های نظری لازم برای تعریف یک حوزه را فراهم می‌آورد، انجام پژوهش کیفی آن حوزه را از یک نقطه انتزاعی به یک مبنای آشکار از آن مفهوم انتزاعی، سو می‌دهد.

استفاده همگام از روش‌قیاسی و استقرایی در تعیین گویه‌های مقیاس توصیه می‌شود [۱۷].

بررسی گستره نظری، دیدگان پژوهشگر رابه تعداد سازه‌ها، چستی^{۱۲} و گونه‌گونی^{۱۳} محتوای آن‌ها می‌گشایدوشامه‌اورا نسبت به ابزارهای مشابه و موجود تیزتر می‌کند [۳، ۱۸]. از مصاحبه و گروه‌کانونی به عنوان دو روش رایج در رویکرد استقرایی استفاده می‌شود. همچنین بهره‌مندی از گروه صاحب‌نظران، دراطمینان‌گیری پژوهشگر از روایی محتوایی ابزار نو و دستیابی به اطلاعات اصیل و ناب بسیار رهگشا است [۹].

آنچه در این گام اهمیت چشمگیر می‌یابد تعیین تعداد بسنده گویه‌های اولیه است. تعداد گویه‌ها بر اساس منابع مختلف در سبد گویه‌های اولیه متفاوت است.

کلاین [۱۹] و اسکانیکار [۲۰] بر این باورند که دست‌کم در سبد گویه‌های اولیه باید دو برابر تعداد گویه‌های مورد نظر در نسخه نهایی مقیاس گویه باشد.

برخی بر این باورند که این تعداد باید ۵ برابر تعداد گویه‌های نسخه پایانی مقیاس باشد [۴].

نگارندگان این مقاله بر مبنای تجارب خود و هماهنگ با کلاین و اسکانیکا ۲ برابر با ۱/۵ برابر را بسنده می‌دانند چنانچه گویه‌ها در سبد گویه‌های اولیه با ریزبینی دست‌چین و نیکوگزیده شده باشد.

عبارت‌پردازی و نگارش گویه‌ها باید ساده و روشن بوده و پژوهشگر باید از گویه‌های دو گانه منفی (من احساس

برای تعیین حوزه یا سازه باید توجه داشت که ابزاری از پیش موجود در آن حوزه وجود نداشته باشد، ابزاری که برای هدف پژوهشی پژوهشگر بسنده است، در چنین احوالی، پژوهشگر مقیاس‌ساز، باید ضرورت ساخت مقیاسی جدید و تفاوت‌های آن با ابزارهای موجود را به درستی توجیه و تبیین کند. پس از تعیین حوزه مورد بررسی پژوهشگر باید کوشش کند تا خرد - حوزه‌ها یا ابعاد آن حوزه را نیز تعیین کند [۴].

باید در نظر داشت که اگر سازه علمی کاملاً انتزاعی است، خرده‌مقیاس‌ها باید دست‌کم سه متغیر را در بر گرفته تا بتوانند هسته واقعی مفهوم را برون آورده و روایی محتوایی را استحکام بخشند [۱۰] گویه‌های تجربی متعدد خطر تاثیرات فرهنگ و سوگیری ترتیب‌گویی‌ها درمقیاس را کاهش می‌دهند.

پس از تعیین و روشن‌سازی حوزه (سازه) (مفهوم) اصلی پژوهش، سبد گویه^۱ اولیه را می‌توان تعیین کرد، گاهی به این فرآیند طراحی پرسش [۱۱] و گاهی گویه - سازی [۱۲] نیز گفته می‌شود.

برای پر کردن سبد گویه‌های اولیه پژوهشگر نیاز به فرآهم‌آوری یک چهارچوب نظری دارد [۴، ۹].

روشهای گویه‌سازی اولیه به سه روش قیاسی^۲، استقرایی^۳ و ترکیبی^۴ دسته‌بندی می‌شوند. روش قیاسی برای گویه‌سازی متکی بر بررسی گسترده و عمیق گستره نظری و مقیاس‌های پیشین است [۷] این روش افزاز منطقی^۵ یا دسته‌بندی از بالا^۶ است [۱۳]. بیشتر پژوهشهای مقیاس‌سازی در ایران از این روش پیروی می‌کنند و عمدتاً برپایه مقیاس‌های موجود در فرهنگ غیر ایرانی بوده و صرفاً ترجمه شده، تنها از چنین رهگذری، سبد گویه‌های پژوهشگر برای گام پیش‌رو پر شده است.

درروش‌های استقرایی، موسوم به روش گروه‌بندی یا دسته‌بندی از پایین^۷، گویه‌سازی بر پایه اطلاعات کیفی است، اطلاعاتی که در قالب یک پژوهش کیفی از عمق ذهن

¹ item pool

² deductive

³ inductive

⁴ combined

⁵ logical partitioning

⁶ classification from top

⁷ classification from bottom

⁸ Thematic analysis

⁹ content analysis

¹⁰ focus group

¹¹ expert panel

¹² nature

¹³ diversity

تعیین می‌شود.

روایی محتوایی براساس داوری متخصصان

متخصصان که به ارزیابی روایی محتوایی گویه‌ها می‌پردازند باید سوای آنانی باشند که گویه‌های اولیه را ساخته و در سبد گویه‌ها نشانده‌اند. عموماً ارزیابی چند تایی^۵ در روایی محتوا به کار گرفته می‌شود و بطور معمول ۷ تا ۵ متخصص برای این هدف نیاز است. [۲۶].

روشهای مختلفی برای محاسبه روایی محتوایی وجود دارد از جمله شاخص روایی محتوایی^۶ (CVI) و نسبت روایی محتوایی^۷ (CVR)

ویلسن، پن و اسکاسکی (۱۰۱۲) معتقد که CVR یک گویه آماره مفیدی در نگهداشت یا برون داشت گویه‌های اولیه در سبد گویه‌ها است و این شاخص را پژوهشگران جهانی روشی مناسب برای رسیدن به روایی محتوایی می‌دانند. CVI برای کل مقیاس برابر یا متوسط CVR گویه‌هایی است که در نسخه نهایی مقیاس می‌مانند. برای تعیین CVR از جدول ۱ استفاده کنید.

جدول ۱. الگوی تعیین CVR

گویه‌ها	ضروری است	مفید است اما ضروری نیست	غیرضروری است
۱-			
۲-			
۳-			
.			
.			

برای محاسبه CVR از فرمول زیر (روش لوشه^۸) استفاده می‌شود:

$$CVR = \frac{ne - \frac{N}{2}}{\frac{N}{2}}$$

ne: تعداد متخصصانی که گویه را ضروری میدانند
N: تعداد کل متخصصان

نامیدی نمی‌کنم)، گویه‌های ترکیبی (من احساس پریشانی و دل‌تنگی می‌کنم) دوری‌گزینند.

در مورد تعداد گزینه‌های پاسخ، کروسیک و پرسر [۲۱] نشان دادند که گزینه‌های دو تا سه درجه‌ای کمترین اعتبار را نسبت به طیف لیکرت پنج تا هفت درجه‌ای دارند.

در این گام دامی نهفته است، این دام، مقیاس پژوهشگر را آن‌گاه در کام خود می‌کشد که او بدرستی حوزه (سازه) مورد نظر خود را مفهوم‌سازی و تعریف نکرده باشد در این احوال، مقیاسی بر ساخته می‌شود که به دلیل ابهام در تعریف حوزه (سازه) یا تعریف نابسنده آن به شدت نارساست. از دیگر پیامدهای آن، در هم آمیختگی^۱ است؛ یک سازه با دیگر سازه‌های موجود در همان قلمرو همپوشانی خواهد داشت [۲۲]. دام پنهان دیگر در این گام، کم‌نمودگی^۲ یک سازه در مقیاس است، کم‌نمودگی یک مقیاس زمانی روی می‌دهد که ابعاد با اهمیت یک سازه به دلیل تنگ‌نگری پژوهشگر در مقیاس نهایی کم‌نمود پیدا می‌کند.

روایی محتوایی گویه‌ها

روایی محتوایی که تحلیل نظری، نیز نام گرفته، ویژگی بسیار مهم است. قضاوت متخصصان، اساسی‌ترین روش برای تحلیل روایی محتوایی است [۲۳]. می‌توان گفت تحلیل محتوا، ربط محتوایی^۳ و نمود محتوایی^۴ را مشخص می‌کند [۲۴] در مورد تعداد متخصصان سه نفر قابل قبول و متداول، ۵ تا ۱۰ نفر مرجح و بیش از ۱۰ کارشناس شاید غیر ضروری باشد.

گوین (۱۹۷۷) بر این باور است که پنج شرط برای رسیدن به هر شکل از روایی محتوا ضروری است [۲۵]:
الف) محتوا، تعریف و معنای عموماً قابل پذیرش داشته باشد.

ب) سازه روشن تعریف شود؛

ج) محتوای سازه با هدف اندازه‌گیری پیوند منطقی داشته باشد.

د) ارزیابی کیفی از حوزه مورد بررسی بر اساس نظر جمعی (اجماع) بسنده باشد و

ه) محتوی پاسخ باید مشاهده و آزمون پذیر می‌باشد داشته باشد.

روایی محتوا با ارزیابی گروهی از متخصصان جامعه هدف

¹ contamination

² underrepresentation

³ content relevance

⁴ content representation

⁵ multiple judies

⁶ content validity index

⁷ content validity ratio

⁸ Lawshe

در این گام واکنش‌های کلامی و رفتاری مورد ریزبینی و تیز بینی قرار می‌گیرد؛ برای نمونه، از پاسخگویان خواسته می‌شود در مورد گویه‌ها نظر دهند یا بگویند که برخی از گویه‌ها چه معنایی برای آنها دارد، رفتارهای پاسخگویان شامل درنگ یا اخم در هم کشیدن که می‌توانند نشانه‌ای از وجود مشکلی در گویه باشد از نظر پژوهشگر پنهان نمی‌ماند. بر اساس بازخورد از این گام باید عبارات، دستورالعمل و طرح پرسشنامه ویرایش شود [۳۳، ۳۴، ۳۵، ۳۶].

مرحله دوم: واریسی مقیاس

در مرحله پیشین گویه‌ها طراحی و تحلیل می‌شوند در این مرحله مقیاس نهایی شده مورد ارزیابی قرار می‌گیرد. در این مرحله گام‌هایی را باید در نظر داشت.

تعیین حجم نمونه و روش نمونه‌گیری

تعیین حجم نمونه در مقیاس‌سازی برای اندازه‌گیری یک سازه پنهان بحث برانگیز است. توصیه اصلی این است که گویه‌های مقیاس مورد نظر در نمونه‌ای ناهمگون آزمون شود؛ نمونه‌ای که گستره‌ای از جامعه هدف را در خود بازنمایی کند [۱۸]. نمونه باید معرف جامعه هدف باشد تا این نمودگی بیشینه گردد.

حجم نمونه به جنبه‌های مختلفی از هر پژوهش متکی است. قانون تقریبی^۲ در مقیاس‌سازی نسبت ۱۰:۱ (به ازای هر گویه، ۱۰ پاسخگو) است [۶].

برخی منابع بر این تکیه دارند که پس از گام پیش - آزمایی، با ۳۰۰ پاسخگو می‌توان پژوهش مقیاس‌سازی را پیش برد. پژوهش‌های شبیه‌سازی نشان می‌دهد که دست کم ۳۰۰ تا ۵۰۰ پاسخگو لازم است تا بتوان الگوهای مقایسه پذیری را فراهم آورد و در نمونه‌های زیر ۳۰۰ نتیجه را باید تکرار کرد.

روشهای آماری مقیاس‌سازی نیازمند نمونه‌های بزرگ هستند؛ بنابراین کمتری و لی [۳۶] بر این باورند که نمونه بالای ۵۰۰ بسیار خوب، بالای ۱۰۰۰ عالی و ۳۰۰ قابل قبول است.

به یاد داشته باشید حجم نمونه بالاتر خطای اندازه‌گیری را کاهش و مانایی ساختار عاملی ابزار را افزون می‌سازد. مقدار KMO^3 بیشتر از ۰/۶ نشان دهنده کفایت حجم نمونه است. همچنین کمتری و لی (۱۹۹۲) بر این باورند که میزان اشتراکات و بارهای عاملی کم نشان دهنده نیاز به افزایش حجم نمونه است. میزان اشتراکات h^2 حداقل ۰/۵ و بار عاملی دست کم ۰/۴ باید باشد تا اطمینان داشت که

گویه‌های نامناسب را می‌توان بر اساس مقادیر CVR هرگویه و مقایسه آن با مقادیر بحرانی که در مقاله Ayre and Scally (2014) در دسترس است، از سید گویه‌ها کنار نهاد. این مقادیر بحرانی بر اساس تعداد متخصصان تعیین شده است. یک راه ساده‌تر این است که گویه‌هایی با CVR برابر با ۰/۷۸ یا بیشتر با دست کم سه متخصص از روایی محتوایی مناسبی برخوردارند.

در حالیکه دویس (۱۹۹۲) مقادیر CVI بیشتر از ۰/۸ را مرجع می‌داند [۲۷]، تیلدن، نلسن و می (۱۹۹۰) CVI بیشتر از ۰/۷ را قابل قبول می‌دانند [۲۸].

گزارش CVI کلی ابزار نسبت به گزارش CVR تک تک به تک گویه‌ها در بسیاری از شرایط کارآمدتر و گویاتر است [۲۹]. به این منظور CVI را میتوان میانگین CVR گویه‌هایی دانست که CVR آنها مطلوب شده‌اند.

تعداد خبرگان	مقدار CVR	تعداد خبرگان	مقدار CVR	تعداد خبرگان	مقدار CVR
۵	۰/۹۹	۱۱	۰/۵۹	۲۵	۰/۳۷
۶	۰/۹۹	۱۲	۰/۵۶	۳۰	۰/۳۳
۷	۰/۹۹	۱۳	۰/۵۴	۳۵	۰/۳۱
۸	۰/۹۵	۱۴	۰/۵۱	۴۰	۰/۲۹
۹	۰/۷۸	۱۵	۰/۴۹		
۱۰	۰/۶۲	۲۰	۰/۴۲		

روایی محتوا براساس داوری جامعه هدف

هدف از این ارزیابی، تعیین روایی ظاهری گویه‌ها است که خود بخشی از روایی محتوایی است [۳۰]، کلارک و واتسن (۱۹۹۵) و مالورتا (۲۰۰۴) این فعالیت را در ساخت مقیاس‌های نو بسیار پر رنگ می‌دانند [۳۱، ۳۲]. لازم است که تعدادی از جامعه هدف در قالب یک (پیش - آزمایی) گویه‌های مقیاس نو را واریسی اولیه نمایند. این فعالیت در تعیین نظر و واکنش پاسخگویان به هر گویه از مقیاس بسیار کارا است. نتیجه این فعالیت به پژوهشگر مقیاس‌ساز، این یاری را می‌رساند که گویه‌هایی که با کژدرکی و مهالودگی احتمالی در فهم در نمونه هدف روبرو می‌شوند، شناسایی و پیش از اجرای نهایی مقیاس در یک نمونه بزرگ از آن مقیاس کنارگذاشته شوند.

شمار پاسخگویان در نمونه پیش - آزمایی می‌تواند کم (بین ۵ تا ۱۰۰) باشد، اما باید تا آنجاکه می‌توان، معرف جامعه هدف باشد. پیش‌آزمایی (مطالعه راهنمایا مقدماتی) می‌تواند از طریق گروه کانونی، مصاحبه شناختی، مصاحبه، پرسشگری^۱ گروهی یا فردی انجام پذیرد.

¹ debriefing

² thumb rule

³ Kaiser-Meyer-Olkin (KMO)

حجم نمونه آسیب‌گر نیست [۳۷، ۳۸].

وارسی کیفیت داده‌ها

برای رسیدن به یافته‌های دقیق و تکرارپذیر، داده را باید پالایش کرد. پژوهشگران باید داده‌ها را به لحاظ داده‌های بدست نیامده، داده‌های دور افتاده و هم خطی چندگانه بطور جدی مورد ریز بینی و تیز بینی قرار دهند [۳۹]. منابع داده‌های به دست نیامده مختلف است؛ گاهی پاسخگویان برخی از گویه‌ها را بی پاسخ رها می‌کنند و برخی از پژوهشگران آنها را از تحلیل کنار می‌گذارند. یک راه حل بهتر این است که مقادیر آنها بر اساس روشهای آماری برآورد شود [۴۰].

داده‌های دور افتاده چنان که چندان زیاد نباشند، حذف یا به طور توجیه‌پذیری تبدیل می‌شوند، چنانچه ۵۰ درصد یا بیشتر از ۵۰ درصد گویه‌های بی پاسخ باشند، کنار گذاشته می‌شوند یا با توجه به الگوی از دست رفتگی آنها را باید مدیریت کرد. به طور کلی روش برآورد توصیه می‌گردد.

وارسی همبستگی میان - گویه‌ای و گویه - نمره کل

یکی از روشهای آماری برای نگهداشت یا کنار گذاری یک گویه یا تعدیل آن، برآورد همبستگی میان گویه‌ای^۱ و گویه-نمره کل^۲ است. همبستگی میان گویه‌ای (که همبستگی چند خانه‌ای و چهار خانه‌ای برای گویه‌های مقوله‌ای نیز نامیده می‌شود)، میزان همبستگی بین نمره‌های یک گویه با نمره‌های گویه‌های دیگر مقیاس را نشان می‌دهد، این همبستگی نشان می‌دهد که گویه‌های یک مقیاس به چه میزانی محتوای یکسانی را اندازه می‌گیرند. گویه‌هایی با همبستگی پایین (کمتر از ۰/۳) نیاز به بازنگری جدی دارند یا از مقیاس نهایی کنار گذاشته می‌شوند [۴۱، ۴۲]. همبستگی گویه - نمره کل (که از آن به دو رشته‌ای نقطه‌ای نیز نام برده می‌شود) برای بررسی رابطه بین هر گویه با نمره کل همه گویه‌های آن مقیاس، به کار گرفته می‌شود. بهتر است از همبستگی تعدیل شده ی گویه - نمره کل^۳ استفاده شود که درنمره کل آن، نمره آن گویه وارد نمی‌شود. چنانچه مقدار آن بسیار پایین باشد (کمتر از ۰/۳)، کنار گذاری آن گویه اولویت می‌یابد. این همبستگی‌ها را می‌توان با نرم افزارهای SPSS، SAS، R، Mplus یا Stata محاسبه کرد.

تحلیل عاملی و اکتشافی و تأییدی

تحلیل عاملی مرحله‌ای است که در آن تعداد بهینه عوامل که گاهی حوزه نام می‌گیرند تعیین می‌شود، تعداد عواملی که با گرد آیه‌ای از گویه‌ها برازندگی یابند.

بررسی گسترده نظری پژوهشهای مقیاس سازی از این نکته پرده بر می‌دارد که تحلیل عاملی اکتشافی^۴ (EFA) بیشترین رویکرد در ارزیابی مقیاس روان‌شناختی است. تحلیل عاملی دسته‌ای از روشهای آماری برای تحلیل ساختاری است که در تعیین همبستگی میان متغیرهای مشاهده شده به کار می‌رود تا داده‌ها به ساختارهای زیربنایی فرو کاسته شوند. این ساختار زیربنایی سازه یا عامل نام می‌گیرند [۴۳] اساساً، EFA داده‌ها را می‌کاود و تعداد عوامل را نشان می‌دهد. در تحلیل عاملی تأییدی^۵ (CFA)، پژوهشگران تعداد عوامل را مشخص می‌کنند و متغیرهایی که با هر عامل ربط دارند را نیز پیش از تحلیل تعیین می‌کنند. EFA بر CFA در مقیاس سازی اولویت دارد؛ زیرا این احتمال وجود دارد که مفروضه‌های پژوهشگر (آن) درمورد ابعاد سازه مورد بررسی و نیز کیفیت گویه‌های مربوط به آنها نادرست باشد.

باید توجه داشت که CFA باید در نمونه‌ای جدا از نمونه‌ای که در آن EFA اجرا شده است به کار برده شود تا ساختار مفروض حاصل از EFA تأیید شود [۴۴، ۴۵، ۴۶، ۴۷، ۴۸، ۴۹، ۵۰].

دو راه حل بهینه نیز در این بخش قابل اجرا است که آن را در این مقاله راهبرد چهار گامی^۶ و دو گامی^۷ می‌نامیم.

راهبرد چهار گامی در واری ساختار عاملی یک مقیاس

از این روش زمانی استفاده می‌شود که مقیاس را پژوهشگر برای اولین بار خود طراحی کرده باشد. چنانچه حجم نمونه کافی باشد، می‌توان به طور تصادفی نمونه موجود را به دو بخش تقسیم کرد.

سپس باید معادل بودن دو گروه را از نظر همگنی جمعیت شناختی (از نظر سن، جنس و تحصیلات و ...) آزمون کرد (که همه ی p ها نباید معنادار باشد $p > 0/05$).

گام یکم بر روی نیمه نخست، تحلیل عاملی اکتشافی (EFA) انجام می‌گیرد.

(در مورد روش استخراج، روش چرخش و روش نگهداشت

¹ inter-item

² item-total correlation

³ corrected item total correlation

⁴ exploratory factor analysis

⁵ Confirmatory factor analysis

⁶ 4-step strategy

⁷ 2-step strategy

متعامد و مقدار ویژه بیشتر از ۱ است. به خاطر داشته باشید این موارد همیشه درست کار نمی‌کند.

در تحلیل عاملی اکتشافی (EFA) باید اطمینان حاصل شود که آیا داده‌ها عامل پذیرند یا خیر؟ بر اساس ماتریس همبستگی، آزمون کرویت بارتلت^۴ و ملاک KMO (کیزر - میر - الکین) اطلاعات مناسبی در مورد عامل پذیری داده‌ها فراهم می‌آورند. ماتریس همبستگی باید دارای مقادیری دست کم با ارزش ۰/۳ باشد و کای - اسکوئر بارتلت در سطح ۰/۰۵ یا کمتر باید معنادار شود و مقدار KMO باید بیشتر از ۰/۶ باشد تا بتوان به نتیجه تحلیل عاملی اکتشافی تکیه کرد [۳۸، ۵۱، ۵۲].

باید به خاطر داشت تحلیل عامل اکتشافی روشی یگانه نیست، دسته‌ای از فنون استخراج و چرخش با هدف دستیابی به سازه‌های پنهان تحلیل عاملی اشتراکی یا تحلیل عاملی اکتشافی در دسترس اند.

در تحلیل عاملی اکتشافی هدف دستیابی به واریانس مشترک میان متغیرها است. یک اشتباه رایج پژوهشگران هنگام کاربست تحلیل مولفه‌های اصلی (PCA) روی می‌دهد [۴۰، ۵۳]. PCA به لحاظ مفهومی و ریاضی از تحلیل عاملی اشتراکی متمایز است. تحلیل عاملی اشتراکی بر تبیین مقدار واریانس متمرکز نیست، بلکه این روش بر واریانس مشترک تکیه دارد، اشکال PCA این است که اگر در مقیاس سازی برای سازه پنهان به کار گرفته شود اندازه مولفه‌ها به دلیل در بر گرفتن واریانس خطا فربه می‌شود. بنابراین، پژوهشگران تعداد بسیار زیادی مولفه را در مقیاس نگه می‌دارند [۴۰، ۴۴، ۴۹، ۵۴، ۵۵].

اگر که ساختار عاملی در قالب تحلیل عاملی تأییدی بررسی شود نتیجه تحلیل عاملی اکتشافی تعمیم پذیرتر است. توصیه می‌گردد که از تحلیل عاملی اشتراکی (اکتشافی) به جای تحلیل مولفه‌های اصلی استفاده شود این تأکید در پژوهشهای روان‌شناختی غایب است.

نکته مهم دیگر به روش استخراج در تحلیل عاملی اشتراکی (اکتشافی) بر می‌گردد. اساساً در همه این روشها، تلاش برای تبدیل فضای متغیری به فضای عاملی است. این روش‌ها در پژوهشهایی مورد مقایسه و بررسی قرار گرفته‌اند. اسبرن (۲۰۱۴) بر این باور است که اطلاعات کمی در مورد نارسایی‌ها و مزایای این روشها در دست است [۵۶].

یا کنارگذاری عوامل در ادامه سخن به میان خواهد آمد).

گام دوم بر روی نیمه دیگر نمونه نیز تحلیل عاملی اکتشافی (EFA) با همان روند گام یکم اجرا می‌شود. انتظار این است که ساختار عاملی اکتشافی در دو گام بسیار مشابه یا یکسان باشد وگرنه، ساختار عاملی مانایی کافی ندارد.

گام سوم در این گام، دو نیمه از نمونه را ترکیب کرده و تحلیل عاملی اکتشافی بر روی نمونه کل (ترکیب گام اول و دوم) انجام می‌شود و نتیجه این تحلیل با جزئیات در مقاله گزارش می‌گردد. اما نتیجه گام یکم و دوم فقط در یک سطر (شامل KMO، درصد واریانس تبیین شده و ارزش‌های ویژه) آورده می‌شود.

گام آخر، نتیجه گام سوم را در نمونه‌ای جدا در قالب تحلیل عاملی تأییدی به کار برده می‌شود

مقاله رشیدی، فراهانی و روشن (۲۰۲۰) نمونه‌ای از این راهبرد در ساخت مقیاس است [۸۸].

راهبرد دوگامی در واریسی ساختار عاملی یک مقیاس

از این روش زمانی می‌توان بهره‌بهینه برد که مقیاس به زبان دیگری موجود باشد یا بر اساس دلایل منطقی و نظری یا تجارب در حوزه پژوهشگر نشان دهد که ساختار عاملی مفروض در مقیاس از مانایی کامل یا نزدیک به کامل برخوردار است.

در این روش نمونه موجود به طور تصادفی به دو نیمه تقسیم می‌شوند.

گام یکم در نیمه اول نمونه تحلیل عاملی اکتشافی انجام می‌گیرد و نتیجه گزارش می‌گردد.

گام دوم در نیمه دوم نمونه تحلیل عاملی تأییدی انجام می‌شود و عوامل بدست آمده در گام یکم در این گام راستی آزمایی تجربی می‌گردند.

یادآوری می‌شود اجرای تحلیل عاملی اکتشافی و تأییدی بر روی یک نمونه واحد یک خطای روش شناختی غیر قابل چشم‌پوشی و مهملک در فرایند مقیاس‌سازی است.

نکات و چالش‌های تحلیل عاملی

به طور کلی در تحلیل عاملی، مهمترین تصمیم در مقیاس‌سازی، تصمیم‌گیری بین تحلیل عاملی اشتراکی^۱ و تحلیل مولفه‌های اصلی^۲ (PCA) است. تعداد عواملی که باید حفظ شوند و روش چرخش (متمایل در مقابل متعامد^۳) است. پیش فرض برخی نرم‌افزارهای آماری، PCA، چرخش

¹ common factor analysis

² principal component analysis

³ oblique vs orthogonal

⁴ Bartlett's test of sphericity

اجرای این روش از طریق <https://analytics.gonzaga.edu/parallellengine/> در خط در اینترنت قابل انجام است.

درصد واریانس تبیین شده

در این روش عاملی قابل قبول است که دست کم ۱۰ درصد از واریانس متغیرهای درون خود را تبیین کند (تاباکنیک و فیدل، ۲۰۱۳). استدلال این درصد را در ادامه درخواهید یافت.

نگارندگان این مقاله بر این باورند که تحلیل موازی در مقیاس‌های علوم رفتاری و روانشناسی به عامل‌های توجیه پذیرتر به لحاظ نظری و تعمیم پذیرتر از نظر تجربی می‌انجامد. در بسیاری از مواقع پژوهشگر با مقدار ویژه ۱/۰۱ یا ۹۹/۰۱ یا نقطه‌ای که در تلاقی خط برش در نمودار سنگریزه واقع شده، روبرو می‌شود که این در مرز قرار گرفتن، تصمیم پژوهشگر در عامل دانستن یا ندانستن آن عامل را در محاق تردیدی فرساینده فرو می‌برد؛ اینجاست که استفاده از تحلیل موازی در تعیین تعداد عوامل موضوعیت و اهمیت چندگانه می‌یابد.

چرخش عوامل در EFA

یکی از محدودیت‌های اصلی تحلیل عامی اکتشافی این است که پاسخ مشخصی ندارد، بنابراین برای رسیدن به عوامل کاملاً مشخص چرخش عوامل ضروری است. هدف چرخش رسیدن به ساختار ساده است (فراهانی و عریضی، ۱۳۸۸)، دو نوع روش چرخش متعامد^۶ و متمایل^۷ برای پژوهشگران وجود دارد. بیشتر پژوهشگران از چرخش واریمکس که یک چرخش متعامد است بهره می‌گیرند که در آن عوامل نامرتب از هم تعیین می‌گردند [۵۹، ۶۷، ۶۸، ۶۹].

مشکل چرخش متعامد این است که دور از انتظار است که عوامل در علوم رفتاری و روان‌شناختی نامرتب با هم باشند [۳، ۵۱].

در چرخش واریمکس که خواستنی‌ترین نوع چرخش از نوع متعامد است، بارهای بالا عاملی (رابطه هر گویه با عامل زیر بنایی) بالاتر و تعداد عوامل کمتر می‌شوند، زیرا امکان مرتبط شدن عوامل با هم در این روش چرخش وجود ندارد [۵۷]. گریز از یافتن یک عامل عمومی در زمانی که یک عامل وجود دارد و ایجاد بارهای عاملی متقابل از نتایج کاربست این

دو روش استخراج؛ عامل دهی محور اصلی^۱ (PAF) و یا بیشینه درست‌نمایی^۲ (ML) رایج‌ترین روش‌ها هستند. PAF برای نمونه‌های کوچک و غیر نرمال و ML برای زمانی که داده‌ها از توزیع نرمال برخوردارند توصیه می‌گردد [۵۷، ۵۸].

روش‌های تعیین تعداد عوامل

راه تعیین تعداد دقیق عوامل (خرده مقیاس‌ها) که باید در مقیاس باقی‌بمانند به ندرت روشن است. به طور کلی چهار روش پیشنهاد می‌شود:

مقدار ویژه بیشتر از یک

شاید بتوان رایج‌ترین روش تعیین نگهداشت یا کنارداشت یک عامل را این روش دانست: قاعده ارزش‌های ویژه^۳ بیشتر از یک [۶۰، ۶۱، ۶۲].

ارزش ویژه نشان‌دهنده مقدار واریانس است که یک عامل تبیین می‌کند. هر چه ارزش ویژه اش بیش، درصد واریانس تبیین‌گرس بیشتر.

اکنون روش شناسان، بر روی عدد یک به عنوان نقطه‌ی سواسازی تعداد عوامل هم‌داستان نیستند، چرا که تعداد عوامل دارای مقدار ویژه‌ی بالای یک عمدتاً با تعداد کل متغیرهای موجود در مدل پیوند دارند که سبب استخراج تعداد بسیار زیاد یا کم عوامل می‌شوند [۴۷، ۴۰، ۴۷، ۵۰]. بنابراین این روش، در مقیاس‌سازی توصیه نمی‌شود.

نمودار سنگریزه

نمودار سنگریزه^۴، روش دیگری برای تبیین تعداد عوامل است و روشی دیداری از ارزش‌های ویژه است. خط برش برای تعداد عوامل جایی است که یک خط از حالت خط مستقیم شکسته می‌شود. این روش نسبت به قاعده ارزش ویژه روش درست‌تری پنداشته می‌شود [۴۸، ۴۹، ۵۱، ۵۳].

تحلیل موازی

تحلیل موازی^۵ در سال ۱۹۶۵ توسط هرن ابداع شد. در این روش ارزش‌های ویژه حاصل از داده‌های واقعی با ارزش‌های ویژه حاصل از داده‌های تصادفی شده مقایسه می‌شود. این روش برای تعیین تعداد عوامل توصیه‌بیشتری شده است [۴۷، ۴۳، ۶۴، ۶۵، ۶۶]. در این روش عواملی قابل قبول پنداشته می‌شوند که ارزش ویژه آنها بیشتر از ارزش ویژه حاصل از داده‌های تصادفی است.

¹ principal axis factoring

² maximum Likelihood

³ eigen values

⁴ scree plot

⁵ parallel analysis

⁶ orthogenal

⁷ oblique

روش است.

اگر عوامل در اساس همبسته هستند راه حل متمایل به پژوهشگران امکان بیشتری می‌دهد که به ساختار نظری مدل مورد بررسی نزدیک گردند [۵۷، ۵۹، ۷۰]

اگر عوامل همبسته شوند، پژوهشگران باید بررسی کنند که آیا آن عوامل دو سازه جداگانه را اندازه می‌گیرند یا یک عامل؟ بنابراین توصیه می‌گردد که از چرخش متمایل استفاده شود، زیرا در روان‌شناسی بیشتر مدل‌ها را درست‌تر بازنمایی می‌کند؛ چرا که این چرخش امکان همبسته شدن عوامل را برای پژوهشگران فراهم می‌آورد. در خروجی نرم افزارهای حاصل از این چرخش دو ماتریس بدست می‌آید، اما می‌توان ماتریس الگو^۱ را برای بررسی ساختار ساده و مهم مورد توجه قرار داد.

چرخش متمایل شامل دو روش رایج ابلیمین مستقیم^۲ و پرومکس^۳ است. در این دو روش امکان همبستگی عوامل وجود دارد. اما روش پرومکس با راه حل متعامد آغاز می‌شود، آنگاه به یک راه حل متمایل تبدیل شده و پایان می‌گیرد [۷۱]. پژوهشها نشان می‌دهند روش پرومکس مقاوم تر^۴ است و بیشتر توصیه می‌گردد.

تعیین ملاک‌های نگهداشت یا کنارگذاری گویه‌ها

در مقاله‌های مقیاس‌سازی عموماً پژوهشگران ملاک‌هایی که بر آن اساس یک گویه را در فرم نهایی نگه می‌دارند مشخص نکرده یا نارسا یا ناروشن بیان کرده‌اند.

ملاک‌های نگهداشت یک گویه

(۱) بار عاملی گویه در گستره ۰/۳ - ۰/۵ یا بالاتر
(۲) نبود بارهای متقابل (یعنی بارگیری معنادار یک گویه در بیش از یک عامل)

(۳) نبود یک عامل با کم‌تر از سه گویه

(۴) نبودارتباط جدی بین گویه‌ای^۵

(۵) همگرایی نظری

بار عاملی یک گویه ممکن است مثبت یا منفی باشد. میزان کمتر از ۰/۳ غیر قابل پذیرش است [۵۲، ۶۲، ۷۲]، بنابراین مقادیر ۰/۳۲ [۳۸، ۵۰]، ۰/۳۵ [۵]، ۰/۴ [۴۵]، ۰/۴ [۷۴] و ۰/۵ [۷۵] مورد پذیرش‌اند.

توصیه می‌گردد که برای هر خرده‌مقیاس، دست‌کم سه گویه به منظور برون‌آوری مرکز اصلی هر بعد لازم است. باید توجه کرد که خرده‌مقیاس‌های دو گویه‌ای یا تک-گویه‌ای فقط زمانی توصیه می‌شود که گویه‌ها به شدت همبستگی قوی داشته باشند ($r > 0.7$) [۵۰]

تحلیل عاملی تأییدی

تحلیل عاملی تأییدی^۶ (CFA) نوعی واریسی نجه‌ای است که امکان مقایسه نظامند یک ساختار جایگزین عاملی اولیه^۷ بر اساس ارزیابی برازندگی را فراهم می‌آورد و روابط میان سازه‌های پنهان که به دلیل خطای اندازه‌گیری همبسته شده‌اند را برآورد می‌کند.

CFA برای آزمودن گسترده‌ای استفاده می‌شود که در آن یک الگوی نظری پیشین از بارهای عاملی سازه‌های معینی، داده‌های واقعی را بازنمایی می‌کنند. به دیگر سخن، CFA به پژوهشگران این امکان را می‌دهد که مدل نظری مورد درک خود را تأیید یا رد کنند و این مهمترین روش بررسی همخوانی ساختار نظری آزمون با ساخت تجربی (دادگانی) آن است.

روایی همگرا، اعتبار سازه و روایی تفکیکی

گویه‌های یک سازه خاص باید همگرا باشند یا نسبت بالایی از واریانس مشترک با هم داشته باشند، این ویژگی مقیاس روایی همگرا^۸ نام گرفته که شکلی از روایی سازه است، در واقع، همگرایی هر گویه میزان اشتراک^۹ آن گویه است که با مجذور بار عاملی آن برابر است [۳۷].

متوسط واریانس استخراج شده^{۱۰} (AVE) بر اساس نتایج CFA یک شاخص مناسب برای روایی همگرا است که بر اساس فرمول زیر بدست می‌آید:

$$AVE = \frac{\sum_{i=1}^n L_i^2}{n}$$

در این فرمول L_i بار عاملی استاندارد هر گویه است و i

¹ pattern matrix

² direct oblimin

³ promax

⁴ robust

⁵ inter-item correlation

⁶ confirmatory factor analysis(CFA)

⁷ Priori factor

⁸ convergent validity

⁹ communalities

¹⁰ Average variance extracted

خطاهای تقریب $(RMSEA) \leq 0/06$ شاخص تاکر -
لوپس $(TLI) \geq 0/05$ ، شاخص برازندگی تطبیقی
 $(CFI) \geq 0/95$ ریشه میانگین مجذور خطاهای
استاندارد شده $(SRMR \leq 0/08)$ و ریشه میانگین
مجذور خطاهای وزنی $WRMR \leq 0/1$ است [۷۶، ۷۷، ۷۸].

روایی افزایشی

یکی از شواهد روایی یک مقیاس که به کلی در مقالات داخلی غایب است، روایی افزایشی^۴ است. هانسل (۲۰۰۳) روایی افزایشی را میزان توانی می‌داند که یک ابزار به پیش بینی کنندگی یک ملاک، فراتر از آنچه بادیگر داده‌ها پیش بینی می‌شده است، می‌افزاید [۷۹]. به دیگر سخن، این روایی پاسخ به این پرسش مهم و فربه است که این ابزار چه دارد که دیگر ابزارهای مشابه از آن بی بهره اند؟ چه لزومی به ساخت یا رواسازی چنین ابزار نویی است؟

این شکل از روایی زمانی که یک آزمون نو ساخته می‌شود یا ابزار موجودی ویراست یا روزآمد می‌شود اهمیت می‌یابد [۸۰]، و اگر یک ابزار در حالی ساخته می‌شود که چندین ابزار از پیش ساخته شده در آن حوزه در دست است، اهمیت بیشتری می‌یابد. این نوع از روایی است که توجیه می‌کند که یک ابزار جدید اطلاعاتی نویی فراهم می‌آورد که آن اطلاعات در ابزارهای در دست، نارسا بوده یا از ریشه نادیده انگاشته می‌شده است.

برای دستیابی به این نوع روایی باید متغیر ملاک و پیش بین (ابزارهای مشابه) را انتخاب کرد و از رگرسیون چند گانه گام به گام^۵ برای تعیین سهم ابزار نو بهره گرفت. متغیر ملاک باید یک پدیده مهم باشد.

متغیرهای پیش بین باید ابزارهای استاندارد و روا و معتبری باشند که به طور رایج به کار بسته می‌شود [۸۰] برای دستیابی به روایی افزایشی از رگرسیون گام به گام استفاده می‌شود و میزان ΔR^2 نقش ابزار نو یعنی همان روایی افزایشی را نشان می‌دهد.

اگر X_0 ابزار در دست و رایج X_n ابزار جدید و y متغیر ملاک باشد آن گاه:

شماره گوپه است، مقدار AVE بیشتر از ۰/۵ نشان از روایی همگرا مطلوب یک مقیاس دارد. AVE را می‌توان برای هر خرده مقیاس نیز جداگانه محاسبه کرد. یکی دیگر از شواهد روایی همگرا یک مقیاس/اعتبار ترکیبی^۱ یا/اعتبار سازه/ی^۲ (CR) آن است. این شاخص از فرمول زیر محاسبه می‌شود:

$$CR = \frac{\sum_{i=1}^n L_i^2}{\sum_{i=1}^n L_i^2 + \sum_{i=1}^n e_i^2}$$

اگر مقدار CR از ۰/۷ بیشتر باشد نشان دهنده روایی همگرا سازه است؛ بدین معنا که گوپه‌ها همگی بطور یکنواختی یک سازه پنهان را بازنمایی می‌کنند. در یک تفسیر دقیق تر باید گفت افزون بر ملاکهای بالا برای CR و AVE باید در همه عوامل $AVE < CR$ باشد تا بتوان از روایی همگرا اطمینان داشت. باید دو چندان بخاطر داشت که روایی همزمان متفاوت از روایی همگرا است.

روایی تفکیکی^۳ نشان می‌دهد که به چه میزان یک سازه بطور کامل از دیگر سازه‌ها جدا است. روایی تفکیکی بالا شاهدهی است بر این نکته که یک سازه یگانه است و پدیده‌هایی را در بر می‌گیرد که دیگر ابزارها نمی‌گیرند [۳۷].

یک رویکرد دقیق برای بررسی این نوع روایی این است که میزان AVE هر عامل با مجذور همبستگی بین آن عامل و عامل مرتبط (همبستگی بین سازه‌ای) مقایسه شود. اگر مقدار AVE هر عامل بیشتر باشد می‌توان به این نتیجه رسید که مشکلی برای روایی تفکیکی مقیاس وجود ندارد.

مروری بر پژوهشهای مقیاس سازی در مجلات داخلی و خارجی‌های کم مایه نشان می‌دهد این دو نوع روایی با این رویکرد از دایره توجه پژوهشگران به کلی برون فتاده‌اند.

برای بررسی برازندگی نظامند ساختار عاملی یک مقیاس لازم است که آماره‌هایی که بعدپذیری را بررسی می‌کنند به سطح کمینه‌ای برسند. این آماره‌ها عبارتند از آزمون کای-اسکور برازندگی دقیق که نباید معنادار باشد ($P > 0/05$) چون این شاخص یک شاخص حساس به حجم نمونه است بنابراین از کای-دو نسبی (کای-دو/درجه آزادی) استفاده می‌شود که مقدار بین ۳ تا ۵ آن مطلوب است، ریشه میانگین مجذور

¹ composite reliability

² construct reliability

³ discriminant validity

⁴ incremental validity

⁵ stepwise multiple regression

انداز این مقاله نیست. مقاله ای با همین منظور توسط نگارندگان این مقاله در حال تدوین است که منتشر خواهد شد.

همچنین مانایی اول و دوم را مانایی ضعیف و مانایی سوم و چهارم را مانایی قوی می نامند.

در تمامی گام ها اگر $\Delta RMSE, \Delta RMR, \Delta CFI, \Delta CFI, \Delta NFI$

دارای اختلاف بیشتر از ۰/۰۱ و یا X^2 معنادار شود ($p < ۰/۰۵$)، نشان از نبود مانایی اندازه گیری دارد [۸۱].

بررسی اعتبار مقیاس

اعتبار ۳ یک مقیاس، میزان همسانی درونی یک ابزار را در شرایطی که آن ابزار دوباره در شرایطی یکسان استفاده شود، نشان می‌دهد [۸۴]. روشهای مختلفی برای تعیین اعتبار یک مقیاس وجود دارد، اما از میان آنها آلفای کرونباخ و باز آزمایی تفوق کاربردی در سنجش اعتبار مقیاس ها دارد [۸۵].

ضریب اعتبار خطای اندازه گیری را نیز نشان می دهد. از آنجا که ضریب آلفای کرونباخ همسانی درونی گویند را نشان می دهد و به طور ساده متوسط ضرایب همبستگی بین گویند ها است، هنگام کاربرد آن باید توجه داشت که اگر مقیاس دارای خرده مقیاس است برای هر خرده مقیاس آلفای جداگانه ای نیاز است و ناهمگنی سبب کاهش آلفا می شود [۸۴، ۸۵]. در نتیجه، یک آلفای کلی همیشه خواستنی و دفاع پذیر نیست. از دیگر سو، اگر همبستگی میان بیشتر گویند ها منفی است این نکته سبب منفی شدن شگفت آور میزان آلفا می شود.

برای حل این مشکل باید همه گویند ها همراستا شوند؛ یعنی نمره بالا در همه آنها مثبت باشد یا منفی و اگر با کاربری چنین شرایطی، دگر باره میزان منفی آلفای کرونباخ رخ نمایاند دو راه حل پیشنهاد می شود:

الف) اگر مقیاس تک عاملی است می توان از ضریب اعتبار تنها استفاده کرد.

$$\theta = \frac{K}{K-1} \left[1 - \frac{1}{\lambda} \right]$$

K: تعداد گویند ها

λ : بزرگترین ارزش ویژه مربوط به عامل کلی مقیاس

ب) استفاده از ضریب امگا که از طریق فرمول زیر محاسبه

$$y_1 = bx_0 + a \quad R_1^2 \quad (۱) \text{ گام اول}$$

$$y_1 = bx_0 + bx_n + a \quad R_2^2 \quad (۲) \text{ گام دوم}$$

$$\Delta R^2 = R_2^2 - R_1^2 \quad (۳) \text{ روایی افزایش}$$

ΔR^2 نشانگر نسبت واریانس یگانه y_1 است که توسط X_n (ابزار نو) تبیین می شود. این نسبت واریانس بایستی بالاتر و فراتر از نسبتی باشد که توسط X_n (ابزار موجود) تبیین می گردد. این ΔR^2 توسط آزمون F آزمون پذیر است (دوس، ۱۹۹۹) لازم به یادآوری است که R^2 تعدیل شده نسبت به R^2 مقاوم تر است و باید از آن در تعیین ΔR^2 استفاده کرد زیرا R^2 تعدیل شده نسبت به حجم نمونه و تعداد متغیرهای پیش بینی تعدیل شده است.

به عنوان نمونه در رواسازی پرسشنامه سواد سلامت جنسی در مقاله رشیدی، فراهانی و روشن (۲۰۲۰) رضایت زناشویی به عنوان متغیر ملاک واز پرسشنامه سواد سلامت و مقیاس دانش جنسی به عنوان متغیرهای پیش بین استفاده شد تا روایی افزایش این ابزار تعیین شود.

روایی ملاک - محور^۱

در این نوع روایی انتظار بر این است که ابزار نو ساخته شده با ابزارهای مشابه همبستگی مثبت (مستقیم) و با ابزارهای وارون همبستگی منفی نشان دهد.

مانایی اندازه گیری

مانایی اندازه گیری^۲ به میزان انتقال پذیری و تعمیم یابندگی ویژگی های روان سنجی یک ابزار در گروههای مختلف یا در طول زمان اشاره دارد. به منظور بررسی مانایی اندازه گیری از ثبات چند گروهی می توان بهره برد.

هدف واری مانایی اندازه گیری این است که به چه میزان ساختار عاملی یک ابزار در گروههای مختلف یا در خلال زمان مقایسه پذیر است [۸۱، ۸۲، ۸۳].

برای بررسی مانایی اندازه گیری ۴ گام لازم است که در آن گام ها به ترتیب (مانایی پیکره ای) (مانایی سنجه ای)، در گام سوم (مانایی عدد ثابت) و در گام آخر (مانایی محض) بدست خواهد آمد. اجرا و توضیح ریزبینانه این گام ها در چشم

¹ criterion – related validity

² measurement invariance

³ reliability

Communication Methods and Measures
(2018) 12(1):25-44
DOI: 10.1080/19312458.2017.1396583

پذیر است:

2- DeVellis RF. Scale Development: Theory and Application. Los Angeles, CA: Sage Publications (2012)

$$\Omega = 1 - \left(a - \sum_{i=1}^n h_c^2 \right) / (k + 2s)$$

3- De Vellis, R. F. Scale Development: Theory and Applications (2nd ed., Vol. 26). Thousand Oaks, CA: Sage Publications (2003).

K: تعداد گویه‌ها

S: مجموع همبستگی میان گویه‌ها

h: اشتراک گویه i ام

4- Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR and Young SL. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Front. Public Health (2018) 6:149. DOI: 10.3389/fpubh.2018.00149

5- Clarke LA, Watson D. Constructing validity: basic issues in objective scale development. Psychol Assess. (1995) 7:309-19. DOI: 10.1037/1040-3590.7.3.309

از آنجا که ضریب آلفای کرونباخ و به طور کلی خانواده آلفا همسانی درونی گویه‌ها را نشان می‌دهد برای بررسی ثبات زمانی از روش بازآزمایی استفاده می‌شود که از محاسبه ضریب همبستگی بین دو بار اجرا با فاصله زمانی ۲-۴ هفته بدست می‌آید.

کمینه میزان خانواده آلفا ۰/۷ است، این مقدار، آستانه قابل قبول بودن ضریب اعتبار است. اما ضریب ۰/۸ تا ۰/۹۵ بسیار مطلوب است [۶، ۸۶، ۸۷].

نتیجه‌گیری

ساخت مقیاس‌های روان‌شناختی کوششی چالش برانگیز است و مرور مقاله‌های منتشر شده نشان از نارسایی برخی از مولفه‌های مهم یا نبود آنها است. بنابراین چنین مقاله‌ای برای تبیین پیرنگ نظری و ترسیم خطوط راهنما برای کاربری آن مبانی نظری در ساخت و رواسازی مقیاس‌ها نه تنها نیاز، که ضرورتی فوری است، این مقاله در پی آن بود که تصویر و تفسیری روشن پیش چشم پژوهشگران مشتاق برکشانند.

در این مقاله به بررسی ساخت آزمون‌ها که در آن‌ها پرسش‌ها دارای پاسخ درست و نادرست واقعی‌اند؛ مانند آزمون‌های هوش و شناختی پرداخته نشده است، هر چند همپوشانی گسترده‌ای بر آن دو سایه افکنده است و اساس این مقاله بر نظریه کلاسیک آزمون (CTT) بنا نهاده شده است و از درخشش‌ها و درعین حال سایه‌های نظریه پاسخ‌سوال (IRT) به دور مانده است، اگرچه بسیاری از پژوهش‌ها نشان می‌دهند هر دو این مسیر به مقصدی نه دور از هم که در همسایگی یکدیگر می‌انجامند. در این مقاله در مورد روش‌های تعیین نقطه برش در مقیاس‌های بالینی سخنی به میان نیامد زیرا بحث دراز دامن و مفصلی است که این مجال آن مقال را برنمی‌تافت.

منابع

14- Hinkin TR. A review of scale development practices in the study of organizations. J Manag.

1- Miller s. Ten Steps in Scale Development and Reporting: A Guide for Researchers.

- Efficacy Scale. *Res Nurs Health* (1999) 22:399–409. DOI: 10.1002/(SICI)1098-240X(199910)22:53.0.CO;2-4
- 28- Tilden V, Nelson, C, May, B. Use of qualitative methods to enhance content validity. *Nursing Research*, 39(3), 172-175
- 29- Gregory E, Gilbert, E, Prion S. Making Sense of Methods and Measurement: Lawshe's Content Validity Index. *Clinical Simulation in Nursing* (2016) 12, 530-531.
- 30- Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. *Psychol Assess.* (1995) 7:238–47. DOI: 10.1037/1040-3590.7.3.238.
- 31- Clarke LA, Watson D. Constructing validity: basic issues in objective scale development. *Psychol Assess.* (1995) 7:309–19. DOI: 10.1037/1040-3590.7.3.309
- 32- Cooper, R., Potter, W. J., & Dupagne, M. (1994). A status report on methods used in mass communication research. *Journalism Educator*, 48(4), 54–61.
- 33- Drennan, J. Cognitive interviewing: Verbal data in the design and pretesting of questionnaires. *Journal of Advancing Nursing* (2003)42(1), 57–63.
- 34- Lewis, B. R., Templeton, G. F., & Byrd, T. A. A methodology for construct development in MIS research. *European Journal of Information Systems* (2005) 14(4), 388–400.
- 35- Reynolds N., Diamantopoulos A., & Schlegelmilch. Pretesting in questionnaire design: A review of the literature and suggestions for further research. *Journal of Market Research Society* (1993) 35(2), 171–182.
- 36- Comrey AL. Factor-analytic methods of scale development in personality and clinical psychology. *Am Psychol Assoc.* (1988) 56:754-61.
- 37- Hair, J., Jr., Black, W. C., Babin, B. J., & Anderson, R. *EMultivariate data analysis* (7th ed.). Upper Saddle River, NJ: Prentice Hall (2010).
- 38- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th Ed.). Boston, MA: Allyn and Bacon
- 39- Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (1995) 21:967–88. DOI: 10.1016/0149-2063(95)90050-0.
- 15- Hunt SD. *Modern Marketing Theory*. Cincinnati: South-Western Publishing (1991).
- 16- Barzegar B, Azadfallah P, Farahani H, Fathi_Ashstiani, A. Development of the Persian Version of the Integrative Hope Scale: A Review of Psychometric Indices. *Int J Behav Sci.* (2018) 12(1): 1-8
- 17- Loevinger J. Objective tests as instruments of psychological theory. *Psychol Rep.* (1957) 3:635–94. DOI: 10.2466/pr0.1957.3.3.635.
- 18- Clarke LA, Watson D. Constructing validity: basic issues in objective scale development. *Psychol Assess.* (1995) 7:309–19. DOI: 10.1037/1040-3590.7.3.309
- 19- Kline P. *A Handbook of Psychological Testing*. 2nd Edn. London: Routledge; Taylor & Francis Group (1993).
- 20- Schinka JA, Velicer WF, Weiner IR. *Handbook of Psychology, Vol. 2, Research Methods in Psychology*. Hoboken, NJ: John Wiley & Sons, Inc. (2012).
- 21- Krosnick JA, Presser S. Question and questionnaire design. In: Wright JD, Marsden PV, editors. *Handbook of Survey Research*. San Diego, CA: Elsevier (2009), pp. 263–314.
- 22- MacKenzie SB, Podsakoff PM, Podsakoff NP. Construct measurement and validation procedures in MIS and behavioral research: integrating new and existing techniques. *MIS Q.* (2011) 35:293. DOI: 10.2307/23044045
- 23- Osborne, J. W. *Best practices in exploratory factor analysis*. Scotts Valley, CA: CreateSpace Independent Publishing. (2014).
- 24- McPhail SM. *Alternative Validation Strategies: Developing New and Leveraging Existing Validity Evidence*. San Francisco, CA: John Wiley & Sons, Inc (2007).
- 25- Guion R. Content validity - the source of my discontent. *Appl Psychol Meas.* (1977) 1:1–10. DOI: 10.1177/014662167700100103
- 26- Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. *Psychol Assess.* (1995) 7:238–47. DOI: 10.1037/1040-3590.7.3.238.
- 27- Dennis C-L, Faux S. Development and psychometric testing of the Breastfeeding Self-

- development research. A content analysis for recommendations for best practices. *The Counseling Psychologist* (2006) 34(6), 806–838. پ.
- 51- McCrosky, J. C., & Young, T. J. The use and abuse of factor analysis in communication research. *Human Communication Research* (1979)5, 375–82.
- 52- Pett, M. A., Lackey, N. R., & Sullivan, J. L. Making sense of factor analysis. The use of factor analysis for instrument development in health care research. Thousand Oaks, CA: Sage Publications, Inc. (2003).
- 53- Reise SP, Morizot J, Hays RD. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Qual Life Res.* (2007) 16:19–31. DOI: 10.1007/s11136-007-9183-7
- 54- Conway, J. M., & Huffcutt, A. IA review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods* (2003) 6(2), 147–168.
- 55- Snook, S. C., & Gorsuch, R. L. Common factor analysis vs. component analysis. *Psychological Bulletin* (1989)106, 148–154.
- 56- Osborne, J. W. Best practices in exploratory factor analysis. Scotts Valley, CA: CreateSpace Independent Publishing (2014).
- 57- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods* (1999), 4, 272–299.
- 58- Nunnally, J. C., & Bernstein, I. Psychometric theory (3rd Ed.). New York, NY, USA: McGraw Hill (1994).
- 59- Ford, J. K., MacCullum, R. C., & Tait, M. The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology* (1986) 39, 291–314.
- 60- Henson, R. K., & Roberts, J. K. Use of exploratory factor analysis in published research. Common errors and some comment on improved practice. *Education and Psychological Measurement* (2006) 66(3), 393–416.
- (2013). Practical considerations for using exploratory factor analysis in education research. *Practical Assessment, Research & Evaluation*, 18(6), 1–13.
- 40- Gottschall AC, West SG, Enders CK. A Comparison of item-level and scale-level multiple imputation for questionnaire batteries. *Multivar Behav Res.* (2012) 47:1–25. DOI: 10.1080/00273171.2012.640589
- 41- Raykov T. Scale Construction and Development. Lecture Notes. Measurement and Quantitative Methods. East Lansing, MI: Michigan State University (2015).
- 42- Piedmont RL. Inter-item correlations. In *Encyclopedia of Quality of Life and Well-Being Research*. Dordrecht: Springer (2014). p. 3303–4. DOI: 10.1007/978-94-007-0753-5_1493.
- 43- Norris M., & Lecavalier L. Evaluating the use of exploratory factor analysis in developmental disability psychological research. *Journal of Autism Development and Disorders* (2010) 40, 8–20.
- 44- Costello, A. B. & Osborne J. W. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation* (2005) 10(7), 1–9.
- 45- Ford, J. K., MacCullum, R. C., & Tait, M. The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology* (1986) 39, 291–314
- 46- Haig, B. D. Exploratory factor analysis, theory generation, and scientific method. *Multivariate Behavioral Research* (2005), 40(3), 303–32.
- 47- Kline, R. B. Exploratory and confirmatory factor analysis. In Y. Petscher, C. Schatschneider, & D. L. Compton (Eds.), *Applied quantitative analysis education and the social sciences* (pp. 171–207). New York, NY, USA: Routledge (2013).
- 48- Pett, M. A., Lackey, N. R., & Sullivan, J. L. Factor analysis. The use of factor analysis for instrument development in health care research. Thousand Oaks, CA: Sage Publications, Inc. (2005).
- 49- Preacher, K. J., & MacCallum, R. C. Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics* (2003)2(1), 13–43.
- 50- Worthington R. L. & Whittaker, T. A. Scale

- 72- Kachigan, S. K. *Statistical analysis: An interdisciplinary introduction to univariate and multivariate methods*. New York: Radius Press. (1986).
- 73- Meyers, L. S., Gamst, G., & Guarino, A. J. *Applied multivariate research: Design and interpretation*. Thousand Oaks, CA: Sage. (2006).
- 74- Reinard, J. C. *Communication research statistics*. Sage Publications, Thousand Oaks, CA. (2006).
- 75- Mertler, C. A., & Vannatta, R. A. *Advanced and multivariate statistical methods: Practical applications and interpretation*. Los Angeles, CA: Pyrczak Publishin. (2001).
- 76- Bond TG, Fox C. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Erlbaum (2013)
- 77- Brown T. *Confirmatory Factor Analysis for Applied Research*. New York, NY: Guildford Press (2014).
- 78- Morin AJS, Arens AK, Marsh HW. A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Struct Equ Model Multidiscip J.* (2016) 23:116–39. DOI: 10.1080/10705511.2014.961800
- 79- Hunsley J, Meyer, G. The Incremental Validity of Psychological Testing and Assessment: Conceptual, Methodological, and Statistical Issues. *Psychological Assessment* (2003), 15, 4, 446–455.
- 80- Haynes, S. N., & Lench, H. C. Incremental validity of new clinical assessment measures. *Psychological Assessment* (2003). 15, 456–466.
- 81- Byrne B. M., & Stewart S. M. The MACS approach to testing for multi-group invariance of a second-order factor structure: A walk through the process. *Structural Equation Modeling*. (2006) 13, 287–321. <http://dx.doi.org/10.1207/s15328007sem13027>.
- 82- Chen F. F. Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling* (2007) 14, 464–504. <http://dx.doi.org/10.1080/10705510701301834>.
- 83- Tylka, T. L. Evidence for the Body Appreciation Scale's measurement equivalence/invariance between U.S. college women and men. *Body Image* (2013) 10, 415–418. <http://dx.doi.org/10.1016/j.bodyim.2013.02.006>.
- 61- Morrison, J. T. Evaluating factor analysis decisions for scale design in communication research. *Communication Methods and Measures* (2009) 3(4), 195–215.
- 62- Russell, D. W. In search of underlying dimensions: The use (and abuse) of factor analysis and scale revision. *Psychological Assessment* (2002) 12(3), 1629–1646.
- 63- Humphreys, L. G., & Montanelli, R. G., Jr. An investigation of the parallel criterion for determining the number of common factors. *Multivariate Behavioral Research* (1975) 10, 193–205.
- 64- Velicer, W. F., Eaton, C. A., & Fava, J. L. Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin, & E. Helmes (Eds.), *Problems and solutions in human assessment: Honoring Douglas Jackson at seventy* (pp. 41–71). Boston, MA: Kluwer. (2000).
- 65- Watkins, M. W. Determining parallel analysis criteria. *Journal of Modern Applied Statistical Methods*. (2006) 5(2), 344–346.
- 66- Zwick, W. R., & Velicer, W. F. Factors influencing four rules determining the number of components to retain. *Multivariate Behavioral Research* (1982) 17, 253–269.
- 67- Borgotta, E. F., Kercher, K., & Stull, D. E. A cautionary note on use of principal components analysis. *Sociological Methods and Research* (1986) 15, 160–168.
- 68- Conway, J. M., & Huffcutt, A. I. A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods* (2003) 6(2), 147–168.
- 69- Norris, M., & Lecavalier, L. Evaluating the use of exploratory factor analysis in developmental disability psychological research. *Journal of Autism Development and Disorders* (2010) 40, 8–20.
- 70- Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment*, 68, 532–560.
- 71- Hendrickson, A. E., & White, P. O. Promax: A quick method for rotation to oblique simple structure. *British Journal of Mathematical Psychology* (1964) 17, 65–70.

- 84- Porta M. A Dictionary of Epidemiology. New York, NY: Oxford University Press (2008).
- 85- Raykov T, Marcoulides GA. Introduction to Psychometric Theory. New York, NY: Routledge, Taylor & Francis Group (2011).
- 86- Brennan RL. A generalized upper-lower item discrimination index. *Educ Psychol Meas.* (1972) 32:289-303. DOI: 10.1177/001316447203200206.
- 87- Sexton JB, Helmreich RL, Neilands TB, Rowan K, Vella K, Boyden J, et al. The Safety Attitudes Questionnaire: psychometric properties, benchmarking data, and emerging research. *BMC Health Serv Res.* (2006) 6:44. DOI: 10.1186/1472-6963-6-44
- 88- Rashid, Kazhal, Farahani, Hojjaollah and Roshan chesli, Rasol (2020) .Developing and valididating sexual health literacy Scale in Iranian Sample .*American Journal of Sexuality Education* (submitted).