

Narrative Criticism of Positivist Researches of Humanities from the Point of View of Generalizability and Reliability of Measures with the Suggestion of Replacing McDonald's Omega (Case Study: Marketing Researches)

Mohsen Moradi (PhD graduate, marketing management, instructor of international research methodology courses, Islamic Azad University, Central Tehran Branch, Department of marketing, mohsen2528@gmail.com)

Bahram Kheiry (Associate Professor, Islamic Azad University, Central Tehran Branch, Department of marketing, bah.kheiry@iauctb.ac.ir)

Aida Miralmasi (Doctor of Philosophy, Islamic Azad University, Tehran Science and Research Branch, Department of Agricultural Development, aidamiralmasi@gmail.com)

ARTICLE INFO

Article History

Received: 2021/12/11

Accepted: 2022/06/26

Key Words:

Quantitative methodology in the humanities,
Measurement of reliability and generalizability,
Omega McDonald's reliability test,
Cronbach's Alpha coefficient,
Limitations of reliability coefficients,
Common mistakes in quantitative research

ABSTRACT

Scientific research includes philosophical, procedural, and ethical steps, and studies with a positivist philosophical premise, which show the recent meta-method reviews, still include a share of more than 70% of studies in humanities and social sciences. From a comparative point of view, it dictates a quantitative approach to researchers. In this approach, entering the field of measurement and then discussing the identification of metrics and linking them with variables or at a more abstract level with structures is, if not the most important, one of the most prominent cases. The vagueness and conceptual immanence of these structures in these sciences makes it more and more difficult for researchers to fully realize validity and reliability. With a scientific review and in a narrative format, focusing on the reliability aspect and the generalizability aspect of the results, the researcher enumerates the weak but unfortunately very common indicators in Iran research literature. Then, with an evaluative and pathological point of view, it proposes a suitable solution to overcome this procedural structural weakness by analyzing and introducing a suitable and operational alternative for it, i.e. McDonald's Omega Coefficient. It was estimated that the results of the present study will create a stable and satisfying perspective in changing the viewpoint of researchers in related humanities fields.

نقد روایتی پژوهش‌های پوزیتیویستی علوم انسانی از نقطه‌نظر تعمیم‌پذیری و قابلیت اطمینان سنج‌ها با پیشنهاد جایگزینی امگا مک‌دونالد (مورد مطالعه: پژوهش‌های بازاریابی)

محسن مرادی (دانش آموخته دکتری مدیریت بازاریابی و مدرس دوره‌های بین‌المللی روش‌شناسی پژوهش، دانشکده مدیریت، دانشگاه آزاد اسلامی، واحد تهران مرکز، ایران؛ mohsen2528@gmail.com)

بهرام خیری (دکترای تخصصی، دانشیار، گروه مدیریت بازاریابی، دانشگاه آزاد واحد تهران مرکز، تهران؛ ایران؛ bah.kheiry@iauctb.ac.ir)

آیدا میرالماسی (دکتری تخصصی، دکتری مدیریت و توسعه کشاورزی، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران؛ ایران؛ aidamiralmasi@gmail.com)

چکیده

پژوهش‌های علمی شامل گام‌های فلسفی،^۱ رویه‌ای^۲ و اخلاقی^۳ است و مطالعاتی با پیش‌فرض فلسفی پوزیتیویستی که مروره‌های فراروش اخیر را نشان می‌دهند، هنوز سهمی بیش از ۷۰ درصد مطالعات رشته‌های علوم انسانی و اجتماعی را شامل می‌شوند. از نظرگاه قیاس‌گونه^۴ رویکردی کمی را به محققان دیکته می‌کند. در این رویکرد ورود به حوزه اندازه‌گیری و به دنبال آن بحث شناسایی سنج‌ها^۵ و پیوند آن با متغیرها و یا در سطحی انتزاعی‌تر سازه‌ها،^۶ اگر نگوییم مهم‌ترین که یکی از برجسته‌ترین موارد است و ابهام و مکتون^۷ بودن مفهومی این سازه‌ها در این علوم، تحقق کامل اعتبار (روایی)^۸ و قابلیت اطمینان (پایایی)^۹ برای پژوهشگران را هر چه دشوارتر می‌نماید. محقق با مروری علمی و در قالبی روایتی^{۱۰} با تمرکز بر وجه پایایی و جنبه تعمیم‌پذیری^{۱۱} نتایج، سستی و ناتوانی شاخص‌های ضعیف اما متأسفانه بسیار رایج در پژوهش‌های ایران را برمی‌شمارد و با نگاهی ارزیابانه و آسیب‌شناسانه، رسالت ارائه راهکاری مناسب برای برون‌رفت از این ضعف ساختاری رویه‌ای را با واکاوی و معرفی جایگزینی مناسب و عملیاتی برای آن، یعنی ضریب امگای مک‌دونالد^{۱۲} پیشنهاد می‌کند. پیش‌بینی می‌شود که نتایج مطالعه حاضر، چشم‌انداز پایدار و موفق‌تری را در تغییر دیدگاه پژوهشگران علوم انسانی پدید آورد.

اطلاعات مقاله

تاریخ دریافت: ۱۴۰۰/۰۹/۲۰

تاریخ پذیرش: ۱۴۰۱/۰۴/۰۵

واژگان کلیدی:

روش‌شناسی کمی علوم انسانی،
سنجش پایایی و تعمیم‌پذیری،
سنجش پایایی امگای مک‌دونالد،
ضریب آلفای کرونباخ،
محدودیت‌های ضرایب سنجش پایایی،
اشتباهات رایج در پژوهش‌های کمی

1. Philosophical
2. Praxis
3. Ethics
4. Deductive
5. Item
6. construct
7. Latent
8. Validity
9. Reliability
10. Narrative
11. Generalization
12. McDonald's omega

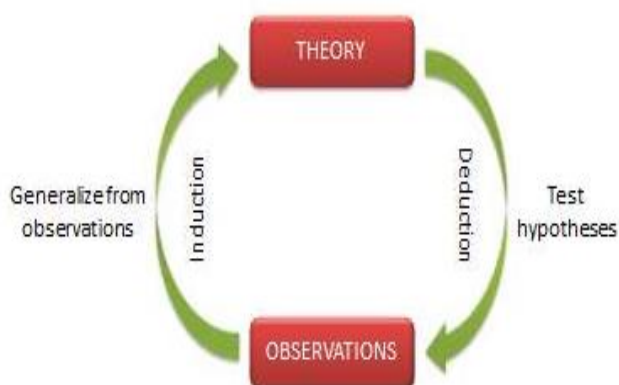
۱. مقدمه

علم عبارت است از پیگیری و به‌کارگیری دانش در درک جهان طبیعی و اجتماعی که از راه یک روش‌شناسی نظام‌مند مبتنی بر شواهد و آزمایش‌ها به دست می‌آید و به نیازهای یک حوزه پژوهشی در حال توسعه پاسخ می‌دهد (Johnson Burke & Christensen, 2014, p. 15)؛ از این رو علوم را می‌توان به دو دسته گسترده تقسیم کرد: الف) علوم طبیعی و ب) علوم اجتماعی و انسانی (Bhattacharjee, 2012, p. 1). صاحب‌نظرانی اعتقاد دارند که علوم طبیعی مانند زیست، فیزیک، شیمی و... مبنای فن‌آوری‌های نوین هستند و بیشتر افراد وقتی با واژه علم روبه‌رو می‌شوند، در گام نخست علوم طبیعی را به یاد می‌آورند (Neuman, 2014, p. 15)؛ در مقابل علوم انسانی و اجتماعی^۱ دسته‌ای از تخصص‌های دانشگاهی و مرتبط با اجتماع و روابط بین افراد آن در جامعه می‌باشند. علوم اجتماعی به‌عنوان یک کل، شاخه‌های فراوانی دارد. علوم اجتماعی شاخه‌ای از مطالعات انسانی است که موجودات از نوع انسان را از جنبه‌های مختلف مورد مطالعه قرار می‌دهد. اکنون، چنانچه از تفاوت‌های علوم انسانی و اجتماعی ورشته‌ها و مباحث بین‌رشته‌ای آنها بگذریم (در مجال این مطالعه نیست)، علوم انسانی و اجتماعی را می‌توان در رشته‌هایی مانند روان‌شناسی (علم رفتارهای انسانی)، جامعه‌شناسی (دانش گروه‌های اجتماعی) و اقتصاد (علوم بنگاه‌ها، بازارها و اقتصاد)، مدیریت و بازاریابی و... طبقه‌بندی کرد (Bhattacharjee, 2012, p. 2).

مقدمه به این گزاره می‌انجامد که هدف علم ایجاد دانش علمی است. دانش علمی به مجموعه قوانین و نظریه‌های کلی برای توضیح یک پدیده یا رفتار موردعلاقه اشاره دارد که با استفاده از روش علمی به دست می‌آید؛ در واقع قوانین^۲ الگوهای مشاهده‌شده از پدیده‌ها یا رفتارها هستند؛ در حالی که نظریه‌ها^۳ توضیحات منظم از پدیده یا رفتارهای اساسی هستند (Smith & Albaum, 2013). روشن‌تر اینکه نظریه‌ها و شواهد (مشاهدات) دو ستون و در واقع تنها دو ستونی هستند که دانش علمی بر پایه آنها بنا شده است. در علم، نظریه‌ها و مشاهدات به هم مرتبط

1. social science
2. laws
3. theories

هستند و بدون یکدیگر وجود ندارند. نظریه‌ها به آنچه ما مشاهده می‌کنیم، معنا و اهمیت می‌بخشند و مشاهدات به اعتبارسنجی و قابلیت اعتماد، اصلاح نظریه موجود یا ساختن نظریه جدید کمک می‌کنند (Moradi & Miralmasi, 2020, p. 9).



شکل ۱. چرخه کامل تحقیق علمی

به هر حال هدف از این مقدمه کلی و اجمالی درک این نکته بسیار مهم است که این دانش شاید ناقص یا حتی کاملاً دور از حقیقت باشد؛ یعنی ما باید درک کنیم که نظریه‌هایی که دانش علمی براساس آنها بنا شده است، تنها توضیحات یک پدیده خاص می‌باشد؛ بنابراین همان‌طور که دانشمندان معتقدند که توضیح و تبیین بسیار خوب تا بسیار ضعیف وجود دارد، پس باید قرارگیری در درجات مختلف این پیوستار و طیف را از این امر منتج کرد که این توضیحات به چه اندازه با واقعیت سازگار است تا بتوان این نظریات را قدرتمند یا ضعیف دانست. البته نباید کتمان کرد که پیشرفت علم و حرکت در طول زمان از نظریه‌های فقیرتر به غنی‌تر با مشاهدات بهتر و استفاده از ابزار دقیق‌تر و استدلال منطقی آگاه‌تر امکان‌پذیر شده است (Johnson Burke & Christensen, 2014, p. 57). پس نمی‌توان کتمان کرد که رسیدن به بلوغ در علم در قالب این چرخه و دستیابی به نظریه‌هایی دقیق که به بهترین صورت ممکن دنیایی واقعی را توضیح و تبیین نماید، نیازمند دنبال کردن و رسیدن به سنجه‌هایی حتی الامکان ایدئال است که هم با اعتبار خود تناسب ایده‌ها با واقعیت را تضمین کند و هم پایا و قابل اعتماد باشد و پیامدهای نامنظم، ناپایدار و ناسازگار نداشته باشد (Aikin & Talisse, 2016).

شد که برخلاف وجود انبوهی از نظریه‌ها، تعداد بسیار اندکی از آنها مورد آزمون قرار گرفته است و این احساس موجب گرایش روزافزون پژوهشگران به سمت رویکرد کمی شد و عملاً کشف تئوری‌های جدید کاهش یافت و در بسیاری از دانشگاه‌ها به بوتۀ فراموشی سپرده شد (Wertz & et al., 2011). اما متأسفانه موضوعی که بیش از بی‌توجهی به جنبۀ ساخت نظریات جدید از راه رویکرد کیفی آشکار شد، این بود که بسیاری از پژوهشگران در بسط و عملیاتی کردن مفاهیم و قضایای تحقیق کیفی و یا آزمون نظریات بسیار ضعیف عمل کردند (Singh, 2019)؛ یعنی اگر از نظر بلیکی گذر کنیم که این پرسش را مطرح می‌کند: «آیا می‌توان از روش‌های علمی با پارادایم‌های اثبات‌گرا در علوم انسانی و اجتماعی استفاده کرد؟»، بزرگان این حوزه معتقدند روش‌های علمی قدیمی مثل آزمون‌های سنجش پایایی از جمله آلفای کرونباخ، روش‌های دونیم‌سازی و فرم‌های موازی^۴ با رشد مباحث تازه‌ای چون سنجه‌های تک‌آیتمی^۵، مدل‌های واریانس‌محور^۶، مدل‌های مرتبه بالاتر^۷ و مفاهیم دارای خرده‌مقیاس‌های گوناگون، سنجه‌های انعکاسی و ترکیبی^۸ (Marko Sarstedt, Bengart, Shaltoni & Lehmann, 2018)، اساساً بازتاب‌دهنده یک روش علمی مناسب برای آزمون اعتبار و قابلیت اطمینان ابزارهای جمع‌آوری اطلاعات برای تأیید و اصلاح فرضیات برآمده از نظریات در یک جمعیت به شمار نمی‌رود (Bielik, 2019).

در سال‌های گذشته محققان در این علوم از رویکردهای بهبودیافته زیادی برای برآورد قابلیت اطمینان حمایت می‌کنند، اما با وجود انتشار و توسعه این گزینه‌ها، مقاومت بسیار زیادی در برابر پیشرفت توسعه و تفسیر، کاربرد و گزارش قابلیت اطمینان مقیاس وجود دارد (Dunn, Baguley & Brunnsden, 2014). در سال ۱۹۹۹ ویلکینسون^۹ بر تأکید گروه ویژه‌ای پی‌ای^{۱۰} در

مسئله پژوهش حاضر نیز آشخور همین مطلع است و محقق برای روایتی علمی‌تر به دلیل گستردگی این مباحث در هر دو رویکرد استدلالی قیاسی و استقرائی که نیومن در سال ۲۰۱۴ بر تفاوت در شیوه‌های ارزیابی آنها اشاره کرده است، تصمیم می‌گیرد مسئله خود را در حوزه اندازه‌گیری و آزمون نظریه‌ها متمرکز کند و بکوشد که با آسیب‌شناسی شیوه‌های رایج ارزیابی پایایی (قابلیت اطمینان) در علوم اجتماعی و با مطالعه پژوهش‌های بازاریابی در ایران که به دلیل بین‌رشته‌ای بودن آن در برگیری کلیه روابط و مناسبات انسانی در سطحی فردی، گروهی و اجتماعی می‌باشد، در نهایت به موردی قابل بسط و تعمیم به سایر مطالعات رشته‌های دیگر علوم انسانی و اجتماعی با پیش‌فرض‌های فلسفی اثبات‌گرایانه بدل شود؛ درحقیقت با این انگاره، انحراف وضع موجود بیشتر پژوهش‌های این حوزه را از وضعیت مطلوبی که تقریباً دو دهه در پژوهش‌های علوم انسانی و اجتماعی دنیا پابرجاست، مسئله را از بعدی تاکتیکی و محتوایی به اثبات رسانیده و بنا بر سنت پیش‌فرض‌های فلسفی انتقادی و رای نقد بر این پژوهش‌ها، راه‌حلی برای پر کردن این خلأ عمیق به پژوهشگران تحقیقات علوم انسانی و اجتماعی پیشنهاد گردد؛ زیرا همچنان که بیان شد مطالعات فراروش نشان می‌دهد با وجود همه نقدها هنوز سهمی بالای ۷۰ درصد مطالعات دانشگاهی در این علوم را پژوهش‌های کمی‌گرا به خود اختصاص داده است.

۲. بیان مسئله پژوهش و ادبیات نظری

از سال ۱۹۶۷ با ارائه کتاب کشف تئوری داده‌بنیاد^۱ تلاش شد میان دو رویکرد اصلی تحقیق، یعنی پژوهش کمی و کیفی توازن ایجاد شود؛ اگرچه مباحثات بسیاری چون نظریه ناسازگاری آنها برای علوم انسانی و اجتماعی از سوی اندیشمندان چون توماس کوهن^۲ ارائه شد (McCain, 2016)، اما از همان آغاز نویسندگانی چون اشتراوس و گلیسر^۳ (۱۹۶۷) اظهار کردند که اثبات صحت نظریه‌ها به موضوع اصلی علوم انسانی و اجتماعی تبدیل شده است؛ یعنی در بیش از ۴۴ سال پیش این‌گونه احساس

1. Grounded theory
2. Kuhn, Thomas Samuel
3. Glaser and Strauss

4. Parallel forms

5. Single item

6. Variance-based models

7. Higher order models

8. Reflective and Formative

9. Wilkinson

10. APA Task Force on Statistic Inference

سنجش سازه‌ها را منتج می‌شود، از ضعف‌های این شاخص پایایی سنج مقیاس‌ها بسیار ناچیز و اندک است؛ به گونه‌ای که آیکن^۸ و همکارانش در سال ۲۰۰۸ در مطالعه خود نه تنها عدم درک ضریب آلفای کرونباخ را از جانب محققان مطرح کرده‌اند، بلکه به‌طور کلی ایجاد فضایی برای نقد این شاخص و توسعه روش‌های جایگزین برای پایایی سنجی با روش‌های مبتنی بر سازگاری درونی گویه‌های یک مقیاس را بسیار ناچیز و با خلأهای عمیق مطالعاتی برشمردند (Aiken, West & Millsap, 2008).

محقق نیز اگر براساس تأکید بر تجربه شخصی خود برای ایجاد جرقه‌ای جهت انجام این مطالعه تأکید نماید؛ اظهار می‌دارد که در دوره تحصیلات تکمیلی به ما آموزش داده شد که همیشه آلفای کرونباخ را برای سنجش پایایی شاخص‌های یک مقیاس گزارش کنیم و به راستی نیز محققان در کلیه رشته‌های علوم انسانی و اجتماعی این کار را انجام داده‌اند، اما به بیان گودبوی و مارتین (۲۰۲۰)، نه داوران در کمیسیون‌های بررسی پژوهش‌ها و نه محققان در سنجش پایایی خود در پژوهش‌ها از خود نپرسیدند که آیا آلفا بهترین راه برای برآورد قابلیت اطمینان یک مقیاس است؟ آیا هرگز فکر کردیم که چرا این کار را انجام می‌دهیم؟ آیا صادقانه سال‌های متمادی کورکورانه و بدون آموزش درست از آن استفاده نکردیم؟ (Goodboy & Martin, 2020). آیا وقت آن نرسیده است که آگاهانه عمل کنیم و بدانیم این تکنیک آماری هم مثل بیشتر تکنیک‌ها که در پژوهش‌ها استفاده می‌کنیم ضعیف، سست و حتی نادرست است؟ آیا نباید از جایگزین‌های به‌روز با رشد روش‌های آماری و نرم‌افزارهای پیشرفته بهره ببریم؟ (Moradi & Miralmasi, 2020, p.2).

مسئله اصلی این مقاله بی‌شک پیشنهاد جایگزینی و گسترش روش قابل‌اعتمادتری نسبت به آلفای کرونباخ و همچنین روش‌های مشابه برای سنجش پایایی مقیاس‌های سنجش سازه‌ها در علوم انسانی و اجتماعی با تمرکز بر مطالعات پژوهش‌های بازاریابی به‌عنوان یک علم بین‌رشته‌ای است و محقق در قالب یک مرور ارزیابانه پژوهشی و آسیب‌شناسی این شیء ضعیف در

حوزه آمار استنباطی بر استفاده درست و بهبود مستمر تخمین‌های پایایی صحه می‌گذارد و بیان می‌کند که رایج‌ترین نوع برآورد پایایی گزارش‌شده در مقالات منتشرشده در جهان برآورد سازگاری درونی (در مقابل فرم‌های بازآزمایی^۱ یا موازی^۲) بوده است (Wilkinson, 1999). در حقیقت محبوب‌ترین روش‌ها برای سنجش کیفیت عملکرد یک مقیاس سنجش پایایی و یا قابلیت اطمینان با برآورد سازگاری درونی است. به نقل از هوگان و همکارانش در سال ۲۰۰۰ پیمایش‌ها نشان می‌دهد که برآوردهای سازگاری درونی ۷۵٪ از کل بررسی‌های سنجش پایایی مقیاس‌های گزارش‌شده را تشکیل می‌دهند (Dunn & et al., 2014)؛ همچنین در ادامه این موضوع در سال‌های گذشته در مقالات خارجی مطالعات علوم انسانی و اجتماعی و متأسفانه حتی هم‌اکنون در مطالعات داخلی کشور از جمله مطالعات پژوهش‌های بازاریابی پیمایش شده توسط محقق، متداول‌ترین روش ارزیابی سازگاری درونی برای ارزیابی پایایی، ضریب آلفا^۳ است و به آن آلفای کرونباخ^۴ نیز گفته می‌شود که به مقاله تأثیرگذار کرونباخ در سال ۱۹۵۱ استناد می‌کند و بررسی نویسندگان مقاله حاضر براساس پایگاه اطلاعاتی گوگل اسکالر^۵ در هنگام نگارش این مقاله نشان می‌دهد که بیش از ۵۴۵۰۰ بار به این مقاله معروف استناد شده است و این بررسی رایج بودن این ضریب را تأیید می‌کند (Moradi & Miralmasi, 2020, p.70) اما باید گفت که خود کرونباخ اظهار می‌دارد که: «استنادات^۶ فراوان به مقاله‌ام به‌هیچ‌وجه نشان نمی‌دهد که شخصی که به آن استناد کرده، آن را خوانده است و حتی نشان نمی‌دهد که به آن نگاه کرده است» (Cronbach, 2004, p.392). با توجه به نظر خود کرونباخ می‌توان به این مسئله مهم از نگاه تاکتیکی تأکید کرد که اساساً درک محققان در یک پژوهش در حوزه علوم انسانی و اجتماعی با پیش‌فرض‌های فلسفی پوزیتیویستی^۷ که رویکردهای کمی و بهره‌گیری از مقیاس‌های مختلف برای

1. Test and Retest
2. Test-retest or parallel forms
3. Coefficient Alpha
4. Cronbach
5. Google Scholar
6. citations
7. The positivist paradigm

8. Aiken

۱. نخستین دلیل را محقق در ضعف شدید محققان در مقطع ارشد و دکتری و حتی اساتید راهنما و مشاور آنها در مباحث روش پژوهش مدرن و آمار کاربردی در رشته‌های علوم انسانی و اجتماعی از جمله پژوهش‌های بازاریابی می‌داند (Moradi & Miralmasi, 2020)؛

۲. دلیل پراهمیت دوم، وجود تفاوت بین آگاهی از این ضعف آماری و کسب توانایی علمی لازم برای رفع این نقطه‌ضعف بین محققان است (Dunn & et al., 2014) (به‌ویژه در ایران)؛

۳. مقالاتی که ضریب آلفا کرونباخ را محکوم می‌کنند، بسیار فنی هستند؛ بنابراین، نه تنها برای بسیاری از محققان در حوزه‌های مختلف غیرقابل فهم هستند، بلکه راه‌حل‌های جایگزین آن نیز به همین دلیل مورد استقبال قرار نمی‌گیرد (Aikin & Talisse, 2016)؛

۴. چهارمین دلیل این است که بیشتر روش‌های جایگزین آلفای کرونباخ نیازمند استفاده از تکنیک مدل‌سازی معادلات ساختاری^{۱۲} در سطحی پیشرفته می‌باشد که این امر نیز به دلیل ضعف آماری بیشتر محققان در عمل منتفی است (Aiken & et al., 2008)؛

۵. متأسفانه در کشور ما، ضعف شدید محققان در مقطع ارشد و دکتری و حتی اساتید راهنما و مشاور آنها از نظر کار با نرم‌افزارهای پیشرفته آماری و دست‌نویسی در آنها اثر این ناتوانی‌ها را صدچندان کرده است؛

۶. دلیل ششم هم این است که در بهترین حالت برخی از محققان در سطحی ابتدایی با نرم‌افزارهایی چون اس.پی.اس.اس^{۱۳}، اس.اس^{۱۴} یا استتا^{۱۵} می‌توانند کار کنند که این نرم‌افزارهای مادر آماری نیز قادر به اجرای آزمون‌های پایایی سنجی ضعیفی مانند آلفای کرونباخ هستند (Moradi & Miralmasi, 2020).

با بیان این عوامل اکنون مسئله به سمتی تبیین می‌شود که محقق بتواند گزینه‌ای جایگزین را به‌گونه‌ای ارائه و پیشنهاد کند که پژوهشگران با تخصص اندک هم بتوانند به آسانی آن را اجرا نمایند. صاحب‌نظران مختلف از مدل‌سازی معادلات ساختاری به‌عنوان قوی‌ترین ابزار برای ارزیابی پایایی یک مقیاس حمایت

پایایی سنجی، مطابق با نظر نیومن^۱ (۲۰۱۴) موظف به ارائه و پیشنهاد راه‌حل جایگزین برای نقد است (Moradi & Miralmasi, 2020, p. 72). البته باید ذکر کرد که پیش‌ازین آلفای کرونباخ در صدر تکنیک‌های رایج، در مطالعات زیادی مورد بررسی جدی قرار گرفته است (ر.ک: گرین و هرشبرگر^۲، ۲۰۰۰؛ گرین، لیسیتس و مولایک^۳، ۱۹۷۷؛ گرین و یانگ^۴، ۲۰۰۹؛ هویسامن^۵، ۲۰۰۷؛ رایکوف^۶، ۱۹۹۸؛ سیجتسما^۷، ۲۰۰۹؛ زیمرمن، زامبو و لالوند^۸، ۱۹۹۳؛ لی^۹ و همکاران، ۲۰۰۵ و هایس^{۱۰}، ۲۰۲۰)، در آنها نقاط ضعف بسیار زیاد برای آلفای کرونباخ به‌عنوان رایج‌ترین ابزار پایایی سنجی بیان شده و حتی در برخی از آنها تعدادی از روش‌های جایگزین را معرفی و از آنان برای توسعه و انتشار در بین محققان دفاع کرده‌اند. البته نباید از این نکته غافل شد که ضریب آلفای کرونباخ با وجود تاریخچه طولانی در برابر کنار رفتن و کمرنگ شدن بسیار مقاومت کرده و حتی محبوبیت آن به حدی است که از مقاله کشف دی.ان.ای^{۱۱} در پزشکی بیشتر بدان استناد شده است (Sijtsma, 2009).

اما در تبیین مسئله این پژوهش در روایتی معناگرایانه از این ادبیات، نخست باید بررسی کرد که چرا با وجود نظر شخص کرونباخ (۲۰۰۴)، بیشتر پژوهشگرانی که از ابزارش برای سنجش قابلیت اطمینان بهره می‌برند، با وجود ارائه این شاخص در بیش از ۵۰ سال پیش، حتی یک مطالعه ابتدائی درباره آن ندارند و تنها به‌طور مرتب به آن استناد می‌کرده‌اند. به راستی چرا آلفا کرونباخ این‌قدر شایع است و گزینه‌های جایگزین آلفا توسعه و گسترش نمی‌یابند؟

محقق تعدادی از این عوامل را براساس ادبیات نظری این حوزه این‌گونه فهرست می‌کند:

1. Neuman
2. Green & Hershberger
3. Green, Lissitz, & Mulaik
4. Green & Yang
5. Huysamen
6. Raykov
7. Sijtsma
8. Zimmerman, Zumbo, & Lalonde
9. Li
10. Hayes
11. DNA

12. Structural Equation Modeling

13. SPSS

14. SAS

15. STATA

ابزارهای جانشین برای آلفا که همانند یک ماکرو^۸ یا یک نوار ابزار از نرم‌افزارها قابلیت اجرای ساده و محاسبه این جایگزین را داشته باشد، مقاومت محققان شکسته نمی‌شود و محقق نخست باید استدلال لازم برای به کار نبردن آلفای کرونباخ را ارائه نماید (Hayes & Coutts, 2020a) و سپس روش جایگزین را با ابزاری کاربرپسند و قابل استفاده برای همگان به عنوان رسالت این پژوهش به عنوان تحقیقی ارزیابانه با رویکرد انتقادی ارائه نماید (Singh, 2019).

۳. مشاهدات استنادی تصادفی

نخست محقق در یک پیمایش مقدماتی از پژوهش‌های بازاریابی یک نمونه با ۱۰۰ مشاهده را به طور تصادفی با نرم‌افزار اکسل از بین ۱۲۴۰ مطالعه از مقالات چهار پایگاه اطلاعات استنادی علمی جهاد دانشگاهی،^۹ سیویلیکا،^{۱۰} پژوهشگاه علوم و فناوری اطلاعات ایران (ایراندک)،^{۱۱} نشریات علمی پژوهشی دانشگاه تهران، مورد ارزیابی توصیفی قرار می‌دهد و نتایج در کشور ما نسبت به مطالعات مشابه در کشورهای دیگر بسیار وضعیت نامناسب‌تری دارد. نتایج در جدول ۱ قابل ملاحظه است.

جدول ۱. نتایج نمونه‌گیری تصادفی از ۱۲۴۰ مطالعه در حوزه پژوهش‌های بازاریابی ایران

ضرایب پایایی سنجی	تعداد مطالعات	درصد مطالعات
ضریب آلفای کرونباخ	۱۰۰	۱۰۰
ضریب گاتمن	۲	۲٪
ضریب امگای مک‌دونالد	۱	۱٪
ضریب جی ال بی	۰	۰

مشاهده می‌شود که بنابر موارد بیان‌شده در مسئله تحقیق در همه این مطالعات از آلفای کرونباخ استفاده شده و در کنار آن تنها در دو مطالعه در پژوهش‌های بازاریابی از ضریب گاتمن^{۱۲} و تنها یک پژوهش آن‌هم در کنار آلفای کرونباخ (تنها به دلیل جستجوی قضاوتی اولیه محقق با کلیدواژه امگای مک‌دونالد و افزودن به

می‌کنند، عمدتاً به این دلیل که به شما امکان می‌دهد مدل‌های مختلف پایایی سنجی را محاسبه و مقایسه کنند (مثلاً گراهام،^۱ ۲۰۰۶؛ میلر،^۲ ۱۹۹۵؛ یانگ و گرین،^۳ ۲۰۱۱)؛ به عبارت دیگر محقق آگاه است که روش‌هایی مانند مدل‌سازی متغیر پنهان^۴ در تجزیه و تحلیل جریان قابلیت اطمینان بهتر عمل می‌کنند که این امر به ندرت جذاب‌ترین روش برای بیشتر محققان و کاربران علوم انسانی و اجتماعی خواهد بود؛ برای مثال مدل‌سازی معادلات ساختاری حجم نمونه‌های زیادی را می‌طلبد و همچنین شاید از نظر عملیاتی برای همه محققان به سبب لزوم تخصص برای به‌کارگیری درست آن غیرعملی باشد (Aiken & et al., 2008). این امر همراه با مشاهده‌ای که محقق این مقاله از سال ۲۰۱۰ تاکنون انجام داده، نتایج آن خبر می‌دهد که تنها ۳٪ از پژوهش‌های بازاریابی در تحصیلات تکمیلی، پوشش عمیق و کاملی از مدل‌سازی معادلات ساختاری را ارائه می‌دهند و فراموش نشود که منظور از اجرای مدل‌سازی معادلات ساختاری و استفاده عمیق و بروز آن برای آزمون فرضیات و استفاده از آزمون‌هایی برای اطمینان‌سنجی مقیاس‌ها چون پایایی اشتراکی، پایایی ترکیبی و یا حتی پایایی شاخص سازگاری اسپیرمن شاخص‌ها^۵ است و محقق با وجود آموزش گسترده این مفاهیم در طول بیش از ۱۰ سال در آکادمی تحلیل آماری ایران^۶ به بیش از ۷ هزار نفر از دانش‌پذیران تحصیلات تکمیلی بسیار بعید می‌داند که این شاخص‌های پیشرفته برآمده از مدل‌سازی معادلات ساختاری، جایگزین آلفا در آینده نزدیک شود؛ به بیان دیگر به نظر اصل مسئله تلاقی آگاهی از محدودیت‌های آلفا و فقدان گزینه‌های مناسب است (Moradi & Miralmasi, 2020, p. 68).

با توجه به نکات بیان‌شده باید جمع‌بندی کرد که محقق بنابر ادبیات فراوانی در محکومیت آلفا، پایایی سنجی با آلفای کرونباخ را نمی‌پذیرد و سال‌ها در آموزش‌های خود آن را رد کرده است؛ همچنین مطابق با نظر هایز و کات^۷ در سال ۲۰۲۰ بدون ارائه

8. Micro
9. SID
10. Civilica
11. Iran doc
12. Guttman coefficient

1. Graham
2. Miller
3. Yang, and Green
4. Latent variable
5. Rho_a
6. www.analysisacademy.com
7. Hayes & Coutts

کرونباخ از نقطه برش از قاعده حذف آیت‌های اندازه‌گیری‌کننده یک متغیر مکنون استفاده کرد (Goodboy & Martin, 2020). بنابراین براساس این داستان‌ها باید محدودیت‌های ضریب آلفای کرونباخ را به صورت پژوهشی و عملیاتی بیان کرد و محقق نیز می‌کوشد از اصطلاحات آماری کمتر استفاده کند تا برای همه مخاطبان، به ویژه دانش‌پذیران پژوهش‌های بازاریابی قابل درک باشد.

در مورد اینکه ضریب آلفای کرونباخ را با پایایی نتایج آزمون آماری یک مقیاس برابر می‌دانند باید بیان کرد که نخست اینکه آلفا برای مقیاس‌هایی با تعداد آیت‌های بالا به دلیل فرمول آن نتایج کاذب با مقدار بالای $0/7$ ارائه می‌دهد و اکنون که از سال 2010 در علوم انسانی نهضت مقیاس با آیت‌های کم^۳ بین 3 تا 5 آیت رایج شده است (M Sarstedt, Diamantopoulos & Salzberger, 2016)، نتایج معمولاً کمتر از نقطه برش نمایش داده می‌شود. این اعتقاد در علوم انسانی و اجتماعی از جمله پژوهش‌های بازاریابی ایجاد شده است که با حداقل شاخص‌ها باید حداکثر اندازه‌گیری صورت گیرد (Avkiran & Ringle, 2018)؛ بنابراین دوره مقیاس‌ها (ابزارهایی) با تعداد آیت‌های^۴ بالا به سر آمده است، زیرا اثرهای کاذب آنها بر کیفیت و برآزش اندازه‌گیری و نیز اعتبار نتایج بر هیچ‌کس پوشیده نیست (Moradi & Miralmasi, 2020, p.39)؛ به عبارتی، افزون بر باورداشت‌های نادرست در مورد اینکه آلفای کرونباخ تعیین‌کننده پایایی یک مقیاس است، باید بیان کرد که نه تنها این شاخص در اندازه‌گیری‌های کنونی و با رشد آزمون‌های پایایی جدید تنها در نقش پیش‌فرضی ابتدائی برای سنجش پایایی است، بلکه نباید غافل شد که این آزمون کاملاً به تعداد شاخص‌های یک مقیاس وابسته است و با رشد نهضت مقیاس‌های کم آیت‌م و حتی تک آیت‌می^۵ بحث سنجش قابلیت اطمینان مقیاس‌ها نیازمند ابزاری بروز و قدرتمندتر می‌باشد (Salzberger, Sarstedt & Diamantopoulos, 2016).

همچنین یکی از مواردی که بیشتر دانش‌پذیران تحصیلات

مشاهدات تصادفی) استفاده شده است. گفتنی است که این مطالعه به صورت مقدماتی ۶ سال پیش نیز بر روی مطالعات هشت رشته تحصیلی از رشته‌های علوم انسانی در قالب ۴۱۷ مطالعه انجام پذیرفت و در آنجا با توجه به احتمالی بودن انتخاب‌ها، حتی یک مورد مقاله یافت نشد که آزمونی به جز کرونباخ به کار برده باشد.

۳-۱. محدودیت‌ها و نقص‌های آلفای کرونباخ

محققان داستان افسانه‌ای نادرستی را در مورد آلفا منتشر کرده‌اند؛ از جمله اینکه:

۱. ضریب آلفای کرونباخ را با قابلیت اطمینان و پایایی نتایج اندازه‌گیری یک مقیاس برابر می‌دانند؛
۲. مقدار آلفای کرونباخ را مستقل از تعداد شاخص‌ها یا آیت‌ها در آزمون یک مقیاس می‌دانند (Hoekstra, Vugteveen, Warrens, & Kruijen, 2019)؛
۳. آلفای کرونباخ را نشانه‌ای می‌دانند از اینکه متغیرها تک‌بعدی و مرتبه اول^۱ هستند (درحقیقت آلفای کرونباخ پایایی مؤلفه‌ها یا خرده‌مقیاس‌های یک سازه را نمی‌سنجد)؛
۴. آلفای کرونباخ را بهترین انتخاب در بین ضرایب پایایی سنجی می‌دانند (Sijtsma, 2009)؛
۵. سطح خاصی از آلفای کرونباخ وجود دارد که مطلوب یا کافی است، یعنی برای آن یک نقطه برش یگانه و مطلق به مقدار $0/7$ در نظر می‌گیرند که بالای آن مقدار این ادعا وجود دارد که پایایی به حد مجاز رسیده است؛
۶. براساس کمتر بودن مقدار آلفا از نقطه برش $0/7$ به خود اجازه می‌دهند که از شاخص‌هایی استاندارد در یک مقیاس آیت‌هایی را حذف کنند (Hoekstra & et al., 2019)؛ در حقیقت اعتبار یک مقیاس در یک جامعه آماری براساس مشاهدات آن جامعه با حذف آیت‌های یک مقیاس به جهت بهبود مقدار آلفای کرونباخ از بین می‌رود. محققانی چون رایکوف^۲ عمیقاً اعتقاد دارند که نمی‌توان به بهانه کمتر بودن مقدار آلفای

3. short item scale

4. Items

5. Single item scale

1. First-order variables

2. Raykov

می‌گذارند و شاخص‌های کل یک فرم پرسشنامه^۴ را بدون درکی از مفاهیم روایی همگرا و واگرا^۵ در یک آزمون قرار می‌دهد و اصطلاحاً آلفای کل پرسشنامه را به دست می‌آورند که احتمالاً با توجه به تکرار این پدیده در این سال‌ها می‌تواند یک طنز تلخ را درباره این شاخص پایایی سنج قدیم آشکارا رؤیت کرد و جنبه طنز آن نیز اصرار تعداد زیادی از اساتید در مورد آن بر دانشجویان، بدون مطالعه‌ای عمیق در راستای بروز شدن درباره برداشتی نادرست از یک مقاله متعلق به ۷۰ سال پیش است؛ درحقیقت این تکنیک طنزگونه عجیب حتی در مقاله ۷۰ سال پیش نیز وجود نداشته است، اما با جستجوی ابتدائی در بیشتر مطالعات با رویکرد کمی در ایران مشاهده می‌شود (Moradi & Miralmasi, 2020, p. 45).

۲-۳. راه‌حل عملیاتی برای جایگزین کردن آلفای کرونباخ و شاخص‌های سنتی دیگر

همان‌طور که پیش‌تر بیان شد، برخی از محققان مانند زینگبرگ^۶ (۲۰۱۱)، سیتزما^۷ (۲۰۱۴)، رول^۸ (۲۰۰۵) و یانگ و گرین^۹ (۲۰۰۹) به محدودیت‌های آلفا اشاره کرده و آزمون‌های مناسب‌تری را برای برآورد قابلیت اطمینان تشخیص داده‌اند. این جایگزین‌ها را می‌توان در دو دسته روش‌های سنتی و معاصر طبقه‌بندی کرد. روش‌های سنتی ارزیابی پایایی با ضعف‌های ساختاری خود از زوایای مختلف برای شاخص‌هایی با ویژگی‌های کنونی کاربرد ندارد. نمونه‌هایی از این روش‌ها شامل آزمایش مجدد^{۱۰}، فرم‌های معادل یا موازی^{۱۱} و ضرایب دونیم‌سازی^{۱۲} است. متأسفانه این روش‌های سنتی سوگیری‌های^{۱۳} بسیاری دارند و خطای بالایی در برآورد پایایی وارد می‌سازند؛ از این رو توصیه می‌شود که محققان برآوردهای بدیع و مبتکرانه‌تری

تکمیلی در رشته‌های علوم انسانی و اجتماعی نسبت به آن آگاهی ندارند، شکل محتوایی و عینی شاخص‌های یک مقیاس است؛ درواقع با رشد شاخص‌های انعکاسی و ترکیبی^۱ جنگ طرفداران هر یک از این نوع شاخص‌ها در رشته‌های علوم انسانی و اجتماعی (Bollen & Diamantopoulos, 2017; Moradi & Miralmasi, 2020, p. 76) آزمون‌های تعیین‌کننده روایی و پایایی در آنها را به شدت گسترش می‌دهد و اساساً در نیمی از مطالعات دنیا که شاخص‌ها از نوع ترکیبی و سازنده هستند، نمی‌توان از آزمون‌هایی استفاده کرد که بر مبنای همبستگی شاخص‌ها همچون آلفای کرونباخ هستند (Avkiran & Ringle, 2018).

نکته بسیار مهم دیگر که دیدگاه‌های نادرست گذشته را از بین می‌برد، اینکه آلفای کرونباخ متغیرهای مکنونی را که توسط چند مؤلفه یا خرده‌مقیاس اندازه‌گیری می‌کند، تک‌بعدی یا مرتبه اول در نظر می‌گیرد و اساساً از سنجش پایایی سازه‌های دارای خرده‌مقیاس^۲ ناتوان است که در بسیاری از رشته‌های علوم انسانی و اجتماعی از جمله پژوهش‌های بازاریابی به‌وفور یافت می‌شود (Gignac, 2014)؛ برای مثال سازه کیفیت خدمات در پژوهش‌های بازاریابی را در نظر بگیرید که مطابق با مدل سروکوال^۳ دارای پنج خرده‌مقیاس قابلیت اطمینان، پاسخگویی، همدلی، اعتماد و عوامل فیزیکی است و هر یک از خرده‌مقیاس‌ها دارای تعدادی گویه یا شاخص برای اندازه‌گیری آنها است، همچنین سازه‌های این‌گونه مانند عدالت، نگرش، حمایت اجتماعی و... در رشته‌های دیگر علوم انسانی بی‌شمار است. پژوهش‌های بسیاری از محققان، متأسفانه پایایی این سازه‌ها را به صورت مرتبه اول و با آلفای کرونباخ سنجیده شده است که طبیعتاً دانشمندان حوزه سنجش آن را کاملاً غیرعلمی می‌دانند (Avkiran & Ringle, 2018) در موارد بسیاری این خطای بزرگ در کشور ما به دلیل ضعف بسیاری از اساتید و محققان در حوزه روش‌شناسی نوین و آمار مشاهده می‌شود که در مورد سازه‌های دارای خرده‌مقیاس پا را حتی فراتر از خطاهای رایج

4. questionnaire

5. Convergent and discriminant validity

6. Zinbarg

7. Sijtsma

8. Revelle

9. Yang & Green

10. Test and retest

11. Equivalent or parallel forms

12. Half split

13. Biases

1. Reflective and formative indicators

2. Component

3. SERVQUAL

مدل‌سازی معادلات ساختاری و تسلط متوسط بر نرم‌افزارهایی مانند لیزرل،^۷ ایموس،^۸ ای کیو اس،^۹ ام پلاس^{۱۰} و... . محققان نمی‌توانند که شاخص‌های نامبرده را محاسبه کنند و نیاز به تسلط بر دست‌نویسی^{۱۱} یا کار با ماکروهای^{۱۲} پیشرفته پس از اجرای مدل‌ها در این نرم‌افزارها دارند. گذشته از این نرم‌افزارها عمدتاً برای آزمون پژوهش‌هایی با رویکرد کمی که دارای فرضیات با روابط علی و تأثیرسنجی هستند، توسعه یافته‌اند و بیشتر محققان با انواع دیگری از فرضیات که به دنبال بررسی قابلیت اطمینان و یا پایایی سنجی نتایج ابزار خود هستند، ناتوان می‌مانند؛ پس این شاخص‌ها، گرچه از هر نظر فاقد مشکلات آلفای کرونباخ هستند و تصویری دقیق واقعی از پایایی شاخص‌های یک متغیر ارائه می‌دهند، اما تنها برای کسانی راه‌حلی عملیاتی برای جایگزینی آلفای کرونباخ می‌باشد که تسلط بالایی بر روش تحقیق مدرن و نرم‌افزارهای تخصصی آن دارند (Moradi & Miralmasi, 2020, p. 68).

۳-۳. ضریب امگا راه‌حل عملیاتی برای پایان دادن بر مشکل پایایی سنجی

اکنون پس از بررسی جامعی از ضریب آلفای کرونباخ، مشکلاتی که در جوهره این ضریب رایج در علوم انسانی در این ۷۰ سال وجود داشته است، پیشینه‌ای مطالعاتی آن و ایراداتی که صاحب‌نظران بر وجوه مختلف روش‌شناختی آن وارد کرده‌اند و نیز در نهایت ذکر ضرایب و راه‌حل‌های جایگزین برای آلفای کرونباخ، همچنین مشکلاتی که می‌تواند از اجرای همگانی و رایج آنها در علوم انسانی و اجتماعی، به‌ویژه حوزه پژوهش‌های بازاریابی مانع گردد، این سؤال مطرح می‌شود که جایگزین مناسب و عملیاتی برای این ضریب که بتواند همانند آلفای کرونباخ رواج فراوانی یابد و نیز فاقد مشکلات گفته شده باشد، چیست؟ پاسخ این پرسش با تشریح و تبیین دلایل آن پیشنهاد ضریب امگای مک‌دونالد است (Hayes & Coutts, 2020a).

از قابلیت اطمینان را که مورد بحث قرار گرفته است، از نظر عملیاتی بودن ارزیابی کنند (Starkweather, 2012).

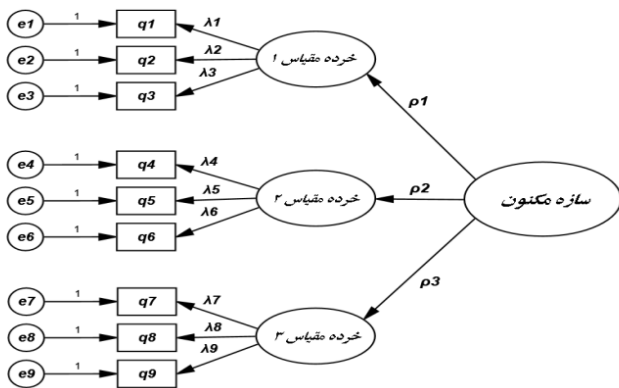
برخی از مشکلاتی را که برای آلفای کرونباخ ذکر شد، می‌توان با به دست آوردن فواصل اطمینان^۱ به‌صورت تقریباً ساده‌ای برطرف کرد. یکی از راه‌های دستیابی به این فواصل اطمینان روشی است که به‌عنوان نمونه‌گیری مجدد خودکار^۲ شناخته می‌شود؛ برای مثال لاکوبسی و دوهاچک^۳ (۲۰۰۳) یک روش ساده بوت استرپینگ را برای به دست آوردن فاصله اطمینان برای آلفا ارائه می‌دهند. بوت استرپینگ شامل نمونه‌برداری مکرر با جایگزینی از یک نمونه برای به دست آوردن توزیع تجربی یک برآوردگر مانند آلفاست (DiCiccio & Efron, 1996). هنگامی که این توزیع به دست آمد، می‌توان فواصل اطمینان یا سایر مقادیر را از آن ایجاد کرد (مثلاً در ساده‌ترین حالت مستقیماً از توزیع تجربی گرفته می‌شود). با وجود جذابیت بوت استرپینگ، محققان به دلیل اتکای این روش به مفروضات توزیعی یا سایر مفاهیم راه‌حلی دشوار می‌پندارند، زیرا به‌ندرت در عمل مفروضات و پیش‌شرط‌های آن وجود خواهد داشت؛ به عبارت روشن‌تر در مورد این روش، فرض نرمال بودن چندمتغیره که با روش‌های تحلیلی معمول فرض می‌شود، به‌ندرت وجود دارد (Revelle & Zinbarg, 2008).

راه‌حل‌های بعدی شاخص‌های سازگاری درونی برای مدل‌هایی با متغیرهای مکنون انعکاسی مانند پایایی ترکیبی^۴ و میانگین واریانس استخراجی^۵ و نیز شاخص پایایی اشتراکی^۶ است که البته هر سه شاخص با وجود دقت بالا و نتایج قابل اعتمادی که دارند، نیازمند تسلط بر روش آماری مدل‌سازی معادلات هستند و همان‌طور که بیان شد در کشور ما ضعف شدید کار با نرم‌افزارهای پیشرفته آماری و دست‌نویسی در صفحات آنها بین محققان در مقطع ارشد و دکتری و حتی اساتید راهنما و مشاور وجود دارد؛ البته حتی با آشنایی نسبی با

7. LISREL
8. AMOS
9. EQS
10. MPLUS
11. Syntax
12. Micro

1. confidence intervals
2. bootstrapping
3. Lacobucci and Duhachek
4. Composite Reliability
5. Average variance extracted (AVE)
6. Commuality

می‌کند و خطاهای اندازه‌گیری را نمی‌تواند محاسبه نماید (Moradi & Miralmasi, 2020, p. 44). اما شاخص امگای مک‌دونالد با توجه به محاسبه بارهای عاملی برای هر آیت‌م و به دنبال آن اندازه‌گیری واریانس تبیین شده^۵ یا همان سهم اندازه‌گیری هر آیت‌م برای یک عامل مکنون و نیز محاسبه خطاها، باید در سال‌هایی بسیار پیش‌تر جایگزین آلفای کرونباخ به‌عنوان ابزار پایایی‌سنجی قرار می‌گرفت (Trizano-Hermosilla & Alvarado, 2016)؛ همچنین نکته بسیار مهم اینکه با پررنگ‌تر شدن سازه‌ها یا متغیرهای مکنونی با درجه انتزاع بالاتر که دارای خرده‌مقیاس یا مؤلفه^۶ در علوم انسانی و اجتماعی هستند و توسعه متغیرهای مرتبه بالاتر (Marko Sarstedt & et al., 2018) در پژوهش‌ها، مشکلات آلفای کرونباخ در پایایی‌سنجی آنها، خطا و عدم دقت خود را هرچه بیشتر نشان می‌دهند. خوشبختانه امگای مک‌دونالد این مسئله مهم را نیز با محاسبه سهم هر خرده‌مقیاس در اندازه‌گیری سازه حل کرده است و با دقت بسیار بالایی قابلیت اطمینان یا پایایی اندازه‌گیری سازه‌هایی را سنجش می‌کند که دارای ابعاد پنهان یا خرده‌مقیاس هستند (در علوم انسانی این سازه‌ها به‌وفور وجود دارند) (Hayes & Coutts, 2020b).



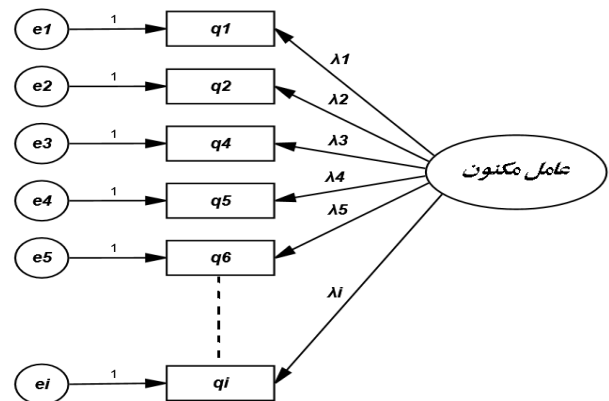
شکل ۳. مدل مرتبه دو برای محاسبه سهم سوالات و سهم

خرده‌مقیاس‌ها در سنجش سازه

5. Explained variance
6. Component

استدلال اصلی برای پیشنهاد این جایگزینی آن است که درست مانند پایایی ترکیبی در مدل‌سازی معادلات، ضریب نیز با استفاده از CFA^۱ و محاسبه بارهای عاملی^۲ بین متغیرهای مکنون، آیت‌ها و واریانس خطاهای اندازه‌گیری^۳ در معادله آن محاسبه می‌شود (McDonald, 1999). اگرچه در این پژوهش محقق به دنبال جذب و مطالعه حداکثری محققان و دوری از استفاده از معادلات و فرمول‌های آماری است، اما ناگزیر باید گفت که فرمول ضریب امگا همان‌طور که مشاهده می‌شود دارای مجذور جمع بارهای عاملی^۴ $(\sum \lambda_i)^2$ بر مجذور جمع بارهایی می‌باشد که با مجموع خطاهای اندازه‌گیری تجمیع شده است؛ در واقع این فرمول درست مشابه با مقدار پایایی ترکیبی در مدل‌سازی معادلات ساختاری، البته بدون قرارگیری در ساختار یک مدل و تأثیر ضرایب مسیر^۴ بر این ضریب سنجش قابلیت اطمینان است.

$$\omega = \frac{(\sum \lambda_i)^2}{[(\sum \lambda_i)^2 + (\sum e_i)]}$$



شکل ۲. مدل محاسبه بارهای عاملی و خطاهای اندازه‌گیری

در یک عامل مکنون

مشاهده می‌شود که در این شاخص برخلاف ضریب آلفای کرونباخ که سهم سوالات در اندازه‌گیری متغیرها را یکسان فرض

1. confirmatory factor analysis
2. Factor loading
3. Variance of measurement errors
4. Path coefficients

سنتی دیگر می‌تواند آن را محاسبه کند. محقق نیز در مکاتباتی با شرکت‌های تولیدکننده این نرم‌افزارها، دستورهای با زبان پایتون^۸ برای محاسبه امگای مک‌دونالد به آنها ارائه کرده که با استقبال برخی از شرکت‌ها در افزودن این ویژگی به نسخه‌های آتی نرم‌افزارهای آنها قرار گرفته است.

۴. بحث و نتیجه‌گیری

روش علمی در قالبی نظام‌مند به محقق رعایت مجموعه‌ای از باورداشت‌ها و قراردادهای مشترک را در مواجهه با مسئله به‌عنوان پارادایم یادآور نمود (Kaushik & Walsh, 2019)؛ بر این اساس محقق به سمت رویکردی کمی و خلاصه‌ای اجمالی از آن حرکت کرد؛ اما در میان این قراردادهای مشترک ضروری در پارادایم‌های پوزیتیویستی، تعمیم‌پذیری نتایج به‌دست آمده از داده‌های یک ابزار کمی به‌عنوان هدف اصلی یک تجزیه و تحلیل استنباطی، نه قابل چشم‌پوشی و نه تساهل در مورد ضعف‌های محاسباتی آن بود. ملاحظه شد که مسئله اصلی پژوهش چه از نظرگاه فنی و چه محتوایی آبشخور همین بخش است و با ظهور ضریبی با نام آلفا از سوی کرونباخ (۱۹۵۱) در بیش از ۷۰ سال پیش و ضعف شدید شاخص‌های سنتی دیگر و در نتیجه رواج آن با توجه به دلایل نامبرده و نیز مقاومت عجیب آن در برابر دیدگاه‌های انتقادی و جایگزینی آن در این سال‌ها، به‌ویژه در علوم انسانی و اجتماعی ازجمله پژوهش‌های بازاریابی به‌عنوان مورد مطالعه، انجام یک پژوهش انتقادی همراه با ارائه و پیشنهاد راه‌حل را ضرورت می‌بخشید. افزون بر آن با مطالعه دقیق ادبیات این حوزه امری ثابت شده است که آلفای کرونباخ در موارد زیر متورم و نتایج کاذب ارائه می‌دهد؛ به عبارتی با افزایش تعداد مشاهدات، متغیرهای آشکار مربوط به هر متغیر مکنون، محتوای آیت‌های با مقادیر تکراری و نیز افزایش مشابهت‌های آنها این ضریب متورم و دچار سوگیری می‌شود (Starkweather, 2012). این موارد در کشور ما خطر جعل در داده‌های پژوهش‌ها را به دلیل فضای نامبارک فروش مقاله و پایان‌نامه و متأسفانه دستکاری در داده‌ها را به شدت افزایش داده است و این نگرانی ایجاد شده که سهولت افزایش کاذب آلفا، این ضریب را برای مدت طولانی‌تری

همچنین بنابر نظر جوالی^۱ و همکارانش (۲۰۱۱) حجم نمونه تأثیر بسیار اندکی بر مقدار ضریب امگای مک‌دونالد دارد و این شاخص در نتیجه کاهش و افزایش در حجم نمونه در پژوهش‌های مختلف درگیر سوگیری نمی‌گردد (Teo & Fan, 2013)؛ در نهایت به دلیل مسئله و خواسته نویسنده مقاله که می‌خواهد از بیان مباحث پیچیده آماری که طیف مخاطبان را محدود می‌کند و این با هدف مقاله در تقابل است، تنها به ذکر چند نکته بالا بسنده می‌شود و بیان می‌دارد که در نرم‌افزارهایی چون SPSS^۲، SAS^۳ و STATA^۴ که برخلاف معادلات ساختاری، نمی‌توانند مدل‌های اندازه‌گیری و تحلیل عاملی تأییدی^۵ را در نسخه‌های متداول خود اجرا کنند، می‌توان با دستوراتی به محاسبه تحلیل عاملی اکتشافی^۶ در قالب برآوردگر حداکثر درست‌نمایی^۷ برای متغیرهای مکنون مرتبه ۱ و حتی مرتبه بالاتر، اقدام به محاسبه شاخص امگای مک‌دونالد به‌جای آلفای کرونباخ به‌عنوان جایگزینی با دقت بالا و خطای اندک جهت دستیابی به قابلیت اطمینان نتایج ابزارهای سنجش کمی در پژوهش‌های اجتماعی ازجمله پژوهش‌های بازاریابی به‌سادگی پرداخت.

همچنین در نهایت شاخص پایایی امگای مک‌دونالد به‌عنوان جایگزینی عملیاتی برای همه مشکلات و مسائلی را که آلفای کرونباخ در طول بیش از ۷۰ سال برای پژوهش‌ها به دنبال داشته، به محقق ارائه می‌کند. مقدار این شاخص مانند پایایی مرکب باید از نقطه برش ۰/۷ بیشتر باشد (Gignac, 2014) و با توجه به مطالعات کلاین ۲۰۱۶ برای مدل‌های تازه تولدیافته، مقادیر بالای نقطه برش ۰/۶ نیز نتایج قابل اطمینانی را از مقیاس ارائه می‌دهد (Kline, 2016). خوشبختانه هایز و همکارانش (۲۰۲۰) دستورات اجرای این ابزار پایایی سنج را برای نرم‌افزارهای مادر مانند اس. پی. اس. اس، استتا و سس نوشته و بسته‌هایی تهیه کرده‌اند که مانند افزونه روی این نرم‌افزارها نصب شده و هر محقق بی‌سازمان ساده‌تر از شاخص آلفای کرونباخ و شاخص‌های

1. Javali
2. statistical package for social science
3. statistical analysis software
4. Software for Statistics and Data Science
5. Confirmatory factor analysis
6. Exploratory factor analysis
7. Maximum likelihood

8. Python

به‌عنوان یک استاندارد پایایی سنجی در علوم انسانی و اجتماعی باقی گذارد.

یکی دیگر از موارد مهم که به‌عنوان یک نتیجه قطعی می‌توان آن را بیان کرد؛ اینکه در نرم‌افزارهای آماری مختلف مانند اس. پی. اس. در بخش محاسبه آلفای کرونباخ گزینه‌ای وجود دارد که مقدار آلفای کرونباخ را با حذف آیتم مورد نظر^۱ می‌خواند و این ویژگی متأسفانه به دلیل ناآگاهی لازم یک ابزار یا مقیاس اندازه‌گیری‌کننده متغیر یا مؤلفه را ناقص می‌کند که در یک فرایند طولانی و پیچیده با رویکردی آمیخته در یک فرایند مقیاس‌سازی استاندارد و با انجام انواع آزمون‌ها و شاخص‌های مختلف اکتشافی و تأییدی ایجاد شده است؛ بنابراین بنا به نظر رایکوف (۲۰۰۸) و یا آلن و همکارانش (۲۰۲۰) نمی‌توانیم برای اینکه در جامعه یک تحقیق قابلیت اطمینان به ابزار را از مقداری مشخص افزایش دهیم، دست به ناقص کردن مقیاس و حذف آیتم‌های یک مقیاس بزنیم و این گزینه نرم‌افزاری تنها هشدار به محقق برای بهبود و اصلاح آن آیتم است و باید این روش حذفی آیتم‌های یک مقیاس در جمعیت آماری هدف کنار گزارده شود.

در نهایت راه‌حل‌های بسیاری در طی سالیان ارائه شد که خود به دلیل مسئله دیگری منتفی شد که همانا ضعف محققان و اساتید آنها به روش‌های پژوهشی و آماری نوین است و محققان آتی قطعاً باید دلایل آن را مورد مطالعه قرار دهند؛ از این رو باید مسئله را از همان نقطه‌ای حل کرد که دست‌کم با رشد نرم‌افزارها تبدیل به مسئله شده است و این نقطه سادگی محاسبه آلفای کرونباخ با نرم‌افزارهای مادر آماری است. پس محقق با در اختیار قرار دادن بسته‌هایی که پیش‌ازین توسط خبرگان حوزه پژوهش مانند هیز و همکارانش (۲۰۲۰) دست‌نویسی شده است و به آسانی قابل نصب و اجراست؛ به دنبال رواج عملیاتی شاخص امگای مک‌دونالد به‌عنوان جایگزین آلفای کرونباخ داده شد. همچنین نباید غافل شد که روش‌های علمی همواره در قالب فن‌آوری‌های نرم‌افزاری در حال رشد است و مقالات انتقادی از این دست می‌تواند در مدتی کوتاه شرکت‌های نرم‌افزاری را به سمت برطرف کردن نقص‌ها رهنمون سازد.

در پایان باید اشاره کرد که این ضریب در مقام قیاس با دیگر جایگزین‌ها که در بخش‌های پیشین بیان شد، هم نقاط ضعف آلفای کرونباخ و شاخص‌های سنتی دیگر را ندارد و هم نیاز به تخصص پیچیده نرم‌افزاری مثل دست‌نویسی در نرم‌افزارهای مدل‌سازی معادلات ساختاری مانند ام‌پلاس را نخواهد داشت؛ همچنین محققان بسیاری پیش‌تر بر برتری این شاخص نسبت به آلفا صحه گذاشته بودند، اما محاسبه آن تا سال گذشته برای عموم پژوهشگران اجرایی نبود. اکنون محققان آتی و نیز توسعه‌دهندگان نرم‌افزارهای آماری نیز وظیفه دارند که با مطالعات تطبیقی هر یک از این شاخص‌ها و استخراج نقاط قوت و ضعف آنها جلوی تثبیت بیشتر و بی‌دلیل یک شاخص آماری ضعیف و بسیار قدیمی را که اندازه‌گیری‌کننده یکی از مهم‌ترین ویژگی‌های تحقیق با رویکرد کمی، یعنی تعمیم‌پذیری نتایج را بگیرند و باید گوشزد کرد که با توجه به سهم بسیار بالای پژوهش‌ها با رویکرد کمی در حوزه‌های علم انسانی، محققان تنها نباید به نقد و واکاوی پژوهش‌های کیفی بپردازند، زیرا همان‌طور که بیان شد دانش علمی از دو ستون اکتشاف و آزمون اکتشافات جهت بسط نظریات به جمعیتی بزرگ‌تر تشکیل شده است. محققان آتی باید در نظر گیرند که با رویکردهای اثبات‌گرایانه برخوردی اثباتی و غیرمنعطف نداشته و همواره فضای نقد دانشگاهی به همه جنبه‌های روش‌شناسی علمی را مدنظر قرار دهند. امید است این مطالعه که رسالت خود برای پیشنهاد کاربست رویه‌های نو و بدیع را در روش‌شناسی رشته‌های مختلف علوم انسانی و اجتماعی انجام داده است، فضا را برای نقادی منصفانه و مبتنی بر چارچوبی نظری البته با ارائه راه‌حل گشوده و با نگاهی فرای روش‌شناسی مرسوم به دنبال شناسایی و رفع خلأها در این بخش‌ها بوده باشد.

1. Alpha if item deleted

منابع

9. Dunn, T. J., Baguley, T. & Brunsten, V. (2014), From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399–412. <https://doi.org/10.1111/bjop.12046>.
10. Gignac, G. E. (2014), On the Inappropriateness of Using Items to Calculate Total Scale Score Reliability via Coefficient Alpha for Multidimensional Scales. *European Journal of Psychological Assessment*, 30(2), 130–139. <https://doi.org/10.1027/1015-5759/a000181>.
11. Goodboy, A. K. & Martin, M. M. (2020), Omega over alpha for reliability estimation of unidimensional communication measures, *Annals of the International Communication Association*, 44(4), 422–439. <https://doi.org/10.1080/23808985.2020.1846135>.
12. Graham, J. M. (2006), Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them, *Educational and Psychological Measurement*, 66(6), 930–944.
13. Green, S. B. & Hershberger, S. L. (2000), Correlated errors in true score models and their effect on coefficient alpha, *Structural Equation Modeling*, 7(2), 251–270.
14. Green, S. B., Lissitz, R. W. & Mulaik, S. A. (1977), Limitations of coefficient alpha as an index of test unidimensionality1, *Educational and Psychological Measurement*, 37(4), 827–838.
15. Green, S. B. & Yang, Y. (2009), Commentary on coefficient alpha: A cautionary tale, *Psychometrika*, 74(1), 121–135.
16. Hayes, A. F. & Coutts, J. J. (2020a), Use Omega Rather than Cronbach's Alpha for Estimating Reliability, But... *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>.
1. Aiken, L. S., West, S. G. & Millsap, R. E. (2008), Doctoral training in statistics, measurement, and methodology in psychology: Replication and extension of Aiken, West, Sechrest, and Reno's (1990) survey of PhD programs in North America. *American Psychologist*, 63(1), 32.
2. Aikin, S. F. & Talisse, R. B. (2016), Pragmatism and pluralism revisited. *Political Studies Review*, 14(1), 17–26. <https://doi.org/10.1177/1478929915607886>.
3. Avkiran, N. K. & Ringle, C. M. (2018), *Partial least squares structural equation modeling (PLS-SEM)*, Springer International Publishing AG. Retrieved from <http://www.emeraldinsight.com/doi/abs/10.1108/EBR-10-2013-0128%5Cnhttp://www.emeraldinsight.com/10.1108/EBR-10-2013-0128>.
4. Bhattacharjee, A. (2012), *Scholar Commons Social Science Research: Principles, Methods, and Practices. Textbooks Collection*. Retrieved from http://scholarcommons.usf.edu/oa_textbookshttp://scholarcommons.usf.edu/oa_textbooks/3.
5. Bielik, L. (2019), *methodology of science an introduction* (1st ed.). Bratislava: comenius university in bratislava.
6. Bollen, K. A. & Diamantopoulos, A. (2017), Notes on measurement theory for causal-formative indicators: A reply to Hardin, *Psychological Methods*, 22(3), 605–608. <https://doi.org/10.1037/met0000149>.
7. Cronbach, L. J. (2004), My current thoughts on coefficient alpha and successor procedures, *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 64(3), 391–418. <https://doi.org/10.1177/0013164404266386>.
8. DiCiccio, T. J., & Efron, B. (1996). Bootstrap confidence intervals. *Statistical Science*, 11(3). <https://doi.org/10.1214/ss/1032280214>.

27. Naik, C. N. K., Gantasala, S. B. & Prabhakar, G. V. (2010), Service quality (SERVQUAL) and its effect on customer satisfaction in retailing, *European Journal of Social Sciences*, 16(2), 231–243.
28. Neuman, W. L. (2014), *Social Research Methods: Qualitative and Quantitative Approaches*. Pearson Education Limited (7th ed.), Printed in the United States of America. <https://doi.org/10.2307/3211488>.
29. Revelle, W. & Zinbarg, R. E. (2008), Coefficients Alpha, Beta, Omega and the glb: Comments on Sijtsma, *Psychometrika*, 74(1), 145. <https://doi.org/10.1007/s11336-008-9102-z>.
30. Salzberger, T., Sarstedt, M. & Diamantopoulos, A. (2016), Measurement in the social sciences: where C-OAR-SE delivers and where it does not, *European Journal of Marketing*, 50(11), 1942–1952. <https://doi.org/10.1108/EJM-10-2016-0547>.
31. Sarstedt, M., Diamantopoulos, A. & Salzberger, T. (2016), Should we use single items? Better notF_. *Journal of Business Research*, 69, 3199–3203.
32. Sarstedt, Marko, Bengart, P., Shaltoni, A. M. & Lehmann, S. (2018), The use of sampling methods in advertising research: a gap between theory and practice, *International Journal of Advertising*, 37(4), 650–663. <https://doi.org/10.1080/02650487.2017.1348329>.
33. Sijtsma, K. (2009), On the use, the misuse and the very limited usefulness of cronbach's alpha, *Psychometrika*, 74(1), 107–120. <https://doi.org/10.1007/s11336-008-9101-0>.
34. Singh, D. (2019), Understanding philosophical underpinnings of research with respect to various paradigms: Perspective of a research scholar, (April).
35. Smith, S. M. & Albaum, G. S. (2013), *Basic Marketing Research Building Your Survey*, Published by Qualtrics Labs, Inc.
36. Starkweather, J. (2012), Step out of the past: Stop using coefficient alpha; there are better ways to calculate reliability, *Research and Statistical Support*.
17. Hayes, A. F. & Coutts, J. J. (2020b), Use Omega Rather than Cronbach's Alpha for Estimating Reliability, But.... *Communication Methods and Measures*, 00(00), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>.
18. Hoekstra, R., Vugteveen, J., Warrens, M. J. & Kruijen, P. M. (2019), An empirical analysis of alleged misunderstandings of coefficient alpha. *International Journal of Social Research Methodology*, 22(4), 351–364. <https://doi.org/10.1080/13645579.2018.1547523>.
19. Huysamen, G. K. (2006), Coefficient alpha: unnecessarily ambiguous; unduly ubiquitous, *SA Journal of Industrial Psychology*, 32(4), 34–40.
20. Johnson Burke., R. & Christensen, L. (2014), *Educational Research: Quantitative, Qualitative, and Mixed Approaches* (5th ed., Vol.4).
21. Kaushik, V. & Walsh, C. A. (2019), Pragmatism as a research paradigm and its implications for Social Work research, *Social Sciences*, 8(9). <https://doi.org/10.3390/socsci8090255>.
22. Kline, R. B. (2016), *Principles and practices of structural equation modelling*, *Methodology in the social sciences*.
23. McCain, K. (2016), *The Nature of Scientific Knowledge, The Nature of Scientific Knowledge*, Springer International Publishing. <https://doi.org/10.1007/978-3-319-33405-9>.
24. McDonald, R. P. (1999), *Test theory: A unified treatment. Test theory: A unified treatment*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
25. Miller, M. B. (1995), Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling.
26. Moradi, M. & Miralmasi, A. (2020), *Pragmatic research method* (F. Seydi, Ed.) (1st ed.), Tehran: School of quantitative and qualitative research, Retrieved from <https://analysisacademy.com/>

- nonhomogeneous items, *Applied Psychological Measurement*, 22(4), 375–385.
<https://doi.org/10.1037/0003-066X.54.8.594>.
42. Yang, Y. & Green, S. B. (2011), Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*, 29(4), 377–392.
43. Zimmerman, D. W., Zumbo, B. D. & Lalonde, C. (1993), Coefficient alpha as an estimate of test reliability under violation of two assumptions, *Educational and Psychological Measurement*, 53(1), 33–49.
44. Zinbarg, R. E., Revelle, W., Yovel, I. & Li, W. (2005), Cronbach's α , Revelle's β and McDonald's ω H: Their relations with each other and two alternative conceptualizations of reliability, *Psychometrika*, 70(1), 123–133.
<https://doi.org/10.1007/s11336-003-0974-7>.
37. Teo, T. & Fan, X. (2013), Coefficient alpha and beyond: Issues and alternatives for educational research, *The Asia-Pacific Education Researcher*, 22(2), 209–213.
38. Trizano-Hermosilla, I. & Alvarado, J. M. (2016), Best Alternatives to Cronbach's Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements. *Frontiers in Psychology*, 7, 769.
<https://doi.org/10.3389/fpsyg.2016.00769>.
39. Wertz, F. J., Charmaz, K., McMullen, L. M., Josselson, R., Anderson, R. & McSpadden, E. (2011), *Five Ways of Doing Qualitative Analysis*, The Guilford Press, New York (1st ed.).
40. Wilkinson, L. (1999), Statistical methods in psychology journals: Guidelines and explanations, *American Psychologist*, 54(8), 594–604.
<https://doi.org/10.1037/0003-066X.54.8.594>.
41. Wilkinson, L. & Raykov, T. (1998), Coefficient alpha and composite reliability with interrelated