

Designing, Constructing, and Analyzing Likert Scale Data

Alireza Mirahmadizadeh (PhD)¹, Hamed Delam (MSc)^{2,*}, Mozhgan Seif (PhD)²,
Reyhaneh Bahrami (BSc)³

¹Non-Communicable Diseases Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

²Epidemiology Department, Health Faculty, Shiraz University of Medical Sciences, Shiraz, Iran

³“Student Research Committee” and “Public Health Department, Health Faculty”, Shiraz University of Medical Sciences, Shiraz, Iran

*Corresponding Author: Hamed Delam, Epidemiology Department, Health Faculty, Shiraz University of Medical Sciences, Shiraz, Iran. Email: hameddelam8@yahoo.com

Abstract

Received: 13/12/2017

Accepted: 13/10/2018

How to Cite this Article:

Mirahmadizadeh A, Delam H, Seif M, Bahrami R. Designing, Constructing, and Analyzing Likert Scale Data. *J Educ Community Health*. 2018; 5(3): 63-72. DOI: 10.21859/jech.5.3.63

Background & Objective: Today, Likert scale data is used in many studies in medical, educational, and psychometric fields to understand the attitudes of individuals. This study was conducted with the aim of correctly understanding how to design a Likert scale and to analyse the issues and challenges in determining the reliability and analysis of Likert scale data.

Instruments & Methods: In the current review study, search engines such as PubMed, Google Scholar, and Science Direct were searched between May and August 2017. Based on the searched terms, 532 articles were primarily obtained; according to the inclusion and exclusion criteria of the study, eventually 60 articles were examined.

Results: Using the 5-point scale is more common than other multi-choice options. Considering the option as intervals encourages people to answer the questions. While using Likert scale data, there are always decision-making challenges, especially when discussing the type of test to be used for parametric or non-parametric tests. If there is a presumption of the normalization of the distance data, independent t test is used for the two separate groups.

Conclusion: In most of the Likert scale studies that have been carried out so far, there is some confusion or disagreement with how to design, construct, and analyse its data. Ultimately, although there is no detailed and comprehensive agreement on the design and analysis of Likert data, this study, with a general summary in this regard, has greatly curtailed the confusion.

Keywords: Attitude; Likert Scale; Cronbach's alpha; Measurement

طراحی، ساخت و تجزیه و تحلیل داده‌های مقیاس لیکرت

علیرضا میراحمدی زاده^۱، حامد دلام^{۲*}، مژگان سیف^۳، ریحانه بهرامی^۴

^۱دکترای تخصصی، مرکز تحقیقات بیماری‌های غیر واگیر، دانشگاه علوم پزشکی شیراز، شیراز، ایران
^۲کارشناس ارشد، گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شیراز، شیراز، ایران
^۳دکترای تخصصی، گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شیراز، شیراز، ایران
^۴کارشناس، "کمیته تحقیقات دانشجویی" و "گروه بهداشت عمومی، دانشکده بهداشت"، دانشگاه علوم پزشکی شیراز، شیراز، ایران

*نویسنده مسئول: حامد دلام، گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شیراز، شیراز، ایران.

ایمیل: hameddelam8@yahoo.com

چکیده

سابقه و هدف: امروزه داده‌های مقیاس لیکرت، در مطالعات گسترده در زمینه‌های پزشکی، آموزشی و روان‌سنجی، برای پی‌بردن به نگرش افراد مورد استفاده قرار می‌گیرد. مطالعه حاضر با هدف شناخت صحیح نحوه ساخت و طراحی مقیاس لیکرت و همچنین تحلیل مسایل و چالش‌های موجود در نحوه تعیین پایایی و تجزیه و تحلیل داده‌های مقیاس لیکرت انجام شد.

ابزار و روش‌ها: در مطالعه مروری حاضر، موتورهای جستجوی پابمد، گوگل اسکالر و ساینس دایرکت در بازه زمانی خرداد تا شهریورماه سال ۱۳۹۶ مورد جست‌وجو قرار گرفتند. براساس واژگان جست‌وجوشده، در ابتدا ۵۳۲ مقاله حاصل شد که با توجه به معیارهای ورود و خروج از مطالعه، در نهایت ۶۰ مقاله مورد بررسی قرار گرفت.

یافته‌ها: استفاده از مقیاس ۵گزینه‌ای نسبت به انواع چندگزینه‌ای دیگر، رایج‌تر است. در نظر گرفتن گزینه‌های حد فاصل، افراد را در پاسخ به سئوالات، دلگرم می‌کند. در استفاده از داده‌های مقیاس لیکرت، همواره چالش‌های تصمیم‌گیری وجود دارد، به‌ویژه زمانی که بحث در مورد نوع آزمون مورد بررسی باشد، اینکه از آزمون‌های پارامتریک یا ناپارامتریک استفاده شود، در صورت برقراربودن فرض نرمالیتی داده‌های فاصله‌ای، برای دو گروه مجزا، از آزمون t مستقل استفاده می‌شود.

نتیجه‌گیری: در اکثر مطالعات مرتبط با مقیاس لیکرت که تاکنون انجام شده است، نوعی سردرگمی یا عدم توافق در نحوه طراحی، ساخت و تجزیه و تحلیل داده‌های آن وجود دارد. در نهایت با وجود آنکه تاکنون پیمان‌نامه دقیق و جامعی در مورد طراحی و تجزیه و تحلیل داده‌های لیکرت وجود ندارد، این مطالعه با ارائه جمع‌بندی کلی در این زمینه، تا حدود زیادی از سردرگمی‌ها کاست.

واژگان کلیدی: نگرش؛ مقیاس لیکرت؛ آلفای کرونباخ؛ اندازه‌گیری

مقدمه

استفاده می‌شود، به‌طور مثال در سال ۱۹۹۸، اسپرینگر نگرانی‌های مربوط به نوجوانان که رفتار فرار از جامعه و خانواده را در آنها تقویت می‌کرد، مورد ارزیابی قرار داد. همچنین در سال ۲۰۰۳، کویلند مشکلات زنان جوانی که در شرایط محروم از امکانات زندگی می‌کردند را سنجید، در سال ۲۰۰۷، لیندورست و همکاران، نگرش افراد را در رابطه با برچسب‌های بسته‌های غذایی سنجیدند، در همان سال، براون و همکاران نگرش و پایبندی افراد به رفتارهای مرتبط با سلامتی از جمله فعالیت ورزشی را بررسی کردند [۱]. محبوبیت مقیاس لیکرت میان تحقیقات بیشتر به دلایلی مانند

مقیاس لیکرت که در اصل توسط رنسیس لیکرت در سال ۱۹۳۲ معرفی شد، به‌طور گسترده در تحقیقات روان‌سنجی، اندازه‌گیری نگرش و باور افراد مورد استفاده قرار می‌گیرد. مقیاس لیکرت به‌طور گسترده برای اندازه‌گیری خواص قابل مشاهده در زمینه‌های علوم اجتماعی و انسانی استفاده شده است [۱]. به‌طور مثال اور و اولسون علاقه به آموزش موسیقی را مورد ارزیابی قرار دادند [۲]. مثال دیگر در مورد اثربخشی داروها در داروشناسی است که توسط محققان مورد بررسی قرار گرفت. علاوه بر این، مقیاس لیکرت برای اندازه‌گیری ساختارهای پنهانی که مستقیماً قابل مشاهده نیستند نیز

ابزار و روش‌ها

در مطالعه مروری حاضر، موتورهای جستجوی پابمد، گوگل اسکولار و ساینس دایرکت با واژگان، "Likert scale"، "Measurement"، "Attitude"، "Cronbach's alpha"، جست‌وجو شدند. معیارهای ورود مطالعات به پژوهش شامل موارد زیر بودند:

(۱) مطالعه پژوهشی اصیل باشد، (۲) گزارش پژوهش کامل باشد، (۳) در انجام مطالعه از یک یا چند مورد از الگوهای آموزش بهداشت و ارتقای سلامت استفاده شده باشد، (۴) تعداد ارجاع نسبت به مقالات مشابه دیگر بیشتر باشد یا محتوای مقالات توسط سه تن از اساتید مجرب در این حوزه در دانشگاه علوم پزشکی شیراز به تایید برسد.

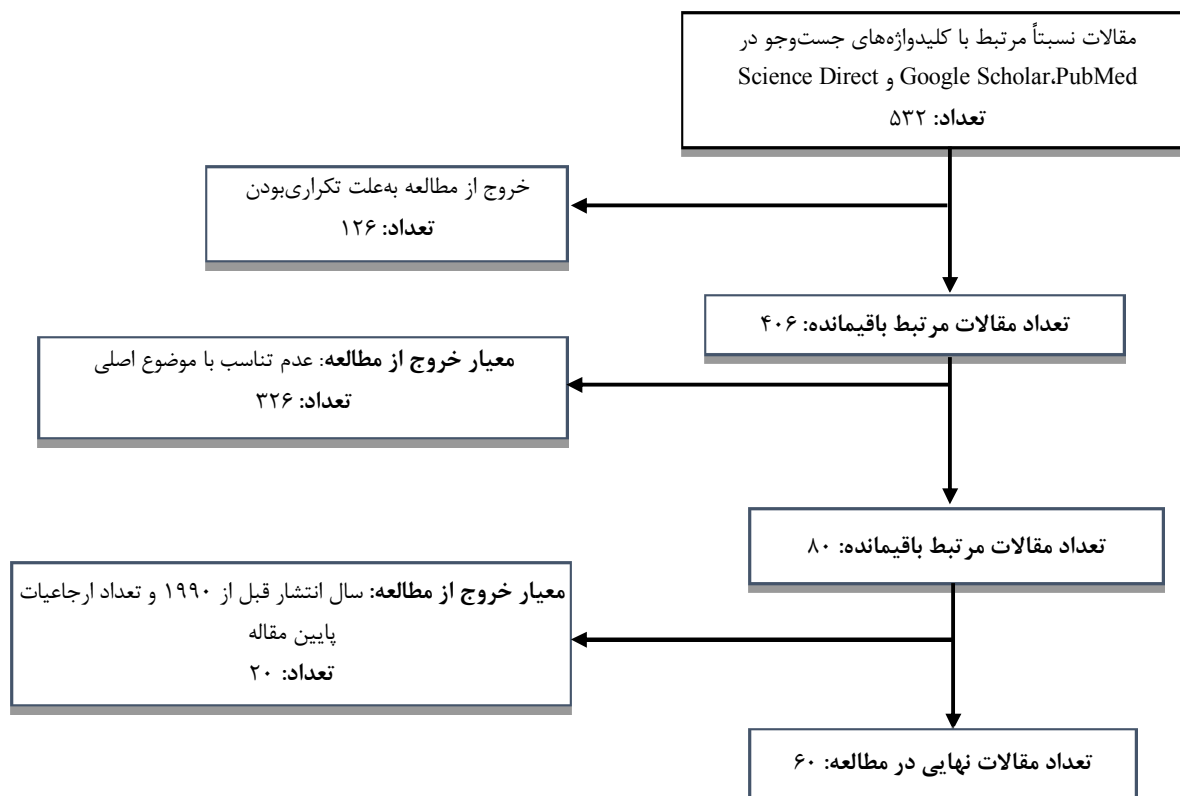
تکراری بودن محتوای مقالات، عدم تناسب با موضوع اصلی مطالعه و همچنین سال انتشار قبل از ۱۹۹۰ با تعداد ارجاعیات کم نیز به عنوان معیارهای خروج از مطالعه در نظر گرفته شد. در ابتدا و با توجه به واژگان جست‌وجو شده، تعداد ۵۳۲ مقاله حاصل شد که تعداد ۱۲۶ مورد به دلیل تکراری بودن محتوا، تعداد ۳۲۶ مورد به دلیل عدم تناسب با موضوع اصلی پژوهش و ۲۰ مطالعه نیز به دلیل قدیمی بودن و تعداد ارجاعات پایین از مطالعه خارج شدند. بنابراین در نهایت تعداد ۶۰ مقاله برای بررسی انتخاب شد (شکل ۱). جمع‌آوری مطالب و جست‌وجوی مقالات در بازه زمانی خرداد تا شهریورماه سال ۱۳۹۶ شمسی در دانشگاه علوم پزشکی شیراز صورت گرفت.

موارد زیر است:

(۱) مقیاس لیکرت را می‌توان به آسانی ساخت و اصلاح کرد، (۲) نتایج اندازه‌گیری‌های عددی می‌تواند به‌طور مستقیم برای تجزیه و تحلیل آماری مورد استفاده قرار گیرد، (۳) اندازه‌گیری‌های مبتنی بر مقیاس لیکرت، اعتبار خوبی دارند، (۴) محققان با مقیاس لیکرت می‌توانند مقدار زیادی از داده‌ها را با صرف زمان و تلاش کمتر جمع‌آوری و مورد تجزیه و تحلیل قرار دهند [۱].

با وجود مزایای بالا، مقیاس لیکرت چندین ضعف نیز دارد که یکی از مشکلات عمده آن این است که مقیاس لیکرت، ماهیتی رتبه‌ای یا فاصله‌ای دارد [۲]، مقیاس لیکرت، فرمت بسته دارد [۴]، پاسخ‌دهندگان مجبور به انتخاب یکی از درجات پاسخ هستند که ممکن است پاسخ داده‌شده با نظر واقعی فرد متفاوت باشد و اینکه در مقیاس لیکرت مقدار قابل توجهی از اطلاعات از دست می‌روند یا تحریف می‌شوند که به دلیل محدودیت‌های موجود در روش طراحی مقیاس لیکرت است [۱]. مقیاس لیکرت حوزه‌های مختلفی از طراحی و تجزیه و تحلیل دارد و برای این مقیاس تاکنون دستورالعمل جامعی در رابطه با نحوه ساخت و آنالیز ارائه نشده است.

مطالعه حاضر با هدف شناخت صحیح نحوه ساخت و طراحی مقیاس لیکرت و همچنین تحلیل مسائل و چالش‌های موجود در نحوه تعیین پایایی و تجزیه و تحلیل داده‌های مقیاس لیکرت انجام شد.



شکل ۱) فلوچارت انتخاب مقالات براساس معیارهای ورود و خروج از مطالعه

یافته‌ها

مراحل ساخت مقیاس لیکرت

۱) تعیین موضوع مورد بررسی: اولین قدم در ساخت مقیاس لیکرت، تبیین ماهیت و چیستی موضوع مورد بررسی است. محقق باید در ابتدا این نکته را مد نظر داشته باشد که می‌خواهد چه چیزی را اندازه‌گیری کند تا براساس آن موضوع، گویه‌ها و گزینه‌های مناسب مطالعه را طراحی کند [۵].

۲) تعیین محورها یا مقوله‌های تشکیل‌دهنده موضوع

مورد بررسی: پس از انتخاب موضوع، با توجه به تعداد مقوله‌ها، تعداد متناسبی گویه مناسب استخراج می‌شود، به‌عنوان مثال در صورتی که بخواهیم سبک زندگی افراد را در یک جامعه بسنجیم، محورهای مختلف سبک زندگی از جمله داشتن فعالیت ورزشی روزانه، رعایت رژیم غذایی مناسب و پرهیز از سیگارکشیدن را مورد بررسی قرار می‌دهیم [۵].

۳) تدوین و ایجاد گویه‌های مرتبط با موضوع و

درجه‌بندی پاسخ به این گویه‌ها

۳-۱- تدوین عبارات نگرشی: پس از آنکه مقوله‌های

مورد بررسی تعیین شد، نوبت به طراحی عبارات نگرشی می‌رسد. برای تدوین عبارات نگرشی مناسب باید توجه نمود که سئوالات برای پاسخ‌دهندگان، با معنی، قابل درک و پاسخ باشد. به‌عنوان مثال، در موضوع سنجش نگرش افراد در رابطه با سبک زندگی و پس از تعیین مقوله‌های تشکیل‌دهنده آن مانند انجام فعالیت ورزشی روزانه، باید سعی شود سئوالات نگرشی کاربردی در این زمینه طراحی شود. به‌عنوان مثال "من از انجام فعالیت‌های ورزشی روزانه لذت می‌برم" [۶].

۳-۲- مشخص کردن تناسب و کارایی عبارات

نگرشی: در این مرحله می‌توان از انواع روش‌های روایی مانند روایی محتوایی و صوری و پایایی آزمون باز آزمون، به‌منظور شناسایی عبارات نگرشی مناسب و متناسب استفاده کرد، البته گاهی با توجه به برخی جنبه‌ها و ظواهر پاسخ به پرسش‌نامه نیز می‌توان به این عمل پی برد. به‌طور مثال اگر تعدادی از سئوالات را پاسخ نداده باشد یا پاسخ به یک گزینه خاص خیلی بیشتر از سایر گزینه‌ها باشد نشان از عدم تناسب در عبارات نگرشی دارد. به‌طور کلی هر چقدر که تعداد سئوالات بدون

پاسخ یا پاسخ "نمی‌دانم" کمتر باشد، نشان‌دهنده وضوح بیشتر عبارات است [۵].

۳-۳- تعداد درجات (گزینه پاسخ) مقیاس لیکرت:

اکثر محققانی که در زمینه مقیاس لیکرت به تحقیق مشغول هستند در مقیاس‌های خود بیشتر از درجه‌بندی‌های پنج‌گانه یا هفت‌گانه استفاده می‌کنند [۷]. این مقیاس از پاسخ‌دهنده می‌خواهد تا سطح توافق خود را با نگرش مورد بررسی نشان دهد. بر حسب اینکه چه مقوله‌ای قرار است اندازه‌گیری شود و همچنین نوع جامعه مورد سنجش و موضوع مورد بررسی، تعداد و نوع درجه‌بندی نیز متفاوت خواهد بود. به‌عنوان مثال در صورتی که بخواهیم توافق افراد را در مورد نگرش مورد بررسی جویا شویم بهتر است از درجه‌بندی پنج‌گانه به‌ترتیب کاملاً مخالفم، مخالفم، نه مخالفم و نه موافقم (نظری ندارم یا تصمیمی نگرفتم)، موافقم، کاملاً موافقم عمل کنیم.

در صورتی که هدف اندازه‌گیری تکرار یا تناوب باشد می‌توان از درجاتی مانند هرگز- همیشه استفاده کرد یا هنگامی که هدف، اندازه‌گیری باور و خصوصیات شخصیتی پاسخ‌دهنده است، درجاتی مانند "نه خیلی زیاد- خیلی زیاد" کاربرد دارند [۱]. به‌طور معمول و همان‌طور که به آن اشاره شد در مقیاس لیکرت استفاده از ۵ یا ۷ درجه رایج است با این وجود عده‌ای از محققان در مطالعات خود فقط از دو گزینه استفاده می‌کنند. تحقیقات نتوانسته توجیه واقعی برای این موضوع پیدا کند که کدام یک بر دیگری برتری دارد. برخی از مطالعات که از مقیاس ۴ درجه‌ای به‌جای مقیاس ۵ درجه‌ای استفاده کرده‌اند در واقع گزینه حد واسط را حذف کرده‌اند و پاسخ‌دهنده مجبور است یکی از گزینه‌های کاملاً موافقم تا کاملاً مخالفم را انتخاب نماید [۸]. رنسیس لیکرت در مطالعه اصلی خود تعداد درجات مربوط به هر سؤال را به‌عنوان یک مساله مهم در نظر نگرفته بود. اگر از ۵ درجه استفاده شود لازم است که درجه سوم به "نظری ندارم یا تصمیمی نگرفتم" اختصاص داده شود. اینکه تعداد درجات در هر مطالعه چقدر باشد به‌صورت ضمنی تعیین می‌شود و در عمل محققان اغلب تعداد گزینه‌ها را به‌صورت اختیاری با توجه به سلیقه شخصی یا پیمان‌نامه قبلی انتخاب می‌کنند [۹]. در جدول ۱ مثال‌هایی از انواع درجه‌بندی‌های پنج‌گانه و هفت‌گانه در مقیاس لیکرت ذکر شده است.

جدول ۱) مثال‌هایی از انواع درجه‌بندی‌های پنج‌گانه و هفت‌گانه در مقیاس لیکرت

خیلی مخالفم (۱)	مخالفم (۲)	تصمیمی نگرفتم (۳)	موافقم (۴)	خیلی موافقم (۵)	-	-
هرگز (۱)	به‌ندرت (۲)	هر چند وقت یک‌بار (۳)	گاهی (۴)	همیشه (۵)	-	-
ضعیف (۱)	نسبی (۲)	متوسط (۳)	خوب (۴)	عالی (۵)	-	-
خیلی مخالفم (۱)	مخالفم (۲)	تا حدودی مخالفم (۳)	تصمیمی نگرفتم (۴)	تا حدودی موافقم (۵)	موافقم (۶)	خیلی موافقم (۷)
هرگز (۱)	به‌ندرت (۲)	گاهی (۳)	اغلب (۴)	مکرر (۵)	تقریباً همیشه (۶)	همیشه (۷)

مثال راسل در سال ۱۹۹۲ پیشنهاد کرد که اگر درجات مقیاس لیکرت افزایش یابد و به یکدیگر نزدیک‌تر شوند تا مقیاس

بسیاری از محققان اعتقاد دارند که هر چه تعداد درجه‌بندی‌ها بیشتر شود، سنجش دقیق‌تر خواهد شد، به‌عنوان

۵-۳- ترکیب پرسش‌ها (گویه‌ها): مقیاس لیکرت باید به‌نحوی طراحی شود که حاوی پرسش‌های مساعد و نامساعد (پرسش‌های مستقیم و معکوس) باشد و بهتر است که تعداد پرسش‌های مساعد و نامساعد به‌طور تقریبی با هم برابر باشند، در غیر این صورت، گرایش افراد به طرف فرضیه‌های تحقیق بیشتر می‌شود (مثلاً پنج پرسش موافق و یک پرسش مخالف) [۱۷، ۱۸]. نکته قابل توجه در نمرات درجه‌بندی پرسش‌های مساعد و نامساعد این است که اگر به‌طور مثال برای پرسش مساعد، گزینه خیلی موافقم نمره بالاتری می‌گیرد، در پرسش نامساعد گزینه خیلی مخالفم نمره بالاتری خواهد گرفت. به‌طور مثال، در پرسش مساعد "انجام فعالیت‌های ورزشی روزانه، لذت‌بخش است" به احتمال خیلی زیاد گزینه انتخابی خیلی موافقم است و در پرسش نامساعد "انجام فعالیت‌های ورزشی دردناک است" گزینه احتمالی خیلی مخالفم است [۱۹]. در صورت استفاده از چند سؤال به‌جای یک سؤال برای اندازه‌گیری ابعاد خاصی، پایایی و روایی پاسخ‌ها به مقدار قابل توجهی افزایش می‌یابد، به همین دلیل اکثر محققین در نگرش‌سنجی‌های خود برای اندازه‌گیری هر مقوله خاص، تعدادی سؤال طراحی می‌کنند و گروه‌بندی سؤالات در ارتباط با ابعاد مختلف را تحلیل عاملی می‌نامند. نکته قابل توجه آن است که پس از تدوین سؤالات، بهتر است که به‌صورت کاملاً تصادفی و نامنظم در پرسش‌نامه نگرش‌سنجی گنجانده شود تا از رخداد خطاهایی مانند خطای توزیع (گرایش زیاد به موافقت، مخالفت یا حد واسط) و القای فرضیه‌ها به پاسخ‌دهندگان جلوگیری شود [۵]. فریدمن در سال ۱۹۹۸، در مطالعه خود نشان داد در صورتی که فقط از پرسش‌های مساعد در طراحی مقیاس لیکرت استفاده شود، نتایج به‌دست‌آمده با زمانی که از پرسش‌های نامساعد استفاده می‌شود، متفاوت خواهد بود.

۶-۳- اصول در طراحی گویه‌ها

۱-۶-۳- اجتناب از طراحی سؤالات جهت‌دار و گرایش: سؤالات طراحی‌شده باید کاملاً بی‌طرفانه مطرح شود تا سبب القای پاسخ یا گرایش به پاسخ خاصی نشود. به‌عنوان مثال "تمام داروهایی که امروزه استفاده می‌شوند، عوارض جانبی زیادی دارند". این گونه سؤالات، افراد را به سمت انتخاب گزینه‌های مشابه سوق می‌دهد [۲۰].

۲-۶-۳- دوری از سؤالاتی که منجر به پاسخ‌های کلیشه‌ای می‌شوند: چنانچه اکثر پاسخ‌دهندگان یک گزینه کاملاً موافقم یا کاملاً مخالفم را انتخاب کرده باشند، می‌توان ادعا کرد که پاسخ‌ها به‌صورت کلیشه‌ای آرایه شده است که این کلیشه‌سازی‌ها می‌تواند با توجه به فرهنگ‌های مختلف کاملاً متفاوت باشند. یکی از راه‌های مقابله با این گونه خطاهای ادراکی، طراحی سؤالات کنترلی (به‌صورت مستقیم یا معکوس) است، بدین معنی که یک یا چند سؤالات مرتبط با سؤال اصلی طراحی شود تا میزان صحت پاسخگویی فرد مشخص شود.

لیکرت بتواند اندازه‌گیری‌های پی‌درپی را تقریب بزند که این مساله باعث می‌شود اطلاعات بیشتری کسب شود [۱۰]. با این حال تعدادی از محققان اعتقاد دارند که افزایش در تعداد درجات مقیاس منجر به تقویت سنجش و اندازه‌گیری نمی‌شود و حتی ممکن است تعداد زیادی از پاسخ‌دهندگان نتوانند به‌خوبی فاصله بین درجات را از هم تشخیص دهند [۵]. چانگ در سال ۱۹۹۴ از طریق مقایسه لیکرت‌های ۴ و ۶ درجه‌ای متوجه شد که تعداد درجات بیشتر در مقیاس در واقع خطای اندازه‌گیری را افزایش می‌دهد و موجب گیج‌شدن پاسخ‌دهندگان می‌شود، همچنین وی به این نکته نیز اشاره کرد که درجات پاسخ طولانی‌تر ممکن است تنبلی در پاسخ به پرسش‌نامه را تشدید کند [۱۱]. بنابراین روش افزایش گزینه‌ها می‌تواند چندین اثر تقدم از جمله اثر ترتیب- پاسخ (گرایش به سمت اولین گزینه موجود بین جواب‌ها توسط پاسخ‌دهنده) [۱۲]. اثر رای ترجیحی (انتخاب پاسخ مشابه برای همه سؤالات) [۱۳] و اثر گرایش مرکزی (انتخاب پاسخ خنثی یا حد واسط) را افزایش دهد [۱۴].

در سال ۱۹۹۷، آلبیوم یک مقیاس لیکرت دومرحله‌ای را پیشنهاد کرد که جایگزین مقیاس قدیمی شد. در این مقیاس اولین مرحله شامل اندازه‌گیری توافق (موافق/مخالف) با یک نگرش بود، مرحله دوم شامل اندازه‌گیری شدت موافقت (خیلی یا ضعیف) با نگرش بود. به نظر می‌رسد که مقیاس لیکرت دومرحله‌ای می‌تواند اطلاعات اضافی بیشتری را نسبت به مقیاس قدیمی‌تر برای محققان فراهم نماید. در یک نگاه، این طراحی برای کاهش اثر گرایش مرکزی موثر است با این حال هنوز مشخص نشده است که این روش چگونه می‌تواند اطلاعات بیشتری را نسبت به روش قدیمی جمع‌آوری کند [۱۵].

۴-۳- نیاز به درجه حد واسط در میان درجات پاسخ:

اغلب به سه دلیل، استفاده از حد واسط ضروری به نظر می‌آید: اول، گماردن درجه حد واسط در میان پاسخ‌ها باعث می‌شود که افرادی که نسبت به آن نگرش خاص بی‌طرف هستند، بی‌طرفی خود را به دقت و آن گونه که دوست دارند بیان کنند و در صورتی که حد واسط وجود نداشته باشد، تمام پاسخ‌دهندگان بی‌طرف مجبور به انتخاب گزینه‌ای می‌شوند که نظرات واقعی آنها را نشان نمی‌دهد. دوم، پاسخ‌دهندگان از اینکه مجبور به انتخاب گزینه‌ای هستند که در واقع نسبت به آن موضوع کاملاً خنثی هستند، موجب می‌شود که در پاسخ به سؤالات سرخورده و دل‌سرد شوند و حتی شاید سؤال را بدون پاسخ رها کنند. سوم، هنگام تجزیه و تحلیل پاسخ‌های داده‌شده به پرسش‌نامه، برخی از شاخص‌های گرایش مرکزی مانند میانه و میانگین می‌توانند دارای ارزشی معادل نقطه حد واسط معیار شوند. برخی از صاحب‌نظران اعتقاد دارند که برای مقیاس‌هایی که پس از آن اقدامات خاصی صورت می‌گیرد، حد واسط کاربردی ندارد [۱۶].

پاسخ‌دهندگان طراحی شود و باید از به‌کاربردن کلمات نامفهوم و ناآشنا برای پاسخ‌دهندگان جلوگیری کرد [۹].

۳-۶-۹- اجتناب از منفی مضاعف: استفاده از این اصطلاحات موجب اشتباهات زیادی در درک سؤال می‌شود و فرد با یک جمله گنگ روبه‌رو می‌شود، به‌عنوان مثال من از فکر اجازه‌ندادن به کارکنان برای کار در نوبت‌های پشت سر هم حمایت نمی‌کنم. این گونه پرسش‌ها را با توجه به اینکه منجر به گیج‌شدن پاسخ‌دهنده می‌شود را باید حذف یا آنها را اصلاح کرد، یا آنالیز چنین سئوالاتی به‌صورت جداگانه انجام شود [۸].

۳-۶-۱۰- اجتناب از درجه‌بندی اعدادی که درجه‌بندی (۱ تا ۵) مقیاس را در پرسش‌نامه نشان می‌دهد: استفاده از این اعداد در پرسش‌نامه موجب گرایش به سمت خطا می‌شود، زیرا اکثر پاسخ‌دهندگان به انتخاب نمرات ابتدایی و انتهایی گرایش دارند و باید نمره‌بندی درجه‌ها را به بعد از انجام مطالعه موکول کرد. پیشنهاد می‌شود برای راحتی و تشخیص بهتر بین گویه و درجات پاسخ، گویه در یک ستون و درجات پاسخ نیز در ستون‌های مجاور آن نوشته شود [۲۰]. به‌طور مثال گویه "انجام فعالیت‌های ورزشی روزانه، لذت‌بخش است" در یک ستون مجزا و درجات پاسخ شامل "خیلی مخالفم"، "مخالفم"، "نظری ندارم"، "موافقم" و "خیلی موافقم" در ستون‌های مجاور قرار گیرند.

۴- تحلیل سئوالات و تعیین کردن سئوالات برگزیده: با توجه به نکات ذکرشده در طراحی سئوالات، در این مرحله باید سئوالات مناسب و پایدار را انتخاب نمود، منظور از پایدار بودن در مقیاس لیکرت، میزان هماهنگی و همبستگی هر یک از سئوالات با سایر سئوالات است که این روش را پایداری درونی نیز می‌نامند.

۵- محاسبه نمره کل برای هر پاسخ‌دهنده: با توجه به تعداد درجات مقیاس، می‌توان طیف مزبور را شماره‌گذاری کرد، به‌طور مثال برای یک مقیاس ۵ درجه‌ای می‌توان از اعداد ۱ تا ۵ یا صفر تا ۴ استفاده نمود که اغلب از درجه‌بندی یک تا ۵ استفاده می‌کنند. نکته قابل توجه آن است که باید توجه داشت که به کدام درجه (کاملاً موافقم یا کاملاً مخالفم) اولین عدد اختصاص داده می‌شود [۵].

مقیاس لیکرت در میان کودکان: اخیراً مقیاس لیکرت در تعداد زیادی از پروژه‌های تحقیقاتی و شرایط بالینی با تاکید بیشتر در کودکان مورد استفاده قرار می‌گیرد، بعضی از نمونه‌های آن با توجه به سن کودکان و تعداد گزینه‌های مقیاس لیکرت شامل تاثیر کودکان از آسیب‌های روانی که توسط ولف بازنگری شد و مقیاس آن سه‌گزینه‌ای (بسیار درست است، تا حدودی درست است، درست نیست) بود [۲۳] که شباهت زیادی با پرسش‌نامه گودمن داشت [۲۴]. مقیاس نگرانی‌های اجتماعی در کودکان که بر ۵ گزینه مبتنی بود (یک=هیچوقت و پنج=همه مواقع) [۲۳]. در سال ۲۰۱۶، بارتولومو در مطالعه

به‌طور مثال وقتی از یک مربی فوتبال در مورد علل شکست تیمش سؤال شود، وی همه تقصیرها را گردن داور یا زمین مسابقه می‌اندازد، در صورتی که با طراحی چند سؤال کنترلی می‌توان به نقطه ضعف مربی‌گری وی پی برد [۲۱].

۳-۶-۳- طراحی سئوالات تک‌موضوعی: یکی از رایج‌ترین اشتباهات در طراحی سئوالات نگرشی آن است که سؤال بیش از یک موضوع را می‌سنجد و قسمت اول و دوم کاملاً با هم متفاوت هستند. در این گونه موارد امکان دارد پاسخ‌دهنده با قسمت اول سؤال موافق بوده ولی با قسمت دوم مخالف باشد و انتخاب گزینه مناسب را برای وی سخت می‌کند. به‌طور مثال، من از حقوق دریافتی در سازمان رضایت دارم و سازمان تمام امکانات رفاهی را برای من مهیا کرده است [۸].

۳-۶-۴- اجتناب از طراحی سئوالات حساس: با توجه به مسائلی مانند قوم و فرهنگ، میزان اطلاعاتی که در اختیار محقق قرار داده می‌شود می‌تواند متفاوت باشد. به‌طور مثال برخی از افراد در مورد پاسخ به سئوالاتی که در مورد حقوق، نژاد، مسائل قومی و یا گرایش‌های سیاسی باشد آکراه دارند. یکی از راه‌های کشف چنین سئوالاتی انجام مطالعه پایلوت در مقیاس کوچک روی گروهی که تحت نگرش‌سنجی قرار خواهند گرفت، است [۲۲].

۳-۶-۵- اجتناب از سئوالات شرم‌آور: هنگامی که در مورد مسائلی بسیار خصوصی یا نگرش‌های خلاف عرف سؤال پرسیده شود، فرد پاسخ‌دهنده احساس شرم می‌کند که درجه حساسیت فرد را افزایش می‌دهد. برخی از کارشناسان معتقد هستند که بهتر است که این گونه سئوالات در انتهای پرسش‌نامه قرار داده شود [۵].

۳-۶-۶- پرسیدن سئوالاتی که پاسخ به آنها در نتایج مد نظر تاثیر داشته و با موضوع مورد بررسی مرتبط باشد و پرهیز از پرسیدن سئوالات بیهوده: یکی از اصول طراحی سئوالات مقیاس لیکرت آن است که محقق از خود بپرسد که چه چیزی را می‌خواهد بداند، نه اینکه پرسش‌شونده چه چیزی را می‌خواهد بداند. به‌عنوان مثال، در صورتی که هدف از انجام مطالعه‌ای، ارزیابی نگرش بیماران دیابتی نسبت به تزریق انسولین باشد، به نظر می‌رسد، پرسیدن سئوالات نگرشی در رابطه با سایر درمان‌های دیابت، محقق را در دستیابی به نتایج مد نظر، گمراه می‌کند.

۳-۶-۷- اجتناب از پرسش‌های مبهم یا پیچیده: در هنگام نگارش سئوالات همواره به این نکته توجه کرد که سئوالات برای همه پاسخ‌دهندگان یک مفهوم را برساند. به‌عنوان مثال، وقتی از فرد سؤال می‌شود که "آیا شما مایلید که امتحان بدهید؟" پرسشی مبهم است، زیرا پاسخ‌دهنده به‌طور دقیق منظور از امتحان دادن را نمی‌داند.

۳-۶-۸- استفاده از زبان مناسب: کلمات و عبارات به‌کاررفته در سئوالات باید متناسب با سطح دانش

هستند. پایایی نمره آزمون اغلب به عنوان قابلیت اعتماد، ثبات و پایداری از نمره تولید شده به وسیله یک ابزار خاص توصیف می شود. یکی از روش های محاسبه پایایی در مقیاس لیکرت، استفاده از روش آزمون باز آزمون است. بدین صورت که پرسش نامه در اختیار چند پاسخ دهنده قرار داده می شود و سپس همین پرسش نامه را در اختیار همان افراد به فاصله مشخص (مثلا دو هفته) قرار می دهند و در نهایت همبستگی بین این اندازه گیری ها را محاسبه می کنند. این روش دو مشکل احتمالی دارد، اول آنکه اگر فاصله بین دو اجرای اندازه گیری خیلی کوتاه باشد، ضریب همبستگی ممکن است به علت تاثیر یادآوری به طور کاذب بالاتر رود، زیرا در اجرای دوم اندازه گیری، پاسخ دهندگان به یاد می آورند که در اجرای اول اندازه گیری کدام گزینه ها را انتخاب کرده اند. به همین ترتیب اگر فاصله بین دو اندازه گیری خیلی طولانی شود، به دلیل تغییر در بخشی از پاسخ دهندگان، ضریب همبستگی محاسبه شده کم می شود. نکته قابل ذکر در مورد این نوع روش محاسبه ضریب پایایی آن است که زمان بین دو اندازه گیری باید حتماً گزارش شود.

برآورد انسجام درونی نمرات جمع شده حاصل از یک مقیاس لیکرت، فقط به یکبار اندازه گیری ابزار نیاز دارد. انسجام درونی نشان دهنده این است که تا چه حد بین پاسخ به سئوالات مختلف در مقیاس لیکرت ارتباط یا وابستگی وجود دارد. کرونباخ این برآورد پایایی را توسعه داد و آن را ضریب آلفا نامید. تعریف ریاضی آلفای کرونباخ در معادله زیر شرح داده شده است:

$$\text{معادله ۴)} \quad \text{واریانس (خطا)} = \frac{n}{n-1} \left(1 - \frac{\text{واریانس (نمرات مشاهده شده)}}{\text{واریانس (خطا)}} \right) \quad \text{آلفای کرونباخ}$$

معادله ۵ برای محاسبه آلفای کرونباخ نشان می دهد که مقدار برآورد انسجام درونی نمرات جمع شده به وسیله میانگین همبستگی های بین سئوالات، در پاسخ به یک سؤال در مقیاس لیکرت و تعداد سئوالات مقیاس تعیین می شود. در معادله ۵، n برابر با تعداد سئوالات مقیاس لیکرت و r میانگین همبستگی بین سئوالات است.

$$\text{معادله ۵)} \quad \text{آلفای کرونباخ} = \frac{nr}{1+r(n-1)}$$

هر چه میانگین همبستگی بین سئوالات و تعداد سئوالات بیشتر باشد، آلفای کرونباخ نیز افزایش می یابد. محققان بر این باور هستند که پایایی مقیاس لیکرت با ضریبی حدود $0/85$ قابل قبول است [۲۶]. مورفی و لیکرت نشان دادند که وقتی تعداد سئوالات از ۴۴ به ۳۶ مورد کاهش می یابد و برعکس، تعداد درجات از ۳ به ۵ مورد افزایش می یابد، به دلیل اینکه تعداد کل درجات افزایش می یابد، یعنی از ۱۳۲ مورد (۴۴ تا ۳ تا) به ۱۸۰ مورد (۳۶ تا ۵ تا) پایایی نیز از $0/88$ به $0/94$ افزایش می یابد [۹].

محاسبه توان افتراقی یا ضریب تمیز: توان افتراقی به قدرت سؤال در تشخیص تفاوت بین دو گروه از افراد که در

خود مهارت های اجتماعی را در کودکان سنجید و در نهایت به این نتیجه رسید که لیکرت سه گزینه ای نسبت به چهار گزینه ای بهتر است، بنابراین مقیاس سه گزینه ای یک روش بهتر برای ارزیابی مهارت های اجتماعی کودکان در نظر گرفته شد زیرا آنها می توانستند ادراکات خود را در این ساختار بهتر نشان دهند [۲۵].

تعیین ضریب پایایی مقیاس لیکرت: پایایی، دقت اندازه گیری را توصیف می کند، در نتیجه با توجه به نظریه آزمون کلاسیک که تصور می کند خطای تصادفی همیشه با اندازه گیری مرتبط است و ساختارهای روان شناختی و اجتماعی را تعیین می کند، نمره واقعی فرد در پرسش نامه نگرش سنجی به صورت زیر است:

معادله ۱) نمره واقعی فرد = نمره مشاهده شده - خطاهای تصادفی از طرفی دیگر و با توجه به اینکه یک گروه از افراد یک ابزار خاص را تکمیل می کنند، واریانس نمرات واقعی برای گروه به صورت زیر است:

معادله ۲) واریانس نمرات واقعی = واریانس نمره مشاهده شده - واریانس خطا

هنگامی که ساختارهای نگرشی و ادراکی با استفاده از مقیاس لیکرت اندازه گیری می شوند، نمره مشاهده شده توسط فرد، برابر است با مجموع نمرات جمع شده در سئوالات مختلف که نشان دهنده واکنش فرد به سئوالات و عبارات نگرشی در مقیاس لیکرت بوده است، در معادله ۲، واریانس نمرات مشاهده شده برای گروهی از افراد که به مقیاس لیکرت پاسخ داده اند، محاسبه می شود. واریانس خطاهای اندازه گیری که به صورت تصادفی فرض شده اند از تغییرات بین افراد به وسیله پاسخ هایشان به هر یک از سئوالات در مقیاس لیکرت برآورد می شود. پایایی به عنوان یک ضریب بیان می شود که نسبتی از واریانس نمرات واقعی به واریانس نمرات مشاهده شده است.

از آنجایی که معادله ۲، واریانس نمرات واقعی را به عنوان واریانس نمرات مشاهده شده منهای واریانس خطاهای تصادفی اندازه گیری تعریف می کند، معادله محاسبه ضریب پایایی به صورت زیر می شود:

$$\text{معادله ۳)} \quad \text{واریانس (خطا)} = \frac{\text{واریانس (نمرات مشاهده شده)}}{\text{واریانس (نمرات مشاهده شده)}} = \text{ضریب پایایی}$$

ضریب پایایی محاسبه شده یک تخمین و برآورد است به دلیل آنکه واریانس خطاهای تصادفی در معادله به صورت تخمینی برآورد شده است. زمانی که ساختار به وسیله مقیاس لیکرت اندازه گیری می شوند، از نظر آماری، ضریب پایایی یک برآورد از نسبت واریانس نمرات مشاهده شده است که به خطاهای تصادفی قابل استناد نیست.

مقادیر ضریب پایایی محاسبه شده از صفر تا یک متغیر است. مقادیر نزدیک به یک نشان دهنده این است که مجموع نمرات مشاهده شده به طور نسبی خالی از خطاهای تصادفی

محقق در مورد سطح اندازه‌گیری داده‌های خود بستگی دارد، اینکه آنها را به‌عنوان مقیاس رتبه‌ای یا فاصله‌ای در نظر بگیرد [۳۱]. گاردنر و مارتین در سال ۲۰۰۷ و جیمسون در سال ۲۰۰۴، ادعا کردند که داده‌های مقیاس لیکرت، ماهیتی رتبه‌ای دارند، بنابراین تنها آزمون‌های ناپارامتریک برای تجزیه و تحلیل آنها کاربرد دارد [۲۸]. با این حال، در سال ۲۰۰۸، کرسول پیشنهاد کرد که با داده‌های لیکرت به‌عنوان مقیاس فاصله‌ای رفتار شود که در اینجا نیاز به دسته‌های چندگانه سؤال در مقیاس، فرض برابری واریانس بین مقادیر و توزیع نرمال داده‌ها وجود دارد [۲۸].

میانگین و انحراف‌معیار برای آنالیز داده‌های ترتیبی نامناسب هستند و باید از میانه یا مد برای اندازه‌گیری در مقیاس لیکرت استفاده کرد. مطالعات استاندارد نیز توصیه کرده‌اند که آمار استنتاجی مناسب برای داده‌های رتبه‌ای، استفاده از تست‌های ناپارامتریک مانند ضریب همبستگی اسپیرمن یا آزمون من‌ویتنی است، به این دلیل که برای آزمون‌های پارامتریک لازم است داده‌ها از نوع فاصله‌ای یا نسبتی باشند [۳]. با این حال، این قانون توسط برخی از نویسندگان در نظر گرفته نشده است. به‌عنوان مثال ساتینگ و هرن در مقالات خود برای توصیف داده‌ها از میانگین و انحراف‌معیار و برای آنالیز داده‌ها از روش‌های پارامتریک مثل آنالیز واریانس استفاده کردند [۳۲]. کوزان و همکاران استدلال کردند که استفاده از آنالیزهای پارامتریک برای داده‌های رتبه‌ای، گناهی نابخشودنی در آنالیزهای آماری به شمار می‌رود [۳۳]. با این حال نپ اعتقاد داشت که حجم نمونه و توزیع داده‌ها نسبت به اینکه از آنالیزهای پارامتریک یا ناپارامتریک استفاده شود، ارجحیت بیشتری دارد [۳۴].

در سال ۲۰۰۴، لوکاج و همکاران در مطالعه خود برای آنالیز داده‌های مقیاس لیکرت از میانگین و انحراف‌معیار و همچنین ضریب همبستگی اسپیرمن استفاده کرد [۳۵]. در سال ۲۰۱۲، د وینتر برای تجزیه و تحلیل داده‌های لیکرت از میانگین و انحراف‌معیار استفاده کرد و به‌منظور بررسی ارتباط آزمون t مستقل و من‌ویتنی-ویلکاکسون را به کار برد و در نهایت به این نتیجه دست یافت که بین مقادیر آزمون t مستقل و من‌ویتنی-ویلکاکسون تفاوت اندکی وجود دارد [۳۶]. مورای در سال ۲۰۱۳، برای داده‌های لیکرت مطالعه خود از روش ضریب همبستگی پیرسون، اسپیرمن و کندال برای آنالیز استفاده نمود و در آخر به این نتیجه دست یافت که مقادیر به‌دست‌آمده از هر سه روش فوق، با هم تفاوت معنی‌داری ندارند [۲۸].

در یک جمع‌بندی کلی، برای مقایسه دو نمونه مستقل در مقیاس لیکرت، در صورتی که فرض نرمال (پارامتری) بودن، برقرار نباشد از آزمون من‌ویتنی استفاده می‌شود (جدول ۳) [۳۷].

مورد یک موضوع خاص، نگرش‌های متفاوتی دارند، اطلاق می‌شود. با توجه به اینکه پاسخ‌های هر سؤال بر حسب نمرات کل پاسخ‌دهندگان مرتب شده‌اند، ۲۵٪ افرادی که کمترین نمره را کسب کرده‌اند را به‌عنوان گروه پایین و ۲۵٪ افرادی که بیشترین نمره را به دست آورده‌اند، به‌عنوان، گروه بالا در نظر گرفته می‌شوند. جدول ۲ نحوه محاسبه توان افتراقی هر یک از سئوالات را در یک نمونه ۹ نفره نشان می‌دهد.

جدول ۲) محاسبه توان افتراقی هر یک از سئوالات در یک نمونه ۹ نفری

گروه	۲۵٪ به بالا	۲۵٪ به پایین
تعداد افراد در گروه	۹	۹
فراوانی گزینه‌ها		
۱	۰	۱
۲	۱	۸
۳	۲	۰
۴	۳	۰
۵	۳	۰
وزن کلی سئوالات	۳۵	۱۷
میانگین وزنی سئوالات	۳/۸۹	۱/۸۹
توان افتراقی	۲	

در جدول ۲ محاسبه وزن کلی سئوالات برای گروه "۲۵٪ بالا" به‌صورت زیر است:

$$\text{وزن کلی داده‌ها} = \text{نمره} \times \text{تعداد گزینه‌های علامت‌زده}$$

$$۳۵ = (۱ \times ۰) + (۲ \times ۱) + (۳ \times ۲) + (۴ \times ۳) + (۵ \times ۳)$$

توان افتراقی نیز از تقسیم میانگین وزنی سئوالات در گروه ۲۵٪ بالا به میانگین وزنی سئوالات در گروه ۲۵٪ پایین به دست می‌آید [۲۷].

تجزیه و تحلیل داده‌های مقیاس لیکرت: بسیاری از محققان و سازمان‌های تجاری و آموزشی همیشه با یک تصمیم‌گیری در رابطه با مقیاس لیکرت مواجه هستند و آن، این است که برای تجزیه و تحلیل داده‌های جمع‌شده مقیاس لیکرت از چه نوع آنالیز آماری باید استفاده کرد. اغلب یک سطح از عدم تصمیم‌گیری در مورد آنالیز مربوطه وجود دارد به‌ویژه اینکه باید از آنالیزهای پارامتریک یا ناپارامتریک استفاده کرد [۲۸]. ویچدراوس معتقد بود که مفروضات مربوط به سطح اندازه‌گیری و آنالیزهای به‌کاررفته در مورد آن داده‌ها، بر نتیجه‌گیری از آن مطالعه تاثیر می‌گذارد [۲۹]. در سال ۲۰۱۰، نورمن پیشنهاد کرد که برای آنالیز داده‌های لیکرت می‌توان از آزمون‌های پارامتریک استفاده کرد بدون اینکه از دستیابی به نتیجه اشتباه ترسی وجود داشته باشد [۳۰]. در سال ۲۰۰۸، پرلا و کاریفیو به‌منظور برطرف کردن بحث و مشاجره در مورد آنالیز مربوط به مقیاس لیکرت، پیشنهاد دادند که استفاده یا عدم استفاده از آزمون‌های پارامتریک و ناپارامتریک به دیدگاه

مقیاس رتبه‌ای، آنالیزهای ناپارامتریک و برای داده‌هایی با مقیاس فاصله‌ای در صورت برقراری فرض نرمال بودن داده‌ها، انجام آزمون پارامتریک پیشنهاد می‌شود. در رابطه با طراحی مقیاس لیکرت برای گروه کودکان نیز، تیم تحقیق پیشنهاد می‌دهد که طراحی سوالات سه‌گزینه‌ای، در تحقق اهداف مطالعه کمک بیشتری می‌کند.

پیشنهاد می‌شود به منظور ارزیابی پایایی داده‌های مقیاس لیکرت از روش آلفای کرونباخ استفاده شود، زیرا نتایج دقیق‌تری را نسبت به دیگر انواع آن ارائه می‌دهد.

دسترسی محدود و عدم دریافت مقالات تمام متن در برخی پایگاه‌های اطلاعاتی از جمله ویلی و اسکوپوس در زمان اجرای پژوهش، موجب شد جست‌وجو فقط در سه پایگاه اطلاعاتی صورت گیرد که از جمله محدودیت‌های مطالعه حاضر به شمار می‌رود.

نتیجه‌گیری

در اکثر مطالعات مرتبط با مقیاس لیکرت که تاکنون انجام شده است، نوعی سردرگمی یا عدم توافق در نحوه طراحی، ساخت و تجزیه و تحلیل داده‌های آن وجود دارد. در نهایت با وجود آنکه تاکنون پیمان‌نامه دقیق و جامعی در مورد طراحی و تجزیه و تحلیل داده‌های لیکرت وجود ندارد، این مطالعه با ارائه جمع‌بندی کلی در این زمینه، تا حدود زیادی از سردرگمی‌ها کاست.

تشکر و قدردانی

نویسندگان از دانشگاه علوم پزشکی شیراز، برای حمایت از این طرح تشکر و قدردانی می‌کنند.

تاییدیه اخلاقی: پژوهش حاضر، حاصل بخشی از پایان‌نامه آقای حامد دلام، دانشجوی مقطع ارشد رشته اپیدمیولوژی در دانشگاه علوم پزشکی شیراز با کد ۹۵/۱۲۴۴۸ بود. این مطالعه با کد IR.SUMS.REC.1395.S1084 توسط کمیته اخلاق مورد تایید قرار گرفت.

تعارض منافع: نویسندگان اعلام می‌دارند هیچ‌گونه تعارض منافی وجود ندارد.

سهام نویسندگان: علیرضا میراحمدی زاده (نویسنده اول)، روش‌شناس/پژوهشگر اصلی (۳۰٪)؛ حامد دلام (نویسنده دوم)، نگارنده مقدمه/روش‌شناس/پژوهشگر اصلی/نگارنده بحث (۶۰٪)؛ مژگان سیف (نویسنده سوم)، تحلیلگر آماری (۵٪)؛ ریحانه بهرامی (نویسنده چهارم)، پژوهشگر کمکی (۵٪)

منابع مالی: منابع مالی پژوهش حاضر توسط دانشگاه علوم پزشکی شیراز تامین شده است.

برای مقایسه بیش از دو گروه، در صورت برقراری فرض نرمال بودن داده‌ها، از روش آنالیز واریانس استفاده می‌شود (جدول ۴) [۳۷].

جدول ۳) انتخاب آزمون مناسب برای مقایسه بین دو گروه

مقیاس اندازه‌گیری	نمونه‌های مستقل	نمونه‌های زوجی
مقیاس فاصله‌ای (برقراری فرض پارامتری)	آزمون t مستقل	آزمون t زوجی
مقیاس رتبه‌ای یا فاصله‌ای (عدم برقراری فرض پارامتری)	آزمون من‌ویتنی	آزمون ویلکاکسون

جدول ۴) انتخاب آزمون (روش) مناسب برای مقایسه بین بیش از دو گروه

مقیاس اندازه‌گیری	نمونه‌های مستقل	نمونه‌های زوجی
مقیاس فاصله‌ای (برقراری فرض پارامتری)	آنالیز واریانس یک‌طرفه	آنالیز واریانس اندازه‌گیری‌های مکرر
مقیاس رتبه‌ای یا فاصله‌ای (عدم برقراری فرض پارامتری)	آزمون کروسکال-والیس	آزمون فریدمن

بحث

مطالعه حاضر با هدف رفع برخی از سردرگمی‌های مرتبط با طراحی و ساخت و تجزیه و تحلیل داده‌های لیکرت طراحی شد و در نهایت به این نتیجه دست یافت که بهترین روش برای سنجش پایایی داده‌های لیکرت آن است که از روش آلفای کرونباخ استفاده شود. درک این مطلب که داده‌های مقیاس لیکرت ماهیتی رتبه‌ای دارند یا فاصله‌ای، به‌نحوه تجزیه و تحلیل داده‌ها کمک می‌کند (اینکه از آنالیزهای پارامتریک یا ناپارامتریک استفاده شود). بحث در مورد اینکه در آنالیز داده‌های لیکرت از چه روش یا آنالیز آماری استفاده شود، همچنان وجود دارد ولی این مطالعه با بررسی مطالعات و روش‌های مختلف مورد استفاده در آن، یک جمع‌بندی کلی از نحوه برخورد با داده‌های مقیاس لیکرت را ارائه کرد.

تاکنون محققان زیادی در زمینه مقیاس لیکرت، فعالیت کرده‌اند که هر کدام از آنها برای داده‌های خود، ترکیب‌ها و آنالیزهای متفاوتی را به کار برده‌اند، در مطالعه حاضر با مقایسه مطالعات مشابه به این نتیجه دست یافته شد که به‌منظور طراحی مقیاس لیکرت با توجه به نوع مطالعه و هدف از انجام مطالعه، به نظر می‌رسد که مقیاس پنج‌گزینه‌ای مناسب باشد. برای تجزیه و تحلیل داده‌ها نیز با توجه به مقیاس داده، برای

REFERENCES

1. Li Q. A novel Likert scale based on fuzzy sets theory.

Expert Syst Appl. 2013;40(5):1609-18.

- DOI:10.1016/j.eswa.2012.09.015
2. Orr MG, Ohlsson S. Relationship between complexity and liking as a function of expertise. *Music Percept Interdiscip J*. 2005;22(4):583-611. DOI:10.1525/mp.2005.22.4.583
 3. Jamieson S. Likert scales: How to (ab) use them. *Med Educ*. 2004;38(12):1217-8. PMID: 15566531 DOI: 10.1111/j.1365-2929.2004.02012.x
 4. Hodge DR, Gillespie D. Phrase completions: An alternative to Likert scales. *Soc Work Res*. 2003;27(1):45-55. DOI:10.1093/swr/27.1.45
 5. Salimi M, Shahbazmoradi S, Bamdad Soufi J. A managerial research approach toward designing the sum of scores scale (Likert). *J Manag Knowl*. 2008;21(80):41-60. [Persian]
 6. Rattray J, Jones MC. Essential elements of questionnaire design and development. *J Clin Nurs*. 2007;16(2):234-43. PMID: 17239058 DOI: 10.1111/j.1365-2702.2006.01573.x
 7. Sullivan GM, Artino Jr AR. Analyzing and interpreting data from Likert-type scales. *J Grad Med Educ*. 2013;5(4):541-2. PMID: 24454995 DOI: 10.4300/JGME-5-4-18
 8. Hartley J. Some thoughts on Likert-type scales. *Int J Clin Health Psychol*. 2014;14(1):83-6. DOI: 10.1016/S1697-2600(14)70040-7
 9. Munshi J. A method for constructing Likert scales. SSRN. 2014 Apr.
 10. Russell CJ, Bobko P. Moderated regression analysis and Likert scales: Too coarse for comfort. *J Appl Psychol*. 1992;77(3):336-42. DOI: 10.1037/0021-9010.77.3.336
 11. Chang L. A psychometric evaluation of 4-point and 6-point Likert-type scales in relation to reliability and validity. *Appl Psychol Meas*. 1994;18(3):205-15. DOI: 10.1177/014662169401800302
 12. Chan JC. Response-order effects in Likert-type scales. *Educ Psychol Meas*. 1991;51(3):531-40. DOI: 10.1177/0013164491513002
 13. Ray JJ. Acquiescence and problems with forced-choice scales. *J Soc Psychol*. 1990;130(3):397-9.
 14. Brown JD. Using surveys in language programs. Cambridge: Cambridge University Press; 2001.
 15. Albaum G. The Likert scale revisited. *Int J Mark Res*. 1997;39(2):1-21. DOI: 10.1177/147078539703900202
 16. Edwards ML, Smith BC. The effects of the neutral response option on the extremeness of participant responses. *Incite*. 2014;6.
 17. Alexandrov A. Characteristics of single-item measures in Likert scale format. *Electron J Bus Res Methods*. 2010;8(1):1-12.
 18. Gliem JA, Gliem RR. Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. 2003 Midwest Research to Practice Conference in Adult, Continuing, and Community Education. Columbus: Ohio State University; 2003.
 19. Friedman H, Herskovitz PJ, Pollack S. Biasing effects of scale - checking styles on response to a Likert scale. Proceedings of the Section on Survey Research Methods. Alexandria VA: American Statistical Association; 1993.
 20. Johns R. Likert items and scales [Internet]. Glasgow: University of Strathclyde; 2010 [cited 2010 Mar 1]. Available from: https://ukdataservice.ac.uk/media/262829/discover_likertfa.ctsheet.pdf
 21. Alreck PL, Settle RB. The survey research handbook. Irwin: Irwin; 1995.
 22. Grimm P. Pretesting a questionnaire. *Wiley Int Encycl Mark*. 2010 Dec.
 23. Mellor D, Moore KA. The use of Likert scales with children. *J Pediatr Psychol*. 2014;39(3):369-79. PMID: 24163438 DOI: 10.1093/jpepsy/ist079
 24. Goodman R. The strengths and difficulties questionnaire: A research note. *J Child Psychol Psychiatry*. 1997;38(5):581-6. DOI: 10.1111/j.1469-7610.1997.tb01545.x
 25. Bartholomeu D, Rocha Da Silva MC, Montiel JM. Improving the Likert scale of the children's social skills test by means of rasch model. *Psychology*. 2016;7(6):820-8. DOI: 10.4236/psych.2016.76085
 26. Warmbrod JR. Reporting and interpreting scores derived from Likert-type scales. *J Agric Educ*. 2014;55(5):30-47. DOI: 10.5032/jae.2014.05030
 27. Sarmad Z, Bazargan A, Hejazi E. Research methods in behavioral sciences. 8th Edition. Tehran: Agah; 2004. pp. 132-7. [Persian]
 28. Murray J. Likert data: What to use, parametric or non-parametric?. *Int J Bus Soc Sci*. 2013;4(11):258-64.
 29. Vigderhus G. The level of measurement and "permissible" statistical analysis in social research. *Sociol Perspect*. 1977;20(1):61-72. DOI: 10.2307/1388904
 30. Norman G. Likert scales, levels of measurement and the "laws" of statistics. *Adv Health Sci Educ Theory Pract*. 2010;15(5):625-32. PMID: 20146096 DOI: 10.1007/s10459-010-9222-y
 31. Carifio J, Perla R. Resolving the 50-year debate around using and misusing Likert scales. *Med Educ*. 2008;42(12):1150-2. PMID: 19120943 DOI: 10.1111/j.1365-2923.2008.03172.x
 32. Hren D, Lukić IK, Marusić A, Vodopivec I, Vujaklija A, Hrabak M, et al. Teaching research methodology in medical schools: Students' attitudes towards and knowledge about science. *Med Educ*. 2004;38(1):81-6. DOI: 10.1111/j.1365-2923.2004.01735.x
 33. Kuzon WM Jr, Urbanek MG, Mc Cabe S. The seven deadly sins of statistical analysis. *Ann Plast Surg*. 1996;37(3):265-72. DOI: 10.1097/0000637-199609000-00006
 34. Knapp TR. Treating ordinal scales as interval scales: An attempt to resolve the controversy. *Nurs Res*. 1990;39(2):121-3. PMID: 2315066 DOI: 10.1097/00006199-199003000-00019
 35. Lukacz ES, Lawrence JM, Burchette RJ, Lubner KM, Nager CW, Buckwalter JG. The use of Visual Analog Scale in urogynecologic research: A psychometric evaluation. *Am J Obstet Gynecol*. 2004;191(1):165-70. PMID: 15295359 DOI: 10.1016/j.ajog.2004.04.047
 36. De Winter JCF, Dodou D. Five-point Likert items: T test versus Mann-Whitney-Wilcoxon. *Pract Assess Res Eval*. 2010;15(11):1-16.
 37. Mc Crum-Gardner E. Which is the correct statistical test to use?. *Br J Oral Maxillofac Surg*. 2008;46(1):38-41. PMID: 17961892 DOI: 10.1016/j.bjoms.2007.09.002