

اجرای فرایند روائی محتوا در طراحی پرسشنامه‌ها

ابراهیم عبدالله پور^۱، سحرناز نجات^۲، مریم نوروزیان^۳، سید رضا مجدزاده^۴

^۱ دانشجوی کارشناسی ارشد اپیدمیولوژی، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران، ایران
^۲ استادیار گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، مرکز بهره برداری از دانش سلامت، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران، ایران
^۳ دانشیار گروه نورولوژی، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران، ایران

^۴ استاد گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، مرکز بهره برداری از دانش سلامت، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران، ایران
نویسنده رابط: سحرناز نجات، نشانی: تهران، انقلاب، خیابان قدس، خیابان پورسینا، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران، دانشکده بهداشت، گروه اپیدمیولوژی و آمار زیستی.
تلفن: ۰۲۱-۶۶۴۹۵۸۵۹، نمابر: ۰۲۱-۶۶۴۹۵۸۵۹، پست الکترونیک: nejatsan@tums.ac.ir

تاریخ دریافت: ۱۳۸۸/۱۲/۲۴؛ پذیرش: ۱۳۸۹/۴/۱۲

مقدمه و اهداف: روائی محتوا نخستین نوع از انواع روائی‌ها است که باید از وجود آن در طول فرایند طراحی یک ابزار اطمینان حاصل شود. پیش نیاز بودن این روائی برای سایر انواع روائی‌ها و ارتباط نزدیک آن با پایایی اجرای آنرا گامی حیاتی در طراحی پرسشنامه معرفی نموده است. با این وجود کم اهمیت شمردن آن در مقالات مختلف (مخصوصاً مقالات داخلی) شفاف‌سازی اهمیت اجرای این فرایند را در طراحی پرسشنامه‌ها ضروری ساخته است. لذا در مقاله حاضر مراحل تهیه و ارزیابی محتوای مناسب برای پرسش نامه‌ها بیان خواهد شد. **روش کار:** مراحل اجرای فرایند روائی محتوا به منظور طراحی پرسشنامه‌های روا و پایا شامل طراحی (شناسایی حیطه‌ها، تولید آیتم‌ها و ساخت ابزار) و کمی‌سازی (تعیین درجه توافق بین متخصصین، تعیین مناسبیت و شفافیت سوالات و مناسبیت، شفافیت و جامعیت کل ابزار) در مقاله حاضر بیان شده است.

نتیجه‌گیری: اهمیت وجود روائی محتوا به هنگام طراحی پرسشنامه‌ها تنها به اطمینان از دستیابی به شاخص‌های روائی محتوای مطلوب (مناسبیت، شفافیت و جامعیت) خلاصه نمی‌شود. در کنار دستیابی به شاخص‌های مذکور بهبود شاخص‌های پایایی ابزار و کمک به کاهش منابع (مالی و زمانی) مورد نیاز برای تهیه پرسشنامه از مزایای دیگر آن به شمار می‌رود. هر چند کیفی بودن پاسخ‌های متخصصین از جمله مشکلات اجرای آن است، ولی کمی‌تر شدن مراحل اجرای آن تا حدود زیادی نگرانی‌ها را در این مورد کاهش داده است. موارد مذکور لزوم اجرای دقیق مراحل فرایند روائی محتوا در طراحی پرسشنامه‌ها توسط محققین را نمایان می‌سازد.

واژگان کلیدی: روائی محتوا، طراحی پرسشنامه، فرایند روائی محتوا

مقدمه

سوالی کیفیت زندگی سازمان جهانی بهداشت WHOQOL-BREF درجه توانائی آیتم‌های انتخاب شده و حیطه‌های تشکیل دهنده (جسمانی، روانی، محیطی و اجتماعی) این پرسشنامه در انعکاس مفهوم کیفیت زندگی در واقع روائی محتوای پرسشنامه مذکور را در بر می‌گیرد (۴). روائی محتوا در عین حال که می‌تواند از حیاتی‌ترین گام‌ها در فرایند طراحی پرسشنامه‌ها بشمار آید (۲،۳) نخستین نوع از انواع روائی‌ها است که بایستی از وجود آن در طول فرایند طراحی یک ابزار اطمینان حاصل شود. در ضمن روائی محتوا را پیش نیاز سایر انواع روائی‌ها نیز معرفی کرده‌اند (۳). هرچند که روائی سازه (Construct Validity) به همراه روائی معیار (Criterion validity) یک ابزار اهمیت زیادی دارند، ولی

با در نظر گرفتن مجموع تعاریف ارائه شده در متون مختلف، روائی محتوا (Content Validity) را می‌توان "توانائی سئوال‌های انتخاب شده در انعکاس ویژگی‌های سازه مورد اندازه‌گیری" تعریف نمود (۱،۲،۳). وجود این توانائی برای حیطه‌های سازه مورد نظر نیز ضروری است. بدین معنی که سوال مهمی که در طول زمان اجرای فرایند روائی محتوا و پس از آن باید پاسخ داده شود این است که: سئوال‌های انتخاب شده برای ورود به ابزار نهائی چقدر می‌توانند گویای تمامی سئوال‌هایی باشند که معیارهای لازم برای ورود به ابزار را دارند. درعین حال حیطه‌های انتخاب شده برای تشکیل ابزار نهائی تا چه اندازه گویای تمامی حیطه‌های سازه مورد بررسی هستند (۱،۳). به عنوان مثال در مورد پرسشنامه ۲۶

Content Validation * OR Content Validity*) AND
(Questionnaire Development*)

بر اساس مقالات متدولوژیک یافت شده مراحل اجرای این فرایند در طراحی پرسشنامه‌ها به صورت زیر ارائه می‌گردد:

۱- طراحی (Development)

۲- کمی‌سازی (Judgment – quantification) (۱،۲،۶)

۱- طراحی

مرحله طراحی در فرایند روانی محتوا را شامل سه بخش زیر می‌دانند (۲):

الف- شناسایی حیطه‌ها

ب- تولید سؤال‌ها

ت- ساخت ابزار

الف. شناسایی حیطه‌ها و تولید سؤال‌ها

شناسایی حیطه‌ها و سپس تولید سؤال‌ها گام‌های اولیه طراحی روای پرسشنامه را تشکیل می‌دهند که خود مستلزم ارائه تعاریف تئوریک (Conceptual) و عملکردی (Operational) مناسب از سازه مورد بررسی می‌باشند (۳).

تعریف تئوریک:

از جمله مشکلات شایع در فرایند طراحی پرسشنامه‌ها عدم اتفاق نظر در مورد مفهوم و حیطه(های) سازه مورد بررسی است. فراهم ساختن تعریف تئوریک برای سازه مورد نظر را می‌توان گام اولیه فرایند روانی محتوا به شمار آورد که معمولاً نادیده گرفته می‌شود (۲). بنابراین فرایند طراحی پرسشنامه بایستی با ارائه یک تعریف تئوریک واضح از سازه مورد بررسی آغاز گردد. این تعریف تئوریک در واقع پایه و اساس گام‌های بعدی طراحی ابزار خواهد بود (۳). به عنوان مثال در صورت طراحی پرسشنامه WHOQOL-BREF با استفاده از فرایند روانی محتوا گام اول استفاده از تعاریف مفهوم کیفیت زندگی است. بدین منظور میتوان تعریف ارائه شده از طرف سازمان جهانی را به عنوان یکی از تعاریف مورد نظر در کنار سایر تعاریف موجود در نظر گرفت. یک تعریف تئوریک مناسب باید حیطه‌های موضوع مورد بررسی را کاملاً مشخص نموده و در عین حال گویای اهمیت هر کدام از این حیطه‌ها باشد (۲). در برخی مواقع اهمیت هر حیطه می‌تواند تعیین کننده تعداد سؤال‌های پوشش دهنده همان حیطه از محتوای مورد بررسی باشد. در صورت جدید بودن مفهوم محتوای مورد بررسی ارائه تعریف تئوریک با استفاده از نظر متخصصین

اطمینان از مطلوب بودن روانی محتوای یک پرسشنامه علاوه بر تأثیر بر بهبود کیفیت ابزار جدید، بر افزایش پایایی پرسشنامه نیز بسیار تأثیرگذار خواهد بود (۱). بدین معنی که ارتباط روانی محتوا با پایایی (Reliability) ابزار را به این صورت عنوان کرده‌اند که در صورتی که یک پرسشنامه از روانی محتوای مطلوبی برخوردار نباشد نمی‌توان انتظار داشت پایایی مطلوبی داشته باشد (۳). این در حالیست که معمولاً چنین مرحله‌ای در طراحی پرسشنامه‌ها آنگونه که شایسته اهمیت آن است در نظر گرفته نشده و اغلب بصورت سطحی و گذرا به آن پرداخته می‌شود و در عین حال در متون علمی و پژوهشی در مقایسه با سایر انواع روانی‌ها توجه کمتری به آن شده است (۳). در این میان فرایند روانی محتوا را می‌توان یک فرایند اساسی در طراحی پرسشنامه به شمار آورد که در طی آن نشان داده می‌شود که سؤال‌های انتخاب شده برای ورود به ابزار می‌توانند به طور کامل حیطه‌های سازه مورد نظر را اندازه‌گیری کنند یا نه (۲). هرچند که ارزیابی معمول روانی محتوا به صورت کیفی و ذهنی (Subjective) را شاید بتوان از علل کم اهمیت شمرده شدن آن قلمداد کرد، ولی امروزه کمی‌تر شدن مراحل اجرای فرایند روانی محتوا با ارائه شاخص‌های کمی شناخته شده، آن را به عنوان فرایندی مناسب و قدرتمند جهت طراحی روای پرسشنامه معرفی نموده است (۱،۲،۳). درحالی که نیاز به ارزشیابی و تجدید نظر در طول زمان انجام فرایند طراحی پرسشنامه مستلزم صرف منابع بسیاری توسط محققین است، کاربرد فرایند روانی محتوا با وجود احتمال صرف منابع بیشتر در ابتدا، بدلیل کاهش تعداد دفعات نیاز به اصلاح و تجدید نظر و نیز کاهش منابع مورد نیاز در طول مراحل اصلاح آن نیاز مطالعه به صرف منابع را به طرز چشمگیری کاهش خواهد داد (۳). با وجود اینکه موارد فوق اجرای فرایند روانی محتوا را گامی حیاتی در طراحی پرسشنامه معرفی می‌کنند، کم اهمیت شمردن آن در مقالات مختلف مخصوصاً مقالات داخلی، شفاف‌سازی اهمیت اجرای فرایند روانی محتوا در طراحی پرسشنامه‌ها را ضروری می‌سازد. لذا در مقاله حاضر مراحل تهیه محتوای مناسب برای پرسش نامه‌ها و همچنین ارزیابی آن بیان خواهد شد.

روش کار

به منظور شناسایی مقالات در برگزیده مراحل اجرای فرایند روانی محتوا پایگاه اطلاعاتی MEDLINE از طریق Pubmed و نیز پایگاه‌های اطلاعاتی Web of science و Scopus با کلید واژه‌های زیر مورد جستجو قرار گرفت:

تعداد آن‌ها ضروری به نظر می‌رسد.

متخصصین شرکت کننده در فرایند روائی محتوا:

در طول فرایند روائی محتوا نظرات ۳ دسته از متخصصین اخذ خواهد شد:

۱- متخصصین محتوا (Content expert)

افرادی که محتوای مورد نظر را به لحاظ نظری می‌شناسند، یا دارای تجارب بالینی (Expertise Clinical) در این زمینه هستند، در مورد موضوعات مرتبط با محتوای مورد نظر پژوهش انجام داده‌اند و یا مقالات پژوهشی مرتبط دارند (۳). برای مثال به هنگام طراحی ابزار «اندازه‌گیری فشار وارده بر مراقبین بیماران مبتلا به دمانس» می‌توان از نورولوژیست‌های با تجربه و متخصص در زمینه "حافظه" و روانپزشکان که دارای مقالات پژوهشی مرتبط هستند یا تجربه بالینی کافی دارند استفاده کرد.

۲- گروهی از مردم آگاه (Lay expert)

هر چند که نقش متخصصین محتوا در طراحی پایا و روایی پرسشنامه‌ها محوری است لیکن بازنگری سئوال‌ها توسط نمونه‌ای از افراد جمعیت هدف که ابزار مورد نظر برای آن‌ها طراحی می‌گردد، بخش مهمی از فرایند روائی محتوا را تشکیل می‌دهد (۲). ممکن است شاخص‌های روائی محتوای سئوال‌های پیشنهادی، از نظر متخصصین محتوا مطلوب و قابل قبول باشد در حالی که از نظر شرکت کنندگان بالقوه در مطالعه (مردم آگاه) این‌گونه نباشد. بنابراین دخالت این دسته از متخصصین در نمره‌دهی به شاخص‌های روائی محتوا و در نهایت تصمیم‌گیری در مورد ورود سئوال‌های مورد نظر به ابزار نهائی و یا حذف آن‌ها ضروری است. به هر حال تصمیم نهائی در صورت وجود اختلاف نظر بین متخصصین محتوا و شرکت کنندگان بالقوه در مطالعه با محقق اصلی است.

۳- متدولوژیست‌ها

افرادی هستند که در زمینه طراحی پرسشنامه کار کرده، مقاله مرتبط دارند و مباحث نظری و علمی اندازه‌گیری آن را می‌دانند.

تعداد متخصصین

تعداد متخصصین محتوا و شرکت کنندگان بالقوه (مردم آگاه) در فرایند روائی محتوا در منابع مختلف متفاوت اعلام شده است. بدین ترتیب که در حالیکه Lynn تعداد متخصصین مورد نیاز برای هر گروه را حداقل ۳ نفر عنوان کرده است و از نظر ایشان نیازی

محتوا، مرحله آغازین فرایند روائی محتوا را تشکیل خواهد داد و در غیر این صورت بجای تهیه و ارائه تعریف تئوریک از تعاریف موجود استفاده می‌شود. دستیابی به یک تعریف تئوریک کامل به ۲ صورت زیر امکان پذیر است:

- مرور منظم و کامل متون علمی و پژوهش‌های مرتبط با محتوای مورد بررسی
- بهره‌گیری از نظرات متخصصین (۳).

تعریف عملکردی

ارائه تعریف عملکردی مستلزم انتخاب حیطه‌های مورد نظر طراح پرسشنامه، از میان حیطه‌های شناسائی شده توسط تعریف تئوریک می‌باشد. تعریف عملکردی مستقیماً از تعریف تئوریک منشا می‌گیرد و همانند تعریف تئوریک مطلوب بودن روائی محتوا بطور مستقیم با تعریف عملکردی نیز در ارتباط است (۳). بنابراین می‌توان نتیجه دستیابی به یک تعریف تئوریک و پیرو آن تعریف عملکردی مناسب از محتوای مورد نظر را شناسائی تمامی حیطه‌های آن و در نهایت تولید سئوال‌های مورد نیاز دانست. در واقع سئوال‌های انتخاب شده برای ورود به ابزار نهائی بایستی برگرفته از یک تعریف عملکردی کامل از محتوای مورد نظر باشند (۱).

ساخت ابزار

ممکن است سوالات تولید شده در مراحل پیشین نیازمند اعمال تغییراتی در زمینه نحوه نگارش و یا نحوه (ترتیب) قرار گرفتن آن‌ها قبل از ارائه به متخصصین باشند. بنابراین این مرحله شامل تصحیح سئوال‌ها و سازماندهی آن‌ها در یک توالی مناسب است (۲).

۲- کمی‌سازی

مرحله دوم از مراحل اجرای فرایند روائی محتوا شامل ارزشیابی و کمی‌سازی شاخص‌های روائی هر یک از سئوال‌ها (مناسبت، شفافیت) به همراه شاخص‌های روائی کل ابزار (مناسبت، شفافیت و جامعیت) است که توسط تعداد مشخصی از متخصصین (متخصصین محتوا و شرکت کنندگان بالقوه در مطالعه) صورت می‌پذیرد. از آنجایی که انتخاب نامناسب متخصصین شرکت کننده در فرایند روائی محتوا را می‌توان از جمله دلایل نامطلوب بودن شاخص‌های روائی و پایایی یک پرسشنامه دانست و صحت انجام مراحل روائی محتوا کاملاً به نحوه انتخاب و بکارگیری متخصصین در طول زمان طراحی ابزار وابسته است (۲) معرفی متخصصین شرکت کننده در فرایند روائی محتوا به همراه معیارهای انتخاب و

ارزشیابی کرده و در صورت نیاز تغییرات نگارشی مورد نیاز را اعمال، سؤال‌های مورد نیاز را به پرسشنامه اضافه و یا در صورت لزوم سؤال‌های اضافی را حذف کند (۲،۳).

از نظر Waltz و همکارانش اعتبار (Accuracy) متخصصین محتوا را می‌توان از طریق وارد ساختن تعدادی از سؤال‌های غیرگویا و نامتجانس به ابزار کنترل کرد. به عقیده وی نمرات تخصصینی که نتوانسته اند این سؤال‌ها را شناسایی کنند در محاسبه نهائی در نظر گرفته نخواهد شد (۵).

شاخص‌های روائی محتوا

قبل از معرفی شاخص‌های روائی محتوا لازم است مفهوم، اهمیت و نحوه محاسبه شاخص درجه توافق کلی بین متخصصین در مورد شاخص‌های روائی کلی ابزار ذکر گردد.

شاخص درجه توافق کلی (Inter Rater Agreement : IRA)

بیان کننده درجه توافق مشاهده شده بین متخصصین شرکت کننده در مطالعه در مورد شاخص‌های مناسبت و شفافیت سؤال‌های ابزار است. شاخص درجه توافق کلی هم برای مناسبت و هم شفافیت سؤال‌های پیشنهادی قابل محاسبه است. توصیه شده است که قبل از آغاز مرحله ارزیابی روائی محتوا ابتدا شاخص مذکور در مورد مناسبت و شفافیت سوالات پرسشنامه تعیین گردد (۲،۳) چرا که این شاخص را یک عامل کنترل کننده در فرایند روائی محتوا می‌دانند و توصیه کرده‌اند در صورتی که متخصصین در مورد مناسبت و شفافیت سؤال‌های ابزار توافق مطلوبی نداشته باشند، قبل از ورود به مرحله ارزیابی تغییرات اساسی در سؤال‌های پیشنهادی و یا متخصصین انتخاب شده صورت پذیرد (۲،۵). در مقالات مختلف محاسبه «درجه توافق کلی» در مورد شاخص‌های مناسبت و شفافیت با ۲ رویکرد زیر انجام گرفته است:

۱- رویکرد محافظه کارانه

در این رویکرد مجموع تعداد سؤال‌هایی که ۱۰۰٪ متخصصین محتوا و مردم آگاه از میان ۴ گزینه کاملاً مطلوب، مطلوب، تاحدودی مطلوب و نامطلوب، گزینه‌های "کاملاً مطلوب" و "مطلوب" را برای مناسبت و (یا شفافیت) آن‌ها انتخاب کرده باشند و یا برعکس ۱۰۰٪ متخصصین مذکور از میان ۴ حالت مذکور گزینه‌های تاحدودی مطلوب و نامطلوب را برای ۲ شاخص مذکور انتخاب کرده باشند بر تعداد کل سؤال‌های پرسشنامه بصورت زیر تقسیم می‌گردد (۲،۳).

به بیشتر از ۱۰ نفر برای هر گروه از متخصصین وجود ندارد (۱،۲،۳) Wolf و Gable به همراه Lenz و Strickland، Waltz این تعداد را ۲ تا ۲۰ نفر اعلام کرده‌اند. از نظر Doris Mc Gartland نیز تعداد متخصصین مورد نیاز برای هر کدام از دو گروه مورد نظر ۳ تا ۱۰ نفر است، هر چند عقیده دارد تعداد بیشتر متخصصین منجر به تولید اطلاعات بیشتر خواهد شد. معیار تعیین کننده تعداد متخصصین از نظر Davis و Grant درجه تخصص و تنوع دانش اعلام شده است (۳) به هر حال تعداد متخصصین مورد نیاز برای شرکت در فرایند روائی محتوا بر اساس معیارهایی از جمله سطح دانش متخصصین و میزان گویائی آن‌ها قابل پیش بینی است (۲). در غالب موارد تعداد متدولوژیست‌های مورد نیاز برای اجرای فرایند روائی محتوا یک نفر است.

معیارهای انتخاب متخصصین

هر چند پژوهشگران و طراحان پرسشنامه معمولاً روائی محتوای ابزارشان را مورد تایید متخصصین شرکت کننده در مراحل فرایند روائی محتوا اعلام می‌کنند، اما شرایط و ویژگی‌های متخصصین و مراحل اجرای فرایند را گزارش نمی‌کنند. این در حالی است که صحت انجام مراحل فرایند روائی محتوا کاملاً تحت تأثیر چگونگی انتخاب و بکارگیری متخصصین در مراحل طراحی ابزار قرار دارد. از نظر Grant و Kinney سابقه انتشارات در مجلات مرجع، تعداد مقالات، انجام تحقیقات در زمینه موضوع مورد نظر و تجارب بالینی (Clinical Expertise) می‌توانند به عنوان معیارهای انتخاب متخصصین محتوا بکار روند (۲،۳). از نظر Davis از آنجائیکه معمولاً شناسایی متخصصین محتوایی که تمامی معیارهای مورد نیاز برای ورود به مطالعه را داشته باشند مشکل و تا حدودی غیر ممکن به نظر می‌رسد بهتر است از گروه‌های مختلفی از متخصصین که بتوانند تمامی جنبه‌های محتوای مورد نظر را مورد قضاوت و ارزیابی قرار دهند استفاده شود (۲).

شواهد روائی محتوای پرسشنامه اساساً قضاوتی است و از پانلی از متخصصین که بطور کامل با محتوای مورد بررسی در ارتباط هستند به دست خواهد آمد. بدین صورت که متخصصین پس از انجام بحث‌های گروهی بر روی تعاریف نظری و عملکردی ارائه شده و سپس انتخاب سؤال‌ها، به کامل بودن تعریف تئوریک نمره داده و درجه تطابق تعریف عملکردی با تعریف نظری را نیز تعیین می‌کنند. همچنین توانائی سؤال‌ها در انعکاس ویژگی‌های تعریف نظری هر دامنه ارزیابی می‌شود. طراح پرسشنامه همچنین می‌تواند از متخصصین درخواست کند تا نحوه نگارش هر سؤال را

توافق کلی در مورد مناسبت^۱ (رویکرد محافظه کارانه) = تعداد توافق مشاهده شده^۲ بین متخصصین / تعداد کل آیت‌ها

۲- رویکرد کمتر محافظه کارانه

در رویکرد کمتر محافظه کارانه مجموع تعداد سئوال‌هایی که ۸۰٪ متخصصین از میان ۴ گزینه کاملاً مطلوب، مطلوب، تاحدودی مطلوب و نامطلوب، گزینه‌های "کاملاً مطلوب" و "مطلوب" را برای مناسبت (یا شفافیت) آن‌ها انتخاب کرده باشند و یا برعکس ۸۰٪ متخصصین مذکور از میان ۴ حالت مذکور گزینه‌های تاحدودی مطلوب و نامطلوب را برای ۲ شاخص مذکور انتخاب کرده باشند بر تعداد کل سئوال‌های پرسشنامه بصورت زیر تقسیم می‌گردد (۲،۳).

توافق کلی در مورد مناسبت^۳ (رویکرد کمتر محافظه کارانه) = تعداد توافق مشاهده شده^۴ بین متخصصین / تعداد کل آیت‌ها

سطح قابل قبول "شاخص درجه توافق کلی" از نظر Davis ۷۰ درصد و از نظر Selby و بقیه ۸۰ درصد عنوان شده است. در صورت مطلوب بودن شاخص درجه توافق کلی تجدید نظر اولیه تا قبل از انجام آنالیز روائی محتوا ضروری نیست و مطلوب نبودن آن بیانگر نیاز به اعمال تجدید نظرهای اساسی در سئوال‌های ابزار است (۴). پس از اطمینان از مطلوب بودن شاخص درجه توافق کلی مرحله تعیین مناسبت و شفافیت سئوال‌ها، مناسبت و شفافیت کلی ابزار به همراه جامعیت پرسشنامه آغاز خواهد شد

(۲،۵).

بطور کلی شاخص‌های روائی محتوا را می‌توان بصورت زیر ارائه کرد:

- مناسبت و شفافیت هر سوال (Item Content Validity Index: I-CVI)
- مناسبت و شفافیت کلی ابزار (Scale Content Validity Index: S-CVI)
- جامعیت کلی ابزار (Comprehensiveness)

متخصصین شرکت کننده در فرایند روائی محتوا به شاخص‌های مذکور نمره داده و نظرات خود را در مورد آن‌ها اعلام می‌کنند. در ادامه مفهوم و نحوه محاسبه هر یک شرح داده خواهد شد.

مناسبت (Relevancy)

درجه توانایی سوال انتخاب شده در انعکاس ویژگی‌های محتوای مورد بررسی، مناسبت سئوال مورد نظر را تشکیل می‌دهد و در واقع بیانگر این نکته است که سوال پیشنهاد شده با محتوای مورد اندازه‌گیری تا چه اندازه مرتبط می‌باشد؟ (۲،۳). معمولاً گزینه‌های شاخص‌های روائی محتوا چهار حالتی است. مناسبت هر سئوال به همراه مناسبت کلی ابزار به چهار حالت ۱ (نامطلوب)، ۲ (تا حدی مطلوب)، ۳ (مطلوب) و ۴ (کاملاً مطلوب) قابل تقسیم است (۲،۱،۳،۵).

جدول شماره ۱- مقایسه مقدار شاخص مناسبت کلی یک پرسشنامه ۹ سوالی بدست آمده از ۲ رویکرد میانگین و توافق کلی

سئوال	نمرات متخصصین (۱ تا ۵) به شاخص "مناسبت" هر کدام از سئوال‌ها					تعداد توافق مشاهده شده بین ۵ متخصص	مناسبت هر سئوال	مناسبت کلی ابزار (رویکرد توافق کلی)	
	۵	۴	۳	۲	۱			(رویکرد میانگین)	(رویکرد توافق کلی)
۱	*۱	۴	۴	۳	۳	۴	٪۸۰		
۲	۴	*۱	۴	۳	۳	۴	٪۸۰		
۳	۴	۳	*۲	۳	۳	۴	٪۸۰		
۴	۳	۳	۴	*۲	۳	۴	٪۸۰		
۵	۳	۳	۴	۴	*۲	۴	٪۸۰		
۶	۳	۳	۴	۴	۴	۵	٪۱۰۰		
۷	۳	۳	۴	۴	۴	۵	٪۱۰۰		
۸	۳	۳	۴	۴	۴	۵	٪۱۰۰		
۹	۳	۳	۴	۴	۴	۵	٪۱۰۰		
								۴۴/۴۴	۸۸/۸۸

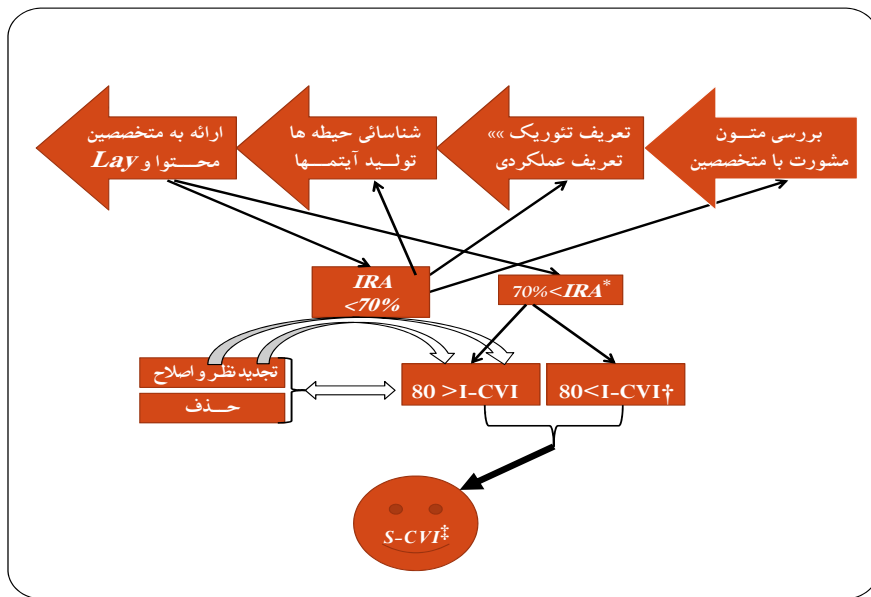
* نمرات متخصصین به مناسبت سئوال پرسشنامه که از نظر آنان مناسبت مطلوبی ندارند.

^۱ توافق کلی در مورد شفافیت (در رویکرد محافظه کارانه) نیز دقیقاً به همین ترتیب محاسبه خواهد شد.

^۲ مجموع آیت‌هایی که مناسبت آن‌ها توسط ۱۰۰٪ متخصصین مطلوب تشخیص داده شده است + آیت‌هایی که مناسبت آن‌ها توسط ۱۰۰٪ متخصصین نامطلوب تشخیص داده شده است.

^۳ توافق کلی در مورد شفافیت (رویکرد کمتر محافظه کارانه) نیز دقیقاً به همین ترتیب محاسبه خواهد شد.

^۴ مجموع آیت‌هایی که مناسبت آن‌ها توسط ۸۰٪ متخصصین مطلوب تشخیص داده شده است + مجموع آیت‌هایی که مناسبت آن‌ها توسط ۸۰٪ متخصصین نامطلوب تشخیص داده شده است.



شکل ۱ - مراحل انجام فرایند روانی محتوا

* شاخص درجه توافق کلی بین متخصصین

† شاخص روانی محتوا (مناسبت یا شفافیت) برای هر سوال

‡ شاخص روانی محتوا (مناسبت، شفافیت و جامعیت) برای کل ابزار

کند (۲). جامعیت کلی ابزار نیز همانند شاخص‌های مناسبت و شفافیت به چهار طبقه ۱ (ناقص)، ۲ (تاحدودی جامع)، ۳ (جامع) و ۴ (کاملاً جامع) قابل تقسیم است (۲).

تعیین شاخص‌های روانی محتوا

از میان شاخص‌های روانی محتوا دو شاخص مناسبت و شفافیت در دو سطح (یکی در سطح تک تک آیتم‌ها و دیگری در سطح کل ابزار) قابل محاسبه هستند. درحالی که شاخص جامعیت ابزار را تنها می‌توان در سطح کل ابزار محاسبه نمود. نحوه محاسبه هر کدام از این شاخص‌ها در ادامه آورده شده است.

تعیین مناسبت و شفافیت هر سوال

جهت محاسبه شاخص شفافیت هر کدام از سوالات مجموع تعداد متخصصینی که شفافیت هر سوال را مناسب و یا کاملاً مناسب تشخیص داده‌اند بر تعداد کل متخصصین تقسیم می‌گردد. در این صورت بدیهی است دامنه شاخص مذکور عددی بین صفر تا یک خواهد بود. شاخص مناسبت هر کدام از سئوال‌ها نیز دقیقاً به همین طریق محاسبه می‌شود (۱،۲،۶). کاربرد شاخص‌های روانی هر سوال در واقع راهنمایی محققین در تصمیم‌گیری برای حذف، اصلاح و یا جانشین سازی سئوال‌ها است. شاخص‌های روانی هر سوال معمولاً تنها در مطالعات

شفافیت (Clarity)

مناسب بودن سوالهای انتخاب شده از نظر نحوه نگارش و مفهوم آنها است (۳). ممکن است یک پرسشنامه بتواند حیطه‌های مرتبط با محتوای مورد بررسی را بدرستی ارائه دهد اما اطلاعات را به دلیل مبهم بودن مفهوم سئوالات و یا نحوه مقیاس‌بندی پاسخ‌ها به درستی منتقل نکند (۲). شفافیت هر سئوال و شفافیت کلی ابزار نیز همانند شاخص مناسبت به چهار حالت ۱ (نامطلوب)، ۲ (تا حدوی مطلوب)، ۳ (مطلوب) و ۴ (کاملاً مطلوب) قابل تقسیم است (۱،۲،۳،۵).

جامعیت ابزار (Comprehensiveness)

توانایی ابزار از نظر پوشش دادن تمامی حیطه‌های مرتبط با موضوع مورد بررسی جامعیت ابزار مورد نظر را تشکیل می‌دهد (۳). یک ابزار کامل بایستی در برگیرنده (یا گویای) تمامی حیطه‌های محتوای مورد بررسی باشد. متخصصین شرکت کننده در مطالعه به هنگام ارزشیابی ابزار، توانایی مجموعه سئوالات پیشنهادی را در انعکاس ویژگی‌های تمامی حیطه‌های محتوای مورد نظر ارزیابی می‌کنند. بر اساس نظرات متخصصین در مورد جامعیت ابزار، محقق می‌تواند سئوال‌های دیگری را به ابزار اضافه

آن‌ها توسط مجموعه متخصصین مطلوب تشخیص داده شده است بر حاصل جمع قضاوت‌های انجام شده (یعنی حاصل ضرب تعداد سؤال‌ها در تعداد متخصصین) توسط مجموعه متخصصین (۱).

جدول شماره ۱ تأثیر احتمالی دو رویکرد موجود در مقدار شاخص‌های مناسبت و شفافیت کلی ابزار را نشان می‌دهد. در این جدول نمرات ۵ متخصص به شاخص مناسبت سؤال‌های یک پرسشنامه ۹ سوالی نشان داده شده است. همانگونه که نشان داده شده است در صورت استفاده از رویکرد توافق کلی (از آنجائی که شاخص مناسبت تنها ۴ سؤال از سئوال‌ات پرسشنامه از نظر تمامی متخصصین مطلوب است) شاخص مناسبت کلی ابزار با وجود مطلوب بودن مناسبت تمامی سئوال‌ات از دیدگاه متخصصین تنها ۴۴/۴۴ درصد بدست خواهد آمد. این در حالیست که در صورت استفاده از رویکرد میانگین شاخص مناسبت کلی ابزار ۸۸/۸۸ درصد خواهد بود.

Denis F Pilot و Cheryl tatano Beck به همراه دیگران از عدم ذکر روش مورد استفاده برای محاسبه شاخص روائی (مناسبت و شفافیت) کلی ابزار توسط محققین انتقاد کرده و ذکر آن را از آنجایی که ممکن است تفاوت‌های قابل توجهی در نتیجه بدست آمده از آن‌ها وجود داشته باشد ضروری دانسته‌اند (۱،۶). همچنین توصیه شده است دامنه مقادیر شاخص‌های روائی سئوال‌اتی که در ابزار نهائی آورده شده‌اند بویژه در مواردی که از رویکرد توافق کلی برای بیان میانگین شاخص روائی کلی ابزار استفاده شده است ذکر گردد. چرا که در این رویکرد مقادیر شاخص‌های مناسبت و شفافیت سئوال‌اتی که در حد مطلوب نیستند نادیده گرفته شده و در محاسبه شاخص‌های روائی محتوای کلی ابزار دخالت داده نمی‌شوند (۱). در مقالات مختلف حداقل مناسبت مطلوب برای یک ابزار جدید ۸۰٪ عنوان شده است (۱،۲،۳،۵). مطلوب نبودن شاخص مناسبت را می‌توان بیانگر نمونه‌گیری غیر گویا و ناقص از حیطه‌های محتوای مورد بررسی دانست (۵).

تعیین شفافیت کلی ابزار

شفافیت کلی ابزار از تقسیم مجموع شاخص‌های شفافیت تک‌تک سؤال‌های پرسشنامه بر تعداد کل سئوال‌ات محاسبه می‌شود. در مقالات مختلف شفافیت مطلوب برای یک ابزار جدید ۸۰٪ عنوان شده است (۱،۲،۳،۵).

متدولوژیک که تمرکز آن‌ها بر توصیف فرایند روائی محتوا قرار دارد گزارش می‌شود و آنچه اغلب در مطالعات طراحی ابزار گزارش می‌شود شاخص روائی محتوا برای کل ابزار است (۱).

تعیین شاخص‌های مناسبت و شفافیت کلی ابزار

تعیین شاخص‌های مناسبت و شفافیت ابزار با ۲ روش شامل رویکرد توافق کلی و رویکرد میانگین امکان پذیر است. در هر دو رویکرد پس از ادغام گزینه‌های نامناسب و تاحدودی مناسب، و همچنین گزینه‌های مناسب و کاملاً مناسب، گزینه‌های دوحالتی (مناسب و نامناسب) برای هر سؤال تشکیل می‌گردد (۱،۳).

رویکرد توافق کلی

در رویکرد توافق کلی، مجموع تعداد سؤال‌هایی که مناسبت (یا شفافیت) آن‌ها توسط همه متخصصین مطلوب تشخیص داده شده است بر تعداد کل سؤال‌ها تقسیم می‌شود. هرچه تعداد متخصصین شرکت‌کننده در مطالعه بیشتر باشد احتمال پایین بودن شاخص مناسبت کلی ابزار (در رویکرد توافق کلی) بیشتر خواهد بود. چرا که تعداد سؤال‌هایی که همه متخصصین بطور همزمان مناسبت آن را مطلوب تشخیص داده باشند با افزایش تعداد متخصصین کاهش خواهد یافت (۱).

رویکرد میانگین

در رویکرد میانگین مجموع شاخص‌های روائی محتوای (مناسبت یا شفافیت) هر یک از سؤال‌های پرسشنامه بر تعداد کل سئوال‌ات تقسیم می‌گردد مناسبت (یا شفافیت) هر سؤال از تقسیم تعداد متخصصینی که مناسبت (یا شفافیت) سوال را مطلوب تشخیص داده‌اند بر تعداد کل متخصصین بدست می‌آید (۱،۴،۶). کاربرد رویکرد میانگین به عنوان میانگین مقادیر شاخص‌های روائی محتوای (مناسبت یا شفافیت) هر یک از سؤال‌های پرسشنامه از آنجائی که تمرکز و تأکید آن بجای میانگین کارایی (عملکرد) متخصصین بر میانگین کیفیت سؤال‌ها قرار دارد نسبت به رویکرد توافق کلی مناسب‌تر است. در ضمن در صورت محاسبه شاخص‌های مناسبت و شفافیت کلی ابزار با رویکرد میانگین با استفاده از هر کدام از روش‌های زیر نتایج یکسانی بدست خواهد آمد:

- میانگین شاخص‌های روائی برای هر سوال
- میانگین شاخص‌های روائی کلی ابزار (بدست آمده بر اساس نظرات هر کدام از متخصصین)
- تقسیم مجموع تعداد سؤال‌هایی که شاخص روائی

تعیین جامعیت ابزار

۳- ذهنی (غیر کمی) بودن نظرات متخصصین در مورد شاخص‌های روانی سؤال‌ها زمینه را برای سوگرا بودن نتایج مطالعه فراهم می‌سازد (۳).

بخش‌شناسی توافق مشاهده شده بین متخصصین در مورد شاخص‌های مناسبت و شفافیت

تلاش‌هایی به منظور اصلاح بخش‌شناسی توافق مشاهده شده بین متخصصین (تطبیق شاخص‌های روانی محتوای سؤالات برای توافق‌شناسی) انجام گرفته که از آن جمله می‌توان به معیار پیشنهادی از طرف خانم Lynn در سال ۱۹۸۶ اشاره کرد. بدین صورت که پس از محاسبه حدود اطمینان ۰/۹۵٪ برای شاخص‌های روانی محتوای هر یک از سؤالات با استفاده از خطای معیار نسبت $(\sqrt{\rho \cdot (1 - \rho)} / N)$ (که در آن P همان شاخص روانی سؤال‌های پرسشنامه و N تعداد متخصصین می‌باشد) حد پائینی شاخص روایی سؤال‌های پرسشنامه بدست آمده با ۰/۵ مقایسه می‌شود و در صورت بزرگتر از نیم بودن آن قابل قبول در نظر گرفته خواهد شد (۴).

رویکرد مذکور به دلایلی مورد انتقاد محققین قرار گرفته است و به همین دلیل روش دیگری از سوی Denise F. Polit و همکارانش پیشنهاد شد. در این روش بخش‌شناسی توافق مشاهده شده بین متخصصین (شاخص‌های روانی محتوای تک‌تک سؤال‌های پرسشنامه) با استفاده از شاخص کاپا حذف می‌شود. در صورتیکه مقدار کاپای بدست آمده برای هر سؤال بزرگتر و یا مساوی ۰/۷۵٪ باشد شاخص روانی مورد نظر عالی در نظر گرفته خواهد شد. توصیه شده است که در این مورد از نقطه برش ۰/۷۵٪ که توسط Sparrow و Cicchetti به همراه Fleiss در سال ۱۹۸۱ ارائه شده است بدلیل محافظه‌کارانه‌تر بودن آن نسبت به سایر مقیاس‌بندی‌های پیشنهادی برای کاپا استفاده شود (۴).

فرایند روانی محتوا در مقابل مطالعه روانی محتوا

فرایند روانی محتوا در واقع شامل مراحل طراحی پرسشنامه با هدف دستیابی به روانی محتوای مطلوب برای آن است که مراحل اجرای آن ارائه گردید، در حالی که مطالعه روانی محتوا (Content Validity Study) مطالعه ارزیابی یک ابزار طراحی شده به منظور تعیین شاخص‌های روانی محتوای (مناسبت، شفافیت و جامعیت) آن بوده و پس از طراحی پرسشنامه انجام می‌پذیرد. بنابراین زمان انجام مطالعه روانی محتوا پس از طراحی ابزار است در حالیکه زمان انجام فرایند روانی محتوا در حین طراحی ابزار می‌باشد (۳).

از تقسیم تعداد متخصصینی که جامعیت ابزار را مطلوب تشخیص داده‌اند بر تعداد کل متخصصین بدست می‌آید. پائین بودن شاخص جامعیت را می‌توان به دلیل یکی از موارد زیر دانست:

۱. حیطه‌ها و در نتیجه سؤال‌های انتخاب شده برای پرسشنامه نتوانسته‌اند تمامی حیطه‌های ابزار مورد نظر را پوشش دهند.
۲. متخصصین شرکت کننده در فرایند (در انجام قضاوت‌های خود و ارزشیابی ابزار) کاملاً خیره نبوده‌اند.
۳. موارد مذکور بر ضرورت انتخاب متخصصین آگاه و خبیره و همچنین انجام تلاش‌های وسیع برای تولید سؤال‌های مناسب توسط محقق در ابتدای فرایند روانی محتوا تأکید دوباره دارد (۴).
۴. بطور کلی بالا بودن شاخص‌های روانی محتوا (مناسبت، شفافیت و جامعیت) را می‌توان دلیلی بر مطلوب بودن روایی محتوای مطالعه مورد نظر (ابزار طراحی شده) دانست. در عین حال در صورت مطلوب نبودن شاخص‌های مذکور و بر اساس نظر متخصصین محتوا سؤال‌های مربوطه حذف و یا پس از اعمال تغییرات مورد نیاز زمینه ورود آن‌ها به ابزار نهایی در صورت صلاحدید متخصصین فراهم خواهد شد (۵).

زمان انجام فرایند روانی محتوا

با توجه به این که روانی محتوا را پیش‌نیاز دستیابی به سایر انواع روایی‌ها دانسته‌اند، دستیابی به روانی محتوای مطلوب برای یک ابزار جدید تنها در صورت اجرای فرایند روانی محتوا همزمان با طراحی پرسشنامه (نه پس از طراحی آن) امکان‌پذیر است (۳). شکل شماره ۱ مراحل انجام فرایند روانی محتوا را نشان می‌دهد.

ایرادات وارده به فرایند روانی محتوا

در کنار مزایای انجام فرایند روانی محتوا به منظور طراحی روای پرسشنامه انجام آن را خالی از اشکال نمی‌توان به حساب آورد. ایرادات وارده به فرایند روانی محتوا را می‌توان در قالب موارد زیر عنوان نمود:

- ۱- از دست دادن بخشی از اطلاعات بدلیل روی هم ریختن نظرات متخصصین محتوا و مردم آگاه و تشکیل پاسخ‌های ۲ حالت "مطلوب و نامطلوب" از پاسخ‌های ۴ حالت "نامطلوب، تاحدودی مطلوب، مطلوب و کاملاً مطلوب" اخذ شده (۱).
- ۲- در نظر نگرفتن بخش‌شناسی توافق مشاهده شده بین متخصصین (۱).

دلایل مشکل بودن فرایند روائی محتوا

- ویژگی‌هایی که اجرای فرایند روائی محتوا را در مقالات مختلف چالش برانگیز معرفی کرده است به صورت زیر قابل عنوان است:
- عدم توصیف دقیق مفاهیم نظری در علوم اجتماعی
 - عدم وجود توافق در مورد حیطه‌های بسیاری از سازه‌های مورد بررسی
 - مشکل بودن نمونه‌گیری مناسب از تمامی حیطه‌های محتوای مورد اندازه‌گیری
 - مشکل بودن حصول اطمینان از پوشش داده شدن تمامی جنبه‌ها (حیطه‌های) محتوای مورد بررسی توسط سوالات پرسشنامه
 - مشکل بودن اختصاص سهم مناسبی به هر کدام از حیطه‌های محتوای مورد نظر در پرسشنامه نهائی (۳).

نتیجه‌گیری

اهمیت وجود روائی محتوا به هنگام طراحی پرسشنامه‌ها تنها به اطمینان از دستیابی به شاخص‌های روائی محتوای مطلوب (مناسبت، شفافیت و جامعیت) خلاصه نمی‌شود. در کنار دستیابی به شاخص‌های مذکور که برای هر ابزار طراحی شده‌ای کاملاً ضروری به نظر می‌رسد، بهبود شاخص‌های پایائی ابزار در کنار

منابع

- 1- Polit DF, Beck CT. The Content Validity Index: Are you sure you know what's Being Reported? Critique and Recommendations. *Research in Nursing & Health* 2006; 29: 489-97.
- 2- Grant JS, Davis LL. Selection and Use of Content Experts for Instrument Development. *Research in Nursing & Health* 1997; 20: 269-274.
- 3- Rubio DMG, Berg-Weger M, Tebb SS, Lee ES, Rauch sh. Objectifying content validity: Conducting a content validity study in social Work Research. *ProQuest Psychology Journals* 2003; 27: 94.
- 4- WHOQOL-Bref Introduction, Administration, Scoring and Generic Version of the Assessment, Field Trial Version. December 1996, Available at: www.who.int/mental-health/media/en/76.pdf
- 5- Schutz AL, Counte MA, Meurer S. Development of a patient safety culture measurement tool for ambulatory health care settings: analysis of content validity. *Health Care Manage Sci* 2007; 10: 139-49.
- 6- Polit FD, Beck CT, Steven VO. Is the CVI an Acceptable Indicator of Content Validity? Appraisal and Recommendations. *Research in Nursing & Health* 2007; 30: 459-67.
- 7- Abdollahpour I, Nedjat S, Noroozian M, Golestan B and Majdzadeh R. Development of a Caregiver Burden Questionnaire for the Patients with Dementia in Iran, *Int J Prev Med*. 2010; 1: 233-241.

کمک به کاهش منابع مورد نیاز برای تهیه پرسشنامه (Efficiency) مزایائی است که به سادگی قابل اغماض نیست. در کنار مزایای اشاره شده کیفی و Subjective بودن پاسخ‌های متخصصین شرکت کننده در مطالعه از جمله مشکلات اجرای آن است. هر چند که کمی‌تر شدن مراحل اجرای آن تا حدود زیادی نگرانی‌ها را در این مورد کاهش داده است. از آنجایی که رویکرد توافق کلی جهت محاسبه مناسبت و شفافیت کلی ابزار آیت‌های نامطلوب را کاملاً نادیده می‌گیرد کاربرد رویکرد میانگین که نمرات تمامی آیت‌ها را در محاسبه شاخص‌های مذکور مد نظر قرار می‌دهد توصیه می‌گردد. در صورت وجود اختلاف در نظرات اعلام شده از طرف متخصصین، تشخیص نهائی صحت و اعتبار نظرات اعلام شده مستلزم برگزاری جلسات بارش افکار با حضور تمامی متخصصین و در نهایت توجه بیشتر به نظرات متخصصین دارای تجارب بالینی و مقالات مرتبط علمی بیشتر می‌باشد لزوم اجرای دقیق مراحل روایی محتوا انجام تجدید نظرهای مورد نیاز در سؤال‌های تولید شده، ارزشیابی ابزار در هر کدام از مراحل تجدید نظر و در صورت نیاز بازگشت از یک مرحله به مرحله قبلی به منظور طراحی روای یک ابزار جهت دستیابی به روایی محتوای مطلوب برای ابزار جدید مورد تأکید است.

Iranian Journal of Epidemiology 2011; 6(4): 66-74.

Review Article

Performing Content Validation Process in Development of Questionnaires

Abdollahpour I¹, Nedjat S², Noroozian M³, Majdzadeh R⁴

1- M.Sc student in epidemiology, Epidemiology and Biostatistics Department of Public Health School, Tehran University of Medical Science, Iran

2- Assistant Professor, Epidemiology and Biostatistics Department of Public Health School, Knowledge Utilization Research Center, Tehran University of Medical Sciences, Iran

3- Associate Professor of Neurology, Tehran University of Medical Science, Iran

4- Professor, Epidemiology and Biostatistics Department of Public Health School, Knowledge Utilization Research Center, Tehran University of Medical Science, Iran

Corresponding author: Nejat S., nejatsan@tums.ac.ir

Background & objectives: Content validity is the first type of validity that has to be ensured during the design of a tool. Its prerequisite for other validities and its close relationship with the tool reliability place it in a vital position in the questionnaire design. However, since it has not been paid enough attention in national publications therefore in this paper we will explain the different steps in preparing an appropriate content for questionnaires and also their assessment.

Methods: The steps of the content validation process for designing valid and reliable questionnaires have been explained in this paper: development (domain identification, items generation and instrument construction), Judgment – quantification (determining the inter rater agreement among experts, determining the clarity and relevancy of each of the questions along with the clarity, relevancy and comprehensiveness of the tool as a whole) have been explained.

Conclusions: The benefits of performing the content validation in questionnaire development are not increasing the possibility of accessing desirable content validity indexes (relevancy, clarity, and comprehensiveness) only, but also improving the Reliability indexes of tool and decreasing need for resources (financial and temporal) are the other benefits of this approach. Although subjective nature of experts' responses is the problematic issue in this process but the more quantitative steps of the content validation has reduced considerable part of previous worries about this process. These issues explain why content validation has become a suitable process for questionnaire

Keywords: Content validity, Questionnaire development, Content validation process