

روایی سازه پرسشنامه WHOQOL-100 با استفاده از تحلیل رایش

الهام صانعی پور*، مسعود کریملو**، عنایت‌الله بخشی***، کامران یزدانی****

مقدمه: با توجه به اهمیت اندازه‌گیری کیفیت زندگی به عنوان یک پیامد مهم برای ارزیابی وضعیت سلامت همچنین ارزیابی نحوه پاسخ‌گویی به مداخلات، لزوم بررسی ویژگی‌های روان‌سنجی پرسشنامه WHOQOL-100 مشهود است. در این مطالعه، روایی سازه پرسشنامه WHOQOL-100 با استفاده از تحلیل رایش بررسی شده است.

روش: جمعیت مورد بررسی در این مطالعه، ۵۰۰ نفر از معلولان جسمی-حرکتی مراجعه‌کننده به مرکز جامع توانبخشی هلال احمر تهران بودند. افراد شرکت‌کننده در این مطالعه پرسشنامه WHOQOL-100 را، که شامل ۶ بعد از کیفیت زندگی است، تکمیل نمودند و سپس داده‌ها با استفاده از تحلیل رایش تحلیل شدند. در این آنالیز، ابتدا بر روی هر یک از ابعاد، یکی از مدل‌های خانواده رایش به نام مدل اعتبار جزئی برازش داده شد و سپس مناسب بودن سئوال‌ها با استفاده از آماره‌های استاندارد برازش سئوال و آنالیز اختلاف عملکرد سئوال بررسی شد.

یافته‌ها: پس از حذف ۴ سئوال کلی و برازش مدل برای هر یک از ابعاد به صورت جداگانه، مشخص شد که ۴ سئوال از ۱۲ سئوال بعد سلامت جسمانی، ۱۴ سئوال از ۲۰ سئوال بعد سلامت روانی، ۱۱ سئوال از ۱۶ سئوال بعد سطح استقلال، ۷ سئوال از ۱۲ سئوال بعد ارتباطات اجتماعی، ۱۵ سئوال از ۳۲ سئوال بعد سلامت محیط و ۲ سئوال از ۴ سئوال بعد مذهب، وضعیت روحی و عقاید شخصی، با توجه به آماره‌های استاندارد برازش سئوال و آنالیز اختلاف عملکرد سئوال، روایی سازه پرسشنامه را دچار مشکل می‌کنند و در نتیجه از پرسشنامه کنار گذاشته شدند. پس از کنارگذاشتن این سئوال‌ها نامناسب، شاخص‌های پایایی افتراق افراد و آلفای کرونباخ برای ابعاد ۶گانه به ترتیب در فاصله (۰/۶۴۴-۰/۹۰۰) و (۰/۵۷۶-۰/۸۷۸) قرار گرفتند.

بحث: سئوال‌های باقی مانده در هر یک از ابعاد پرسشنامه WHOQOL-100 برازش مناسبی با مدل نشان می‌دهند که بر این اساس می‌توان گفت پرسشنامه دارای روایی سازه مناسب است. شاخص افتراق افراد که یک شاخص پایایی است و با استفاده از تحلیل رایش به دست می‌آید، برای تمام ابعاد مقداری بزرگ‌تر از شاخص آلفای کرونباخ به دست آمده است که نشان دهنده برآورد دقیق‌تر شاخص کیفیت زندگی با استفاده از تحلیل رایش است.

کلید واژه‌ها: تحلیل رایش، روایی سازه، کیفیت زندگی، نظریه سئوال - پاسخ

تاریخ دریافت: ۹۱/۲/۲۴ تاریخ پذیرش: ۹۲/۱۲/۲۵

* کارشناس ارشد آمار حیاتی، دانشگاه علوم بهزیستی و توانبخشی، تهران، ایران
** دکترای آمار حیاتی، دانشگاه علوم بهزیستی و توانبخشی، تهران، ایران، <mkarimlo@yahoo.com> (نویسنده مسئول)
*** دکترای آمار حیاتی، دانشگاه علوم بهزیستی و توانبخشی، تهران، ایران
**** دکترای اپیدمیولوژی، دانشگاه علوم پزشکی تهران

مقدمه

در چند دهه اخیر شیوع استفاده از شاخص کیفیت زندگی یا کیفیت زندگی وابسته به سلامت^۱ به عنوان یک پیامد مهم برای ارزیابی وضعیت سلامت و همچنین ارزیابی نحوه پاسخ‌گویی به مداخلات، سیر صعودی داشته است. به منظور اندازه‌گیری کیفیت زندگی مانند هر متغیر پنهان^۲ دیگری، نیاز به مجموعه‌ای از سؤال‌هاست تا بر اساس عملکرد فرد در پاسخ به آن‌ها بتوان میزان این متغیر پنهان را برای هر فرد اندازه‌گیری نمود و در مورد آن قضاوت کرد.

سازمان جهانی سلامت^۳ (WHO) برای ارزیابی کیفیت زندگی، ابزاری را تحت عنوان WHOQOL-100 ارائه داده است. این ابزار ۶ بعد^۴ یا خرده‌مقیاس دارد که عبارتند از: سلامت جسمانی، سلامت روان، سطح استقلال، ارتباطات اجتماعی، سلامت محیط و نهایتاً بعد مذهب و وضعیت روحی و عقاید شخصی. هر یک از این ابعاد چندین زیرمجموعه یا جزء^۵ را شامل می‌شوند. به طوریکه در مجموع ۲۴ زیرمجموعه وجود دارد و هر زیرمجموعه توسط ۴ سؤال سنجیده می‌شود، از طرفی، ۴ سؤال نیز سلامت و کیفیت کلی زندگی را می‌سنجند. بنابراین، در کل ۱۰۰ سؤال در پرسشنامه طراحی شده است که برای پاسخ هر سؤال از مقیاس لیکرت ۵ گزینه‌ای استفاده شده است.

با توجه به اهمیت ارزیابی کیفیت زندگی، بررسی ویژگی‌های روان‌سنجی پرسشنامه‌هایی که این شاخص را می‌سنجند، مورد تأکید قرار گرفته است. بررسی ویژگی‌های روان‌سنجی پرسشنامه WHOQOL-100 به روش تئوری آزمون کلاسیک^۶ انجام شده است، اما در ایران تا کنون بررسی این ویژگی‌ها با روش‌های مدرن، مانند تحلیل راس، به ندرت انجام شده است (کریملو، زائری و صالحی، ۲۰۱۱). معمولاً استفاده از

1. health related quality of life

2. latent variable

3. World Health Organization

4. dimension

5. Facet

6. classical test theory

آنالیزهای کلاسیک، که در آن‌ها نمره‌های خام، ترکیبات خطی این نمره‌ها و پاسخ‌هایی که در مقیاس ترتیبی می‌باشند و در اصل ماهیت رتبه ای دارند، به عنوان داده‌هایی در مقیاس فاصله ای در نظر گرفته می‌شوند، با مشکلاتی همراه است. آنالیزهای مدرن راه حل‌هایی را برای مواجه نشدن با این مشکلات ارائه می‌دهند. تحلیل راش یک تکنیک آماری است که می‌تواند در سؤال‌هایی که دو یا بیش از دو حالت برای پاسخ‌گویی دارند، برای تبدیل داده‌های ترتیبی به داده‌های فاصله ای به کار برده شود (ویلسون و هاسکینز^۱، ۲۰۰۵).

تحلیل‌های استاندارد راش بر اساس مدل‌های تک‌بعدی^۲ انجام می‌شوند. در مدل‌های تک‌بعدی فرض بر این است که یک و فقط یک ویژگی پنهان از شخص، تعیین‌کننده عملکرد شخص روی پرسشنامه می‌باشد. اگر داده‌ها به خوبی با مدل راش برازش نداشتند، فرض تک‌بعدی بودن رد می‌شود. به این معنی که بیش از یک ویژگی پنهان روی عملکرد شخص تأثیر داشته است و بنابراین ویژگی مورد نظر را با استفاده از این پرسشنامه نمی‌توان به خوبی اندازه‌گیری کرد (وانگ و دیگران^۳، ۲۰۰۶).

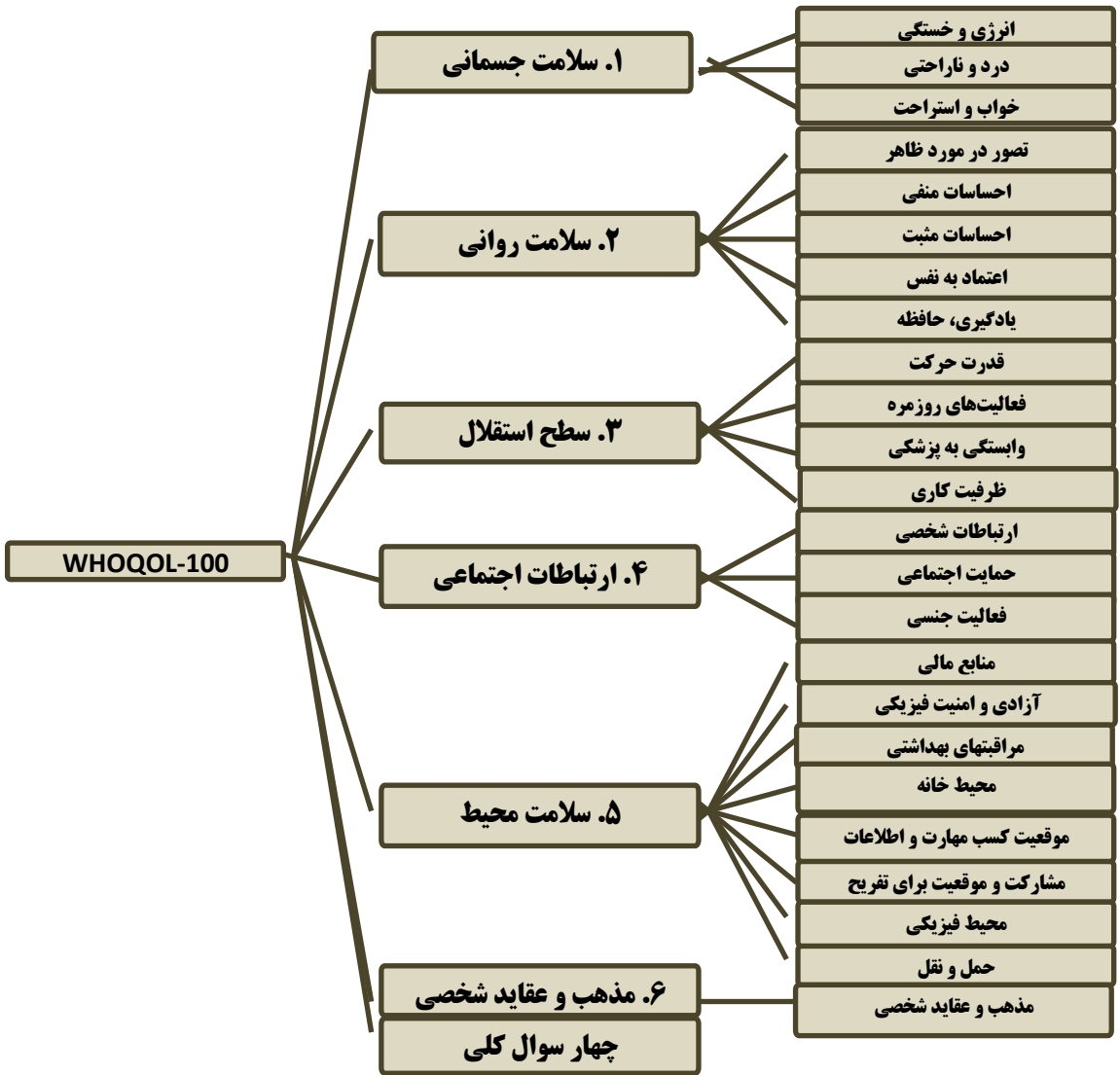
در پرسشنامه‌هایی همچون WHOQOL-100 که شامل چندین خرده‌مقیاس می‌باشند، می‌توان از تحلیل‌های استاندارد راش استفاده کرد. در این صورت می‌توان یک مدل تک‌بعدی روی کل پرسش‌های موجود در پرسشنامه اعمال کرد یا این‌که از مدل‌های راش تک‌بعدی برای هر یک از خرده‌مقیاس‌ها به صورت جداگانه استفاده کرد (همان). هدف از انجام این مطالعه، بررسی روایی سازه^۴ پرسشنامه WHOQOL-100 با استفاده از تحلیل‌های تک‌بعدی راش است که توسط نرم افزار ConQuest انجام می‌شود.

-
1. Wilson and Hoskens
 2. unidimensional
 3. Wang
 4. Construct validity

روش

جامعه مورد بررسی در این مطالعه، جامعه معلولین جسمی-حرکتی می‌باشند که برای نمونه گیری از آن‌ها، ابتدا با هماهنگی به عمل آمده با سازمان بهزیستی استان تهران، به مراکز نگهداری این گونه معلولین که تعداد آن‌ها در شهر تهران انگشت شمار می‌باشد، مراجعه شد. ولی عملاً همکاری پایین معلولین این مراکز در تکمیل پرسشنامه‌ها موجب شد که پس از چند ماه پیگیری مستمر و بی نتیجه، از مراجعه کنندگان به مرکز جامع توانبخشی هلال احمر تهران که یک مرکز چندمنظوره برای مراجعه معلولین کشور می‌باشد، در این زمینه استفاده شود. معلولیت‌هایی که در این مطالعه مد نظر بودند شامل معلولیت در اندام‌های فوقانی، اندام‌های تحتانی، ستون فقرات، استخوان بندی و عضلات می‌باشند. به این دلیل که مرکز جامع توانبخشی هلال احمر یک مرکز مرجع برای مراجعه معلولین سطح شهر تهران و حتی کشور است، به نوعی می‌تواند نمونه‌ای از معلولین کشور تلقی شود. نمونه گیری از این افراد به صورت در دسترس و غیراحتمالی و با حجم نمونه ۵۰۰ نفر از معلولین واجد شرایط انجام شد.

از افراد تحت مطالعه خواسته شد که به سؤالات پرسشنامه WHOQOL-100 بر اساس وضعیتشان در دو هفته گذشته پاسخ دهند و پرسشنامه‌ها توسط پرسشگرها تکمیل شد. این پرسشنامه شامل ۱۰۰ سؤال است که ۴ سؤال آن سطح کلی کیفیت زندگی را می‌سنجند. ۹۶ سؤال باقی‌مانده برای اندازه گیری ۶ بعد از کیفیت زندگی مورد استفاده قرار می‌گیرند. هر یک از این ابعاد شامل یک یا چند خرده‌مقیاس است که هر یک توسط ۴ سؤال سنجیده می‌شوند. به طور کلی ۶ بعد وجود دارد که توسط ۲۴ خرده‌مقیاس پوشش داده شده‌اند (شکل ۱). از آنجایی که ۴ سؤال کلی به هیچ یک از ابعاد تعلق ندارند، از پرسشنامه کنار گذاشته شدند.



شکل (۱) نمای گرافیکی پرسشنامه WHOQOL-100

مدلی که در این مطالعه از آن استفاده شد، مدل لجیت چندجمله ای با ضرایب تصادفی^۱ نام دارد که آن را اصطلاحاً RCMLM می نامند. در این مطالعه اندازه متغیر پنهان را برای شخص n با θ_n نشان می دهیم. همچنین فرض می کنیم افراد از جامعه ای با تابع توزیع $g(\theta_n, \alpha)$ نمونه گیری شده اند و α برداری از پارامترهاست که نشان دهنده ویژگی های توزیع جامعه می باشد. در این مطالعه فرض می شود g تابع چگالی توزیع نرمال است. بنابراین خواهیم داشت $\alpha \equiv (\mu, \sigma^2)$.

مدل RCMLM احتمال این که پاسخ شخص n با سطح توانایی θ_n برای سؤال i م رده z باشد، را این گونه محاسبه می کند:

$$p_{nij} = \frac{\exp(b_{ij}^T \theta_n + a_{ij}^T \xi)}{\sum_{u=1}^{k_i} \exp(b_{iu}^T \theta_n + a_{iu}^T \xi)} \quad (1)$$

که k_i نشان دهنده تعداد رده های موجود در سؤال i است (در این مطالعه برای هر سؤال داریم $k_i=5$)؛ ξ برداری از پارامترهاست که سؤال را توصیف می کند؛ b_{ij} عددی است که برای رده z از سؤال i نمره ای را در نظر می گیرد. اگر نمره های مربوط به کل سؤال ها کنار هم قرار بگیرند، بردار به دست آمده بردار نمره ها^۲ نامیده و با B نشان داده می شود و a_{ij} که آن را بردار طراحی رده z از سؤال i می نامیم، برداری است که ارتباط خطی میان اعضای ξ را توصیف می کند و اگر این بردارها برای کل سؤال های موجود کنار هم قرار داده شوند، ماتریس به دست آمده ماتریس طراحی^۳ نامیده و با A نشان داده می شود.

-
1. Random Coefficient Multinomial Logit Model
 2. scoring vector
 3. design matrix

می‌توان از معادله ۱، معادله ۲ را به صورت زیر به دست آورد (آدامز، ویلسون و وانگ^۱، ۱۹۹۷).

$$\log\left(\frac{P_{nij}}{P_{ni(j-1)}}\right) = (b_{ij}^T - b_{i(j-1)}^T)\theta_n + (a_{ij}^T - a_{i(j-1)}^T)\xi \quad (2)$$

$$\equiv b_{ij}^{*T}\theta_n + a_{ij}^{*T}\xi$$

استفاده از a_{ij} و b_{ij} یا به طور معادل استفاده از a_{ij}^* و b_{ij}^* برای تعریف ارتباط میان سؤال‌ها و افراد، برای ما یک مدل کلی فراهم می‌کند که می‌تواند بسیاری از مدل‌های تک‌بعدی از جمله مدل مقیاس درجه بندی^۲ و مدل اعتبار جزئی^۳ را فراهم کند (آندریچ^۴، ۱۹۷۸؛ مسترز^۵، ۱۹۸۲).

از میان مدل‌های موجود در خانواده^۶ راش، مدل‌های اعتبار جزئی و مقیاس درجه بندی که می‌توانند خاصیت رتبه ای بودن پاسخ‌ها را در سؤال‌هایی که بیش از دو حالت برای پاسخ‌گویی دارند حفظ کنند، برای سؤال‌هایی که از مقیاس لیکرت پیروی می‌کنند مناسب‌ترند.

در مطالعه حاضر، هر یک از ۶ بعد موجود در پرسشنامه WHOQOL-100، یک بعد از کیفیت زندگی را می‌سنجد، بنابراین ۶ بعد از کیفیت زندگی توسط ۶ مقیاس پرسشنامه WHOQOL-100 سنجیده می‌شوند. از آنجایی که رده‌های موجود برای پاسخ دهی در تمام سؤال‌ها یکسان نیست، از مدل اعتبار جزئی به جای مدل مقیاس درجه بندی استفاده کرده‌ایم.

-
- | | |
|----------------------------|-----------------------|
| 1. Adams, Wilson, and Wang | 2. rating scale model |
| 3. partial credit model | 4. Andrich |
| 5. Masters | |

مدل اعتبار جزئی لگاریتم شانس انتخاب رده j به $j-1$ را در سؤال i و برای فرد n که دارای سطح متغیر پنهان θ_n می باشد، این گونه تعریف می کند:

$$\log\left(\frac{P_{nij}}{P_{ni(j-1)}}\right) = \theta_n - (\delta_i + \tau_{ij}) \quad (۳)$$

که در آن δ_i سختی کلی^۱ سؤال i نامیده می شود و τ_{ij} پارامتر مرز رده^۲ یا گام^۳ نامیده می شود.

برای مدل مقیاس درجه بندی معادله^۳ به معادله^۴ تقلیل می یابد،

$$\log\left(\frac{P_{nij}}{P_{ni(j-1)}}\right) = \theta_n - (\delta_i + \tau_i) \quad (۴)$$

که در آن پارامترهای گام برای تمام سؤال های موجود در پرسشنامه یکسان در نظر گرفته می شوند (آدامز و دیگران^۴، ۱۹۹۷).

با مقایسه معادله های ۳ و ۴ با معادله^۲، به راحتی در می یابیم که مدل های اعتبار جزئی و مقیاس درجه بندی حالت های خاصی از RCMLM می باشند.

RCMLM عضوی از خانواده توزیع های نمایی است که می تواند به عنوان یک مدل خطی آمیخته تعمیم یافته در نظر گرفته شود. از چندین برنامه کامپیوتری می توان برای برآورد پارامترهای این مدل استفاده کرد. از آن جمله می توان به ConQuest (وو، آدامز و

-
- | | |
|-----------------------|-------------------------|
| 1. overall difficulty | 2. categorical boundary |
| 3. step | 4. Adams |

ویلسون^۱، (۱۹۹۸)، SAS NLMIXED (وولفینگر^۲، ۱۹۹۹)، STATA (اسکروندال، راب-هسکیچ و بوکاراتون^۳، ۲۰۰۴)، MIXOR (هدیکر و گیونز^۴، ۱۹۹۶) و MIXNO (هدیکر، ۱۹۹۹) اشاره کرد. در این مطالعه، از نسخه دوم نرم افزار ConQuest استفاده شده است.

برای بررسی مناسب بودن سؤال‌ها، دو روش بررسی آماره برازش سؤال‌ها و آنالیز اختلاف عملکرد سؤال^۵ (DIF) مورد استفاده قرار گرفته‌اند. برای بررسی برازش سؤال‌ها، از دو شاخص outfit MNSQ و infit MNSQ استفاده شده است. در outfit MNSQ، باقی‌مانده‌ها مستقیماً با هم جمع می‌شوند و این شاخص نسبت به پاسخ‌های غیرمنتظره در سؤال‌هایی که میان پارامتر آن‌ها و پارامتر شخص اختلاف زیاد وجود دارد، حساس‌اند. اما در محاسبه infit MNSQ، به داده‌ها قبل از جمع‌شدن وزن داده می‌شود و آماره به‌دست آمده نسبت به پاسخ‌های غیرمنتظره در سؤال‌هایی که پارامتر آن‌ها به پارامتر شخص نزدیک است، حساس است. زمانی که داده‌ها با مدل برازش داشته باشند، مقدار مورد انتظار برای هر یک از این دو پارامتر برابر یک است. اندازه MNSQ میزان انحراف از سیستم اندازه‌گیری را نشان می‌دهد. مقادیر خیلی کم‌تر از یک برای این شاخص‌ها در یک سؤال نشان دهنده این است که پاسخ‌ها برای این سؤال بیش از حد معمول توسط مدل قابل پیش‌بینی است و مقادیر خیلی بزرگ‌تر از یک برای این شاخص‌ها گویای این است که پاسخ‌های داده شده به سؤال توسط مدل قابل پیش‌بینی نیست. از لحاظ آماری، MNSQ آماره‌ی کای-دویی است که بر درجه آزادی تقسیم شده است. برای مقیاس‌های درجه بندی، فاصله (۰/۶-۱/۴) به عنوان فاصله‌ی معیار برای MNSQ پیشنهاد شده است. به سؤال‌هایی که آماره‌های outfit MNSQ یا infit MNSQ برای آن‌ها خارج از این فاصله

1. Wu, Adams and Wilson

2. Wolfinger

3. Skronidal, Rabe-Hesketh and Boca Raton

4. Hedeker and Gibbons

5. Differential Item Functioning

قرار بگیرد، به عنوان سؤال‌هایی که برازش مناسبی با مدل ندارند، برخورد می‌شود (رایت، لیناگر، گوستافسون و مارتین-ال^۱، ۱۹۹۴). در این مطالعه، اندکی سخت‌گیرانه‌تر عمل شده و از فاصله (۰/۷-۱/۳) به عنوان فاصله معیار استفاده شده است.

یکی از معیارهایی که در ساختن آزمون‌های خودسنج مطرح می‌شود، داشتن آزمونی است که برای گروه‌های مختلف قابل استفاده باشد. از آنجایی که افراد با ویژگی‌های متفاوت می‌توانند برداشت‌های متفاوتی از سؤال‌ها و رده‌ها داشته باشند این موضوع اهمیت پیدا می‌کند. آنالیز DIF آنالیزی است که عملکرد یکسان سؤال‌ها را در گروه‌های مختلف بررسی می‌کند. اگر یک سؤال عملکرد یکسانی در گروه‌های مختلف نداشته باشد، به این معنا است که گروه‌های مختلف دید متفاوتی نسبت به این سؤال دارند و بنابراین اندازه‌های به دست آمده برای افراد در گروه‌های متفاوت به طور مستقیم قابل مقایسه نیستند. برای مثال، اگر برخی از سؤال‌های WHOQOL-100 برای گروه‌های سنی متفاوت عملکرد یکسانی نداشته باشند، اندازه‌های حاصل از کاربرد این پرسشنامه برای گروه‌های سنی مختلف قابل مقایسه نمی‌باشند. برای داشتن اندازه‌های قابل مقایسه بین گروه‌های مختلف، باید تمام سؤال‌ها فاقد DIF باشند یا حداقل مقدار کمی DIF را نشان دهند.

از لحاظ آماری، سؤال‌ی دارای DIF محسوب می‌شود که احتمال پاسخ برای آن سؤال به طور کامل توسط متغیر پنهان و پارامترهای سختی سؤال قابل توضیح نباشد. آنالیز DIF سؤال‌هایی را که پس از کنترل سطح توانایی، به نظر می‌رسد برای بعضی گروه‌ها خیلی سخت و برای بعضی دیگر خیلی آسان می‌باشند، را شناسایی می‌کند (هولند و واینر، واینر^۲، ۱۹۹۳). برای افراد شرکت کننده در این مطالعه، آنالیز DIF برای ۳ ویژگی دموگرافیک مهم جنسیت (زن و مرد)، تحصیلات (ابتدایی، دیپلم، دانشگاهی) و سن (۱۵-

1. Wright, Linacre, Gustafson and Martin-L
2. Holland and Wainer

۲۴، ۲۵-۳۴، ۳۵-۴۴، ۴۵-۵۴، ۵۵-۶۴، ۶۵) انجام شد. سختی کلی سئوال‌ها بین دو گروه زن و مرد، بین ۳ سطح تحصیلات و بین ۶ گروه سنی مورد بررسی قرار گرفت. سئوال‌هایی که حداقل برای یکی از این متغیرها، در دو گروه اختلافی بزرگ‌تر از ۰/۵ نشان دادند، دارای DIF تشخیص داده شده و برای آنالیزهای بعدی کنار گذاشته شدند.

برای حل مشکل عدم تعیین^۱ در آنالیز DIF، میانگین پارامترهای سختی کلی سئوال‌ها، برای هر یک از گروه‌های یک متغیر، برابر صفر تنظیم شده است. به این صورت پارامترهای سئوال در گروه‌های مختلف می‌توانند به صورت مستقیم با هم مقایسه شوند. در واقع با برابر قراردادن میانگین پارامترهای سئوال در گروه‌های مختلف، اثری که اختلاف موجود در سطح توانایی گروه‌های مختلف می‌تواند روی آنالیز DIF داشته باشد، حذف می‌شود (رایت و استون^۲، ۱۹۷۹).

در این مطالعه، اختلافات برابر یا بزرگ‌تر از ۰/۵ میان پارامترهای برآورد شده برای گروه‌های مختلف به عنوان وجود DIF تلقی شده است. هر گاه سئوالی دارای DIF تشخیص داده شده، آن سئوال حذف شده، سپس مدل اعتبار جزئی روی سئوال‌های باقی مانده برازش داده شده است. آنالیزها زمانی متوقف شده اند که برای تمام سئوال‌ها آماره‌های outfit MNSQ و infit MNSQ در فاصله (۰/۷-۱/۳) قرار گرفته اند و هیچ یک از سئوال‌ها هم DIF نشان نداده است.

یکی از مهمترین شاخص‌هایی که در پرسش‌نامه‌ها مطرح می‌شود شاخص پایایی است. در روش‌های کلاسیک برای بررسی پایایی آزمون از شاخص آلفای کرونباخ استفاده می‌شود. اما در روش‌های مدرن، که شامل تحلیل رانش نیز می‌شوند، شاخصی تحت عنوان شاخص تمیز افراد برای بررسی پایایی مورد محاسبه قرار می‌گیرد. هر دوی این شاخص‌ها

1. indeterminacy

2. Wright and Stone

با فرمولی مشابه و برپایه خطاهای استاندارد برآورد خصیصه محاسبه می‌شوند. از آنجایی که تحلیل رانش برپایه مدل عمل می‌کند، انتظار می‌رود برآوردهای دقیق‌تری برای خصیصه فراهم کند و در نتیجه شاخص پایایی بزرگ‌تری نسبت به روش کلاسیک فراهم کند.

یافته‌ها

از نظر توزیع جمعیت تحت مطالعه، ۵۷/۶٪ پاسخ‌گویان مرد و ۴۲/۴٪ زن بودند، بیش‌تر آن‌ها در گروه‌های سنی ۳۵-۴۴ سال (۲۸/۲٪) و ۱۵-۲۴ سال (۲۳/۸٪) قرار داشته و بیش از نیمی از آن‌ها (۵۳/۸٪) از تحصیلات راهنمایی تا دیپلم برخوردار بودند (جدول ۱).

جدول (۱) توزیع فراوانی جمعیت تحت مطالعه برای سه متغیر جنس، سن و تحصیلات

متغیر	جنسیت		سن							تحصیلات		
			۱۵-۲۴	۲۵-۳۴	۳۵-۴۴	۴۵-۵۴	۵۵-۶۴	۶۵≤	ابتدایی			راهنمایی
گروه	زن	مرد										
f	۲۱۲	۲۸۸	۱۱۹	۱۱۰	۱۴۱	۸۳	۳۸	۹	۳۶	۲۶۹	۱۹۵	
%	۴۲/۴	۵۷/۶	۲۳/۸	۲۲	۲۸/۲	۱۶/۶	۷/۶	۱/۸	۷/۲	۵۳/۸	۳۹	

پس از این‌که مدل برای ۱۲ سؤال موجود در بعد سلامت جسمانی برازش داده شد، آنالیز DIF برای بررسی تفاوت سختی سئوال‌ها در زیر گروه‌های مختلف متغیرهای جنس، سن و سطح تحصیلات انجام شد و سئوال‌هایی که، حداقل در یکی از متغیرهای زمینه‌ای مورد بررسی، تفاوت بیش‌تر از ۰/۵ را در دو سطح از متغیر نشان دادند حذف شدند. در مرحله اول، ۳ سؤال از ۱۲ سؤال موجود برای متغیر سن DIF نشان داده و حذف شدند. این بار مدل روی ۹ سؤال باقی مانده برازش داده شد و مشاهده شد که فقط یکی از سئوال‌ها، باز هم در متغیر سن، DIF نشان داد. پس از حذف این سؤال و برازش مجدد

مدل روی ۸ سؤال باقی مانده، دیده شد که هیچ یک از سؤال‌ها DIF نشان ندادند. علاوه بر این بررسی شاخص‌های outfit MNSQ و infit MNSQ نیز نشان داد که این دو شاخص برای تمام سؤال‌ها در فاصله (۰/۷-۱/۳) قرار دارند. بنابراین می‌شود گفت از ۱۲ سؤال موجود، ۸ سؤال باقی مانده برازش مناسبی با مدل دارند و می‌شود از آن‌ها برای اندازه‌گیری بعد سلامت جسمانی استفاده کرد.

برای سایر ابعاد نیز به همین ترتیب آنالیز DIF انجام شد و شاخص‌های outfit MNSQ و infit MNSQ مورد بررسی قرار گرفتند. مشاهده شد که برای ابعاد سلامت روانی، سطح استقلال، ارتباطات اجتماعی، سلامت محیط و بعد مذهب و وضعیت روحی و عقاید شخصی به ترتیب ۶، ۵، ۵، ۱۷ و نهایتاً ۲ سؤال برازش مناسبی با مدل نشان دادند. به طور کلی، با کاربرد آنالیزهای outfit MNSQ و infit MNSQ و آنالیز DIF، ۵۳ سؤال به دلیل عدم برازش از پرسشنامه WHOQOL-100 حذف شدند.

در جدول (۲)، بیش‌ترین مقادیر اختلاف در پارامتر سختی سؤال برای زیرگروه‌های مختلف متغیرهای سن، جنس و سطح تحصیلات و همچنین شاخص‌های outfit MNSQ و infit MNSQ برای سؤال‌های باقی مانده در هر یک از ابعاد پرسشنامه دیده می‌شوند. طبق این جدول، حداکثر اختلاف میان گروه‌ها در هر یک از متغیرها از ۰/۵ کم‌تر است، از طرفی مقادیر شاخص‌های outfit MNSQ و infit MNSQ برای هر یک از این سؤال‌ها در بازه (۰/۷-۱/۳) قرار گرفته است، پس می‌توانیم بگوییم سؤال‌های باقی مانده انتظارات مدل را برآورده کرده‌اند. به این معنا که ۹۶ سؤال در فرم اولیه این پرسشنامه، بیش از ۶ بعد را پوشش می‌دادند. وقتی سؤال‌های نامناسب حذف شدند، سؤال‌های باقی مانده توانستند ۶ بعدی را که در پرسش‌نامه مطرح شدند، پوشش دهند.

جدول (۲) بیشترین مقدار اختلاف در پارامترهای سؤالها در زیرگروههای جنس، سن و تحصیلات در سؤالهای باقی مانده همچنین شاخصهای **infit MNSQ** و **outfit MNSQ** برای این سؤالها

infit	Outfit	تحصیلات	سن	جنس	سؤال	بعد
۱/۰۰	۱/۰۲	۰/۰۸۶	۰/۴۷۳	۰/۰۲۰	۴. آیا زود خسته می شوید؟	۱. سلامت جسمانی
۰/۹۵	۰/۹۴	۰/۰۸۸	۰/۱۸۷	۰/۱۴۴	۵. آیا خستگی شما را آزار داده است؟	
۰/۹۳	۰/۹۴	۰/۲۴۲	۰/۴۳۲	۰/۰۰۴	۶. آیا هیچ گونه مشکلی در خوابیدن داشته اید؟	
۰/۹۲	۰/۸۹	۰/۳۰۸	۰/۲۶۶	۰/۱۵۶	۷. آیا مشکل خواب شما را نگران کرده است؟	
۱/۰۳	۱/۰۳	۰/۱۰۶	۰/۲۷۲	۰/۰۵۶	۴۰. آیا برای زندگی روزانه تان انرژی کافی داشته اید؟	
۱/۰۳	۱/۰۳	۰/۳۲۶	۰/۲۲۳	۰/۱۵۰	۵۶. از مقدار انرژی که داشته اید چقدر راضی بوده اید؟	
۱/۰۹	۱/۰۸	۰/۲۹۸	۰/۱۴۵	۰/۰۰۴	۸۳. چقدر خوب می خوابید؟	۲. سلامت روانی
۱/۰۸	۱/۰۹	۰/۲۶۸	۰/۳۱۹	۰/۱۲۲	۸۶. آیا هر چند وقت یک بار از درد جسمی رنج برده اید؟	
۱/۱۸	۱/۱۶	۰/۱۵۰	۰/۳۹۱	۰/۱۰۲	۹. آینده را مثبت دیده اید؟	
۱/۰۲	۱/۰۱	۰/۲۴۲	۰/۳۴۰	۰/۰۰۲	۱۰. از زندگی تان چقدر احساس مثبت داشته اید؟	
۱/۲۲	۱/۲۷	۰/۱۶۱	۰/۳۷۱	۰/۰۵۴	۱۵. آیا هیچ بخشی از زندگی تان موجب ناراحتی تان شده است؟	
۰/۸۷	۰/۸۵	۰/۱۲۵	۰/۲۳۵	۰/۰۹۴	۱۷. آیا مکان زندگی تان راحت بوده است؟	
۰/۸۵	۰/۸۳	۰/۳۶۴	۰/۲۶۱	۰/۰۹۸	۱۸. آیا احساس افسردگی شما را آزار داده است؟	۳. سطح استقلال
۰/۹۴	۰/۹۴	۰/۱۶۵	۰/۳۵۰	۰/۰۴۰	۸۸. هر چند وقت یک بار احساسات منفی همچون غم، یأس، اضطراب و افسردگی داشته اید؟	
۰/۹۴	۰/۹۶	۰/۰۶۵	۰/۳۹۵	۰/۰۲۶	۲۲. برای انجام فعالیت های روزمره خود چقدر به درمان پزشکی نیاز داشته اید؟	
۰/۹۵	۰/۹۶	۰/۱۰۵	۰/۲۹۲	۰/۱۲۶	۳۳. کیفیت زندگی تان چقدر به مصرف مواد دارویی یا کمک های پزشکی وابسته بوده است؟	
۰/۹۶	۰/۹۶	۰/۳۶۱	۰/۳۴۱	۰/۰۵۲	۹۰. آیا احساس می کردید که توانایی انجام وظایف تان را دارید؟	
۰/۹۸	۰/۹۷	۰/۱۲۰	۰/۴۲۹	۰/۲۲۶	۹۲. چه امتیازی (رتبه ای) به توانایی خود برای انجام کار می دهید؟	
۱/۱۷	۱/۱۸	۰/۳۴۰	۰/۴۸۵	۰/۰۲۲	۹۴. مشکلات مربوط به جابجایی (تحرک) چقدر شما را اذیت کرده است؟	۴. ارتباطات اجتماعی
۱/۰۰	۱/۰۱	۰/۱۱۷	۰/۲۷۹	۰/۲۷۰	۲۴. در زندگی چقدر احساس تنهایی کردید؟	
۱/۱۲	۱/۱۸	۰/۱۱۹	۰/۴۳۱	۰/۱۵۸	۲۶. آیا هیچ مشکلی در زندگی جنسی تان شما را رنجانده است؟	
۱/۰۲	۱/۰۲	۰/۲۶۹	۰/۳۳۲	۰/۰۳۶	۴۵. چقدر توانسته اید در صورت لزوم روی دوستان حساب کنید؟	
۰/۹۱	۰/۹۲	۰/۱۸۴	۰/۱۰۵	۰/۱۱۲	۶۷. چقدر از حمایت و کمکی که از دوستان دریافت می نمایید راضی بوده اید؟	
۰/۹۸	۰/۹۶	۰/۲۱۷	۰/۳۳۲	۰/۰۳۴	۸۰. آیا از روابط خود با اعضای خانواده رضایت داشتید؟	
۰/۹۱	۰/۹۱	۰/۱۵۲	۰/۳۰۸	۰/۰۱۸	۲۷. تا چه حد در زندگی روزمره خود احساس رضایت کردید؟	۵. سلامت محیط
۱/۱۰	۱/۱۴	۰/۲۸۶	۰/۳۸۸	۰/۰۱۴	۲۸. آیا احساس کردید که در یک محیط بی خطر و امن کردید؟	
۰/۹۳	۰/۹۷	۰/۰۴۵	۰/۳۱۳	۰/۰۵۴	۳۰. آیا مکان زندگی تان راحت بوده است؟	
۱/۱۷	۱/۲۶	۰/۲۷۰	۰/۲۲۳	۰/۰۹۶	۳۲. آیا مشکلات مالی داشتید؟	
۱/۰۵	۱/۰۵	۰/۱۷۴	۰/۲۳۴	۰/۰۷۰	۳۴. آیا به راحتی به مراقبت های پزشکی خوب و مناسب دسترسی داشته اید؟	
۰/۹۲	۰/۹۵	۰/۳۸۵	۰/۲۷۵	۰/۰۳۲	۳۵. از وقت آزاد خود چقدر لذت برده اید؟	
۱/۰۳	۱/۰۳	۰/۱۹۰	۰/۳۱۵	۰/۰۶۰	۳۶. محیط فیزیکی زندگی تان تا چه حد سالم و پاک بوده است؟	
۱/۲۹	۱/۳۰	۰/۳۲۶	۰/۱۶۰	۰/۰۴۸	۳۸. در رفت و آمد با وسایل نقلیه چقدر مشکل داشته اید؟	
۱/۱۹	۱/۲۳	۰/۰۹۴	۰/۰۸۰	۰/۰۰۸	۳۹. مشکلات رفت و آمد با وسایل نقلیه چقدر زندگی تان را محدود کرده است؟	
۰/۸۷	۰/۸۹	۰/۲۱۴	۰/۳۹۷	۰/۱۴۰	۴۶. چقدر امکانات و شرایط خانه تان نیازهای شما را برآورده کرده اند؟	
۰/۷۷	۰/۷۷	۰/۲۵۰	۰/۴۰۰	۰/۰۲۰	۴۷. آیا پول کافی برای برآورده کردن نیازهای تان داشته اید؟	
۰/۸۱	۰/۸۲	۰/۴۰۰	۰/۳۲۵	۰/۰۸۸	۴۹. برای بدست آوردن اطلاعات مورد نیاز تان چقدر امکانات داشته اید؟	
۰/۸۳	۰/۸۲	۰/۱۰۲	۰/۲۷۰	۰/۰۲۴	۵۰. برای فعالیت های تفریحی چقدر فرصت و امکان داشته اید؟	
۰/۷۷	۰/۷۷	۰/۱۹۱	۰/۱۶۹	۰/۰۵۶	۵۱. تا چه حد توانسته اید راحت و خوش باشید؟	
۰/۹۰	۰/۹۲	۰/۰۵۸	۰/۲۷۹	۰/۰۰۴	۷۱. چقدر از وضعیت مالی خود راضی بوده اید؟	
۱/۱۹	۱/۲۳	۰/۱۴۴	۰/۳۱۳	۰/۱۲۸	۷۳. چقدر از وجود یا دریافت خدمات حمایت های اجتماعی راضی بوده اید؟	
۱/۲۴	۱/۲۶	۰/۳۹۹	۰/۲۱۳	۰/۱۴۶	۸۵. به کیفیت خدمات حمایت های اجتماعی در دسترس تان چه امتیازی می دهید؟	
۰/۹۹	۰/۹۷	۰/۱۶۶	۰/۱۹۱	۰/۰۲۲	۹۷. آیا اعتقادات شخصی تان به زندگی شما معنی و مفهوم بخشیده اند؟	۶. مذهب و وضعیت روحی و عقاید شخصی
۱/۰۰	۰/۹۹	۰/۱۶۶	۰/۱۹۱	۰/۰۲۲	۹۹. چقدر اعتقادات شخصی تان به شما در درک و فهم مشکلات زندگی کمک کرده است؟	

با استفاده از اندازه‌های برآورد شده برای افراد همچنین خطای معیار این برآوردها شاخصی به نام شاخص تمیز افراد^۱ (PSI) قابل محاسبه است. این شاخص یک شاخص پایایی است که اندازه آن میان صفر و یک قرار می‌گیرد و مقادیر نزدیک به یک، مقادیر مطلوبی برای این شاخص می‌باشند. همانطور که در جدول (۳) مشاهده می‌شود، مقادیر PSI برای تمام ابعاد از مقادیر آلفای کرونباخ متناظر با آن بیش تر است.

جدول (۳) شاخص‌های آلفای کرونباخ و PSI برای ابعاد

بعد	تعداد سوال	α	PSI
۱. سلامت جسمانی	۸	۰/۸۴۳	۰/۸۶۰
۲. سلامت روانی	۶	۰/۸۰۳	۰/۸۴۶
۳. سطح استقلال	۵	۰/۷۲۰	۰/۷۵۲
۴. ارتباطات اجتماعی	۵	۰/۵۷۶	۰/۶۴۴
۵. سلامت محیط	۱۷	۰/۸۷۸	۰/۹۰۰
۶. مذهب، وضعیت روحی و عقاید شخصی	۲	۰/۸۰۲	۰/۸۳۱

بحث

در برخورد با پرسشنامه WHOQOL-100 جنبه‌های متعارف به کار بسته شده و با این پرسشنامه به عنوان پرسشنامه‌ای با ۶ بعد برخورد شده است. پس از برازش مدل‌های جداگانه برای هر یک از ابعاد و کاربرد آنالیزهای *infit MNSQ* و *outfit MNSQ* و آنالیز DIF، ۵۳ سوال برای روایی سازه پرسشنامه نامناسب تشخیص داده شده و حذف شدند. از این ۵۳ سوال به ترتیب ۴، ۱۴، ۱۱، ۷، ۱۵ و ۲ سوال به ابعاد سلامت جسمانی، سلامت

1 Person Separation Index

روانی، سطح استقلال، ارتباطات اجتماعی، سلامت محیط و در نهایت بعد مذهب و وضعیت روحی متعلق بودند. پس از حذف این سئوال‌ها می‌توان ادعا کرد پرسشنامه از روایی سازه مناسب برخوردار است. از طرفی مشاهده شد با کاربرد تحلیل راش به عنوان روشی که برپایه مدل عمل می‌کند، شاخص پایایی پرسشنامه، برای هر یک از ابعاد در مقایسه با روش کلاسیک مقدار بزرگ‌تری به دست آمد که این موضوع گواهی بر برآوردهای دقیق‌تر خصیصه پنهان در تحلیل راش است.

دلیل حذف سئوال‌هایی که عدم برازش نشان داده اند، این نیست که این سئوال‌ها جنبه‌هایی از کیفیت زندگی را نشان نمی‌دهند، بلکه علت این است که این سئوال‌ها ابعادی را که ۴۳ سئوال دیگر نشان می‌دهند، ارزیابی نمی‌کنند. از طریق آنالیز DIF می‌شود نسبت به، دلیل عدم برازش سئوال‌ها درک دقیق‌تری پیدا کرد. این درک دقیق‌تر می‌تواند منجر به تجدیدنظر در سئوال‌ها یا حتی ساختن پرسشنامه‌های جدید شود. برای مثال، حذف سئوال‌های ۱۲، ۱۳، ۶۰ و ۶۱، که مربوط به اعتماد به نفس می‌باشند، از بعد سلامت روان به این معنا نیست که اعتماد به نفس یک جنبه کم اهمیت از کیفیت زندگی است. بلکه باید گفت، سئوال‌های مربوط به اعتماد به نفس هماهنگ با سایر سئوال‌های موجود در بعد سلامت روانی عمل نمی‌کند. بنابراین اگر واقعا اعتماد به نفس جنبه ای از کیفیت زندگی است، یا باید جنبه‌های زبان شناختی این سئوال را مورد بازبینی قرار داد، به طوری که بعد روانی از کیفیت زندگی را بهتر منعکس کند و هماهنگ با سایر سئوال‌های موجود در این مدل عمل کند یا این که باید یک بعد مجزا برای اعتماد به نفس در پرسشنامه لحاظ کرد. مطالعات بعدی روی پرسشنامه WHOQOL-100 می‌توانند به بازبینی زبان شناختی سئوال‌های حذف شده و یا ساختن ابعاد جداگانه برای جنبه‌های مهمی که سئوال‌های حذف شده نماینده آنها بوده اند بپردازند.

از دیدگاه توسعه دهندگان آزمون‌ها، کنار هم قراردادن جنبه‌های مختلف از یک متغیر

پنهان، می‌تواند گستردگی یک سازه را افزایش دهد. برای مثال، هر یک از ۳۲ سؤال موجود در بعد سلامت محیط، برای نشان دادن یک جنبه قابل اهمیت از بعد سلامت محیط وارد پرسشنامه شده‌اند. خارج کردن هر یک از این سؤال‌ها می‌تواند روایی سازه‌ای که مد نظر سازنده پرسشنامه بوده است را مورد تهدید قرار دهد.

به هر صورت، این مسئله که آیا سؤال‌های موجود یک متغیر پنهان را مورد ارزیابی قرار می‌دهند یا نه، باید مورد آزمون قرار بگیرد. از طریق تحلیل رانش می‌شود آشفتگی‌های غیرمنتظره موجود در داده‌ها را پیدا کرد. فقط زمانی که داده‌ها با مدل برازش مناسبی داشته باشند، می‌شود ادعا کرد که ویژگی پنهان مورد نظر توسط سؤال‌ها سنجیده می‌شود. بنابر استدلال ارائه شده، باید سؤال‌ها و افرادی را که برازش مناسبی با مدل ندارند، تشخیص داد و آن‌ها را اصلاح یا حذف کرد. پس از حذف ۱۵ سؤالی که برازش مناسبی با مدل ندارند، از ۳۲ سؤال موجود برای سنجش بعد سلامت محیط می‌توان ادعا کرد که پرسش‌های باقی مانده می‌توانند به درستی بعد محیط را مورد ارزیابی قرار دهند. هرچند که این ارزیابی نسبت به ارزیابی‌ای که با ۳۲ سؤال انجام می‌شد چشم اندازهای کم‌تری را در بر می‌گیرد، اما نسبت به حالت قبل معنادار تر است.

بر روی نسخه فارسی این پرسشنامه تنها مطالعه یافت شده با استفاده از تحلیل رانش، مطالعه یزدانی و همکاران می‌باشد که روی نمونه افراد سالم انجام شده است. در مطالعه یزدانی مدل اعتبار جزئی بر روی تمام سؤال‌های پرسشنامه برازش داده شده است و پس از انجام آنالیزهای مربوط، تعداد ۳۰ سؤال باقی مانده است (یزدانی، ۲۰۱۰). مطالعه‌ای نیز با استفاده از روش تحلیل عاملی اکتشافی توسط کریملو و همکاران روی ۵۰۰ نمونه معلول موجود در این مطالعه انجام شده است. در این مطالعه، از تحلیل عاملی به روش مولفه اصلی و با استفاده از چرخش نامتعامل پروماکس استفاده شده است. دیده شده که ۴ سؤال کلی به همراه ۹۰ سؤال دیگر در پوشش ۱۰ عامل توانستند حدود ۵۶٪ از واریانس

مشاهده شده را بیان کنند (کریملو، نجات، صالحی و رهگذر، ۱۳۸۷).
با توجه به نتایج حاصل از این پژوهش توصیه می‌شود، جهت برآورد خصیصه کیفیت زندگی، از نسخه اصلاح شده پرسشنامه WHOQOL-100 استفاده شود. همچنین توصیه می‌شود جهت داشتن برآوردهای دقیق‌تر برای خصیصه پنهان، تحلیل راش به جای آنالیز کلاسیک، استفاده شود.

- کریملو، م.، نجات، س.، صالحی، م. و رهگذر، م. (۱۳۸۷). *استانداردسازی پرسشنامه کیفیت زندگی سازمان بهداشت جهانی-۱۰۰ و مطابقت آن با شرایط اجتماعی و فرهنگی ایران و کاربرد آن در مقایسه کیفیت زندگی معلولین جسمی-حرکتی و افراد سالم*. ایران.
- Adams, R. J., Wilson, M. and Wang, W. C. (1997). The Multidimensional Random Coefficients Multinomial Logit Model. *Applied Psychological Measurement*, 21(1), 1-23.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.
- Hedeker, D. (1999). MIXNO. Chicago: University of Illinois Press.
- Hedeker, D. & Gibbons, R. (1996). MIXOR: a computer program for mixed-effects ordinal regression analysis. *Comput Methods Programs Biomed*, 157-176.
- Holland, P. and Wainer, H. (1993). *Differential Item Functioning*.
- Karimlou, M., Zayeri, F., and Salehi, M. (2011). *Archives of Iranian Medicine*, 14(4), 281-287.
- Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.
- Skrondal, A., Rabe-Hesketh, S., and Boca Raton, F. (2004). *Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models*: Chapman & Hall/ CRC Press.
- Wang, W. C., Yao, G., Tsai, Y. J., Wang J.D., and Hsieh, C. L. (2006). Validating, improving reliability, and estimating correlation of the four subscales in the WHOQOL-BREF using multidimensional Rasch analysis. *Quality of Life Research*, 15, 607-620.
- Wilson, M. and Hoskens, M. (2005). *Multidimensional Item Responses: Multimethod-Multitrait Perspectives Applied Rasch Measurement: A Book Of Exemplars*. California: Springer.
- Wolfinger, R. (1999). *Fitting nonlinear mixed models with the new NLMIXED procedure*. Paper presented at the Proceedings of the 99 Joint Statistical Meetings.
- Right, B., Linacre, J., Gustafson, J. E. and Martin, L. P. (1994). Reasonable mean-square fit values. *Rasch Measurement, Transactions. Rasch Meas Trans*, 8, 370.
- Wright, B. and Stone, M. (1979). *Best Test Design*. Chicago: Measurement, Evaluation, Statistics, and Assessment Press.
- Wu, M., Adams, R., and Wilson, M. (1998). ConQuest (Version 2). Australia: *Australian Council for Educational Research*.
- Yazdani, K. (2010). *Constructing a "quality of life" questionnaire using the WHOQOL-100 questionnaire and the Rasch analysis*. Ph.D, Tehran.