

Sobre la estabilidad temporal de la Calidad de Vida: medición de la invarianza longitudinal del cuestionario WHOQOL - AGE

Estefanía Sánchez-Pastor Martín

Máster en Metodología de las Ciencias del Comportamiento y de la Salud



MÁSTERES
DE LA UAM
2017 - 2018

Facultad de Psicología



Facultad de Psicología

Departamento de Psicología Social y Metodología

**Sobre la estabilidad temporal de la Calidad de Vida: medición de
la invarianza longitudinal del cuestionario WHOQOL-AGE**

Trabajo Fin de Máster

Autora: Estefanía Sánchez-Pastor Martín

Máster en Metodología de las Ciencias del Comportamiento y de la Salud

Madrid, septiembre 2018

Sobre la estabilidad temporal de la Calidad de Vida: medición de la invarianza longitudinal del cuestionario WHOQOL-AGE

Autora: Estefanía Sánchez-Pastor Martín

Tutor académico: Francisco José Abad

Tutor profesional: Francisco Félix Caballero

Índice

Contenidos

Resumen.	2
Introducción.	3
WHOQOL: midiendo la calidad de vida.	4
Evaluación longitudinal de la calidad de vida y el papel de la invarianza de medida.	5
Marco del modelo del AFC longitudinal.	7
Objetivo.	9
Método.	9
Diseño.	9
Muestra y procedimiento.	9
Variables e instrumentos.	10
Análisis estadísticos.	11
Variables sociodemográficas y estadística descriptiva.	11
Análisis factorial.	12
Análisis factorial exploratorio.	12
Análisis factorial confirmatorio.	13
Índices de bondad de ajuste.	13
Análisis de la invarianza longitudinal (AFC longitudinal).	14
Software estadístico.	15
Resultados.	15
Estadística descriptiva.	15
Análisis factorial.	16
Análisis factorial exploratorio.	18
Análisis factorial confirmatorio.	19
Análisis de la invarianza longitudinal.	21
Discusión y conclusiones.	22
Financiación.	26
Referencias bibliográficas.	27
Anexos.	32

Resumen

La naturaleza múltiple y compleja del concepto de calidad de vida ha reforzado el uso de medidas que incorporan distintas facetas en su evaluación única, así como de aquellas que tienen en cuenta su carácter cambiante. El presente trabajo tuvo por objetivo principal analizar la invarianza de medida del cuestionario WHOQOL-AGE a lo largo de dos momentos temporales en una muestra de adultos españoles. En primer lugar, se realizó un estudio completo de la dimensionalidad de la escala, explorando distintas estructuras mediante un análisis factorial exploratorio. Una vez seleccionada la mejor estructura en base a criterios estadísticos y sustantivos, se testó un modelo factorial confirmatorio sobre los datos de ambas olas de medición (basal y seguimiento). En tercer lugar, se llevaron a cabo pruebas de invarianza factorial en el contexto del análisis factorial confirmatorio longitudinal. Los resultados sugieren que el WHOQOL-AGE es un instrumento invariante a través los dos momentos temporales, algo que permite concluir que las diferencias observadas en las puntuaciones a lo largo del tiempo se deben únicamente a diferencias en las variables latentes subyacentes.

Palabras clave: calidad de vida, WHOQOL-AGE, análisis factorial confirmatorio, invarianza de medida, invarianza longitudinal.

Introducción

La noción general de calidad de vida (CV) fue considerada inicialmente como un complemento útil a los conceptos tradicionales de salud y estado funcional. Afortunadamente y desde hace tiempo, la investigación sobre CV ha adquirido una relevancia genuina debido a que este concepto se ha emancipado como principio organizador de la sociedad, siendo posible aplicarlo en la mejora de la misma (Schalock, 1996).

Para algunos autores, esta variable se origina con el objetivo de distinguir resultados relevantes para la investigación en salud, derivada de los estudios tempranos en bienestar subjetivo y satisfacción con la vida (Smith, Avis y Assman, 1999). El hecho de que en sus comienzos haya estado vinculada a otros constructos psicológicos, ha provocado que, por su rápido repunte de las últimas décadas, se hayan evidenciado problemas tanto en su definición, en las dimensiones que la comprenden, en su medición, así como en los factores que la pueden influenciar (Taillefer, Dupuis, Roberge y Le May, 2003).

Al no existir una definición única sobre el concepto de CV, se ha propuesto una categorización de estas definiciones en función de otras variables tenidas en cuenta (Urzúa y Caqueo, 2013), a saber: *condiciones de vida objetivas* (salud física, condiciones de vida, relaciones sociales, actividades funcionales u ocupación); *satisfacción con la vida* (en distintos dominios); *condiciones de vida + satisfacción con la vida* (incluiría aquellas definiciones que tienen en cuenta indicadores biológicos, sociales, materiales, conductuales y psicológicos, sumado a las emociones evocadas por cada área); [*condiciones de vida + satisfacción con la vida*] * *evaluación* (el nivel de CV percibido estaría modulado por los procesos cognitivos vinculados a la evaluación que la persona realiza de los otros dos aspectos mencionados); [*condiciones de vida + satisfacción con la vida*] * *valores* (incluiría aquellas definiciones que tienen en cuenta todo lo anterior, mediado por los valores personales). Dentro de esta última categoría de definiciones, podría incluirse la que fue propuesta en la Organización Mundial de la Salud (OMS) por el WHOQOL (World Health Organization Quality Of Life), un grupo de trabajo dedicado expresamente al estudio del concepto de calidad de vida: “percepción que un individuo tiene de su lugar en la existencia, en el contexto de la cultura y el sistema de valores en los que vive, en relación con sus expectativas, sus normas y sus inquietudes” (The WHOQOL Group, 1995).

Smith, Avis y Assman (1999) plantearon que la percepción de la calidad de vida estaría basada en un proceso que involucraría: (1) identificar los dominios relevantes que la

conforman, (2) determinar los estándares de comparación en cada uno y (3) integrar los juicios separados de cada uno de los dominios en una evaluación global de calidad de vida. Según esta perspectiva, se trataría de un concepto multidimensional cuya evaluación única implicaría incorporar distintas facetas (Urzúa y Caqueo, 2013).

WHOQOL: midiendo la calidad de vida

La evaluación de la CV recoge la problemática que descansa en su aproximación conceptual: su naturaleza múltiple y compleja, y la dualidad entre lo objetivo y lo subjetivo implican utilizar, en función de dónde se ponga el foco de atención, diversos criterios a la hora de medirla (Moreno y Ximénez, 1996).

El panel de investigadores que formaban el Grupo WHOQOL, buscando evaluar dicho constructo, creó una serie de instrumentos que permitían su aplicación en distintas poblaciones. El primero de ellos es el WHOQOL-100, compuesto por 100 ítems, desarrollado durante varios años en diversos entornos culturales (The WHOQOL Group, 1994). En su versión breve de 26 ítems, el WHOQOL-BREF (The WHOQOL Group, 1998) mostró buenas propiedades psicométricas y mantuvo cuatro de las seis dimensiones originales: salud física, salud psicológica, relaciones sociales y condiciones ambientales.

Se sabe que, según el rango de edad, existen facetas de la CV más relevantes que otras. La escala resultante de tener en consideración a la población anciana fue el WHOQOL-OLD (Power, Quinn y Schmidt, 2005), para la cual existen también distintas versiones breves. Cuenta con 26 ítems y buenas propiedades psicométricas y en ella se consideran 6 dimensiones: habilidades sensoriales; autonomía; actividades pasadas, presentes y futuras; participación social; muerte; e intimidad.

Para salvar los problemas relacionados con la duración y complejidad tanto del WHOQOL-100 como del WHOQOL-BREF se creó el EUROHIS-QOL, con 8 ítems (Schmidt, Mühlán y Power, 2006). Esta medida de la calidad de vida es una adaptación del segundo de ellos, manteniendo sus cuatro dimensiones representadas por dos ítems cada una de ellas.

Finalmente, el WHOQOL-AGE es un instrumento creado a partir de otros de la familia del WHOQOL, desarrollado en el marco del proyecto COURAGE in Europe –*Collaborative Research on Ageing in Europe*– y validado en muestras representativas a nivel nacional en Finlandia, Polonia y España (Caballero et al., 2013). Este instrumento se compone de 8 ítems

del EUROHIS-QOL y 5 de una versión abreviada del WHOQOL-OLD. En algún caso, la redacción de las preguntas fue modificada, dando lugar, así, a una prueba de 13 ítems positivos medidos en una escala Likert de cinco puntos.

La principal ventaja de esta escala es que reduce el tiempo necesario para completar otras similares que evalúan la CV, concretamente en población de avanzada edad, pudiendo ser aplicada en proyectos a gran escala que implican poblaciones de mayor tamaño. En el trabajo de Caballero et al. (2013) se evidenciaron buenas propiedades psicométricas y se encontró que la estructura factorial que mejor se ajustaba era la de un factor de segundo orden con dos factores de primer orden.

Dentro del proceso de validación de un instrumento, no solo es necesario comprobar la fiabilidad de las puntuaciones u otras propiedades psicométricas, sino también probar que éstas se mantienen invariantes en distintos grupos. De esto se ha ocupado un reciente estudio (Santos et al., 2018) el cual tuvo por objetivo poner a prueba la invarianza de las puntuaciones del WHOQOL-AGE en diferentes países, utilizando un análisis factorial confirmatorio multigrupo sobre las tres muestras disponibles. Su conclusión a este respecto fue que los encuestados de los países participantes –entre ellos, España– atribuyeron el mismo significado al constructo estudiado, pudiendo imputarse los cambios encontrados a diferencias reales en la calidad de vida.

También en este estudio, y en un paso previo, se puso a prueba una estructura bifactor con un factor general y dos factores específicos, mostrando ésta un mayor ajuste que la de dos factores de primer orden correlacionados. Las dos dimensiones específicas resultantes fueron dotadas de significado e interpretadas como factores de formato de respuesta (bipolar y unipolar).

Evaluación longitudinal de la calidad de vida y el papel de la invarianza de medida

Como ya se ha comentado más arriba, la percepción de la CV es un reflejo de cómo el individuo canaliza sus juicios sobre situaciones objetivas de cada uno de los ámbitos evaluados, de acuerdo a modelos de comparación basados en emociones, aspiraciones o expectativas (Urzúa y Caqueo, 2013). Bajo este prisma, la CV sería el resultado de la discrepancia entre un estado ideal o deseado y uno real (Felce y Perry, 1995). Además de esto, las comparaciones entre uno mismo y los demás, o entre uno mismo en distintos momentos en el tiempo se posicionan como puntos clave a tener en cuenta a la hora de evaluar la CV.

Lo más frecuente es encontrar mediciones transversales, en las que se atiende a las discrepancias entre sujetos; por su parte, los estudios longitudinales –en los que se observan las diferencias intrasujetos–, debido a su gran coste y al carácter cambiante y dependiente del concepto de CV, son menos comunes.

Este mismo carácter dinámico es el que advierte de la necesidad de comprobar que lo que se está midiendo en cada momento es un mismo constructo; y es que es importante asegurarse de que las personas implicadas perciben el significado subyacente de los ítems de la misma forma a lo largo del tiempo, así como tener en cuenta la ocurrencia de algún cambio ambiental (por ejemplo, una crisis económica) durante un periodo concreto que pudiera tener un impacto sobre la percepción del constructo. En el caso del proyecto que nos ocupa, en el que hay un seguimiento de población adulta mayor, es importante conocer el impacto que tiene el envejecimiento en múltiples variables con el objetivo de tomar las medidas oportunas.

De este modo, el diseño longitudinal es condición necesaria para evaluar el cambio y los factores implicados en él, siendo éste uno de los principales objetivos de las ciencias del comportamiento y de la salud (Widaman, Ferrer y Conger, 2010). El cambio en la puntuación verdadera de la variable latente que se pretende encontrar cuando utilizamos estos diseños es lo que Chan (1998) llamó cambio *alfa*, frente a los cambios *beta* y *gamma*, consistentes en discrepancias en las propiedades de medida de los indicadores y modificaciones en el significado del constructo según el momento temporal, respectivamente.

Estudiar el cambio o la relación entre procesos a través del tiempo implica asumir la estabilidad temporal del constructo; sin embargo, en lugar de hacer esta suposición, podemos probar formalmente dicha hipótesis a través de la invarianza factorial. De hecho, en ausencia de tal comprobación, no es posible determinar si el cambio observado se debe a un cambio real en el constructo o en la estructura y medición del mismo. Por lo tanto, el examen de la invarianza factorial debe preceder a la aplicación de procedimientos SEM (*Structural Equation Modeling*) con datos longitudinales (p.ej., modelos de curva de crecimiento latente; Bollen y Curran, 2004; citado en Brown, 2006). Lo que se pretende demostrar con esto es que las relaciones entre las variables latentes y sus indicadores son invariantes independientemente del momento de medición. Dicho de otra forma, el valor esperado de la puntuación de una persona en la variable observada j en el momento t debe ser función de su puntuación en la variable latente y del factor específico asociado a esa variable en el momento t , no debiendo depender adicionalmente del momento de medición (Meredith, 1993).

Marco del modelo de análisis factorial confirmatorio longitudinal

Una definición de invarianza de medida puede ser el proceso de verificación de que las propiedades de medida de los instrumentos o sus ítems, son independientes de las características de los grupos evaluados, pero no del constructo que el instrumento está midiendo (Meredith, 1993). Ésta ha sido ampliamente estudiada en el contexto del AFC multigrupo, comparando grupos definidos por variables nominales como el sexo o el entorno cultural. Sin embargo, la flexibilidad del marco del análisis factorial confirmatorio (AFC) ha inspirado multitud de modelos empleados en los diseños longitudinales de medición, entre ellos, la invarianza factorial (Koch, Schultze, Eid y Geiser, 2014).

Las relaciones de medición entre los k ítems en el g^o momento temporal ($g = 1, 2$) bajo el modelo factorial confirmatorio de ajuste a los datos podría representarse como (Jöreskog y Sörbom, 1996):

$$X_k^g = \tau_k^g + \Lambda_k^g \xi^g + \delta_k^g \quad (1)$$

donde X_k^g se refiere al vector de ítems, Λ_k^g es la matriz de pesos que relaciona X_k^g con el vector de constructos ξ^g , τ_k^g indica el vector de los interceptos de la regresión, y δ_k^g representa el vector de los errores de medición de los ítems. La ecuación de covarianzas de (1) puede escribirse como:

$$\Sigma^g = \Lambda_X^g \Phi^g \Lambda_X^{g'} + \Theta_\delta^g \quad (2)$$

donde, Σ^g es la matriz de varianzas-covarianzas entre los K ítems en el g^o momento temporal, Λ_X^g es la matriz de pesos factoriales de los ítems en ξ^g , Φ^g recoge las varianzas y covarianzas entre los constructos ξ^g , y Θ_δ^g es la matriz de varianzas únicas.

Al operar en un entorno de AFC, es posible separar el error de medición de la puntuación verdadera, y por otro lado, generar una serie de restricciones de manera consecutiva que dan lugar a lo que se identifica como niveles de invarianza de medida. Concretamente, el enfoque de Widaman y Reise (1997) adapta fácilmente al contexto longitudinal la evaluación de la invarianza de medición en distintos grupos:

a. *Invarianza configural*: en este modelo se pone a prueba, a través de un AFC, que la estructura de covarianzas entre los ítems es similar en los distintos momentos temporales. Se permite que los factores estén correlacionados y que, además, el error de cada indicador

(unicidad), correlacione con el error del mismo indicador en otros puntos en el tiempo. Dicha especificación asume que el error de cada indicador está compuesto de una combinación de error aleatorio y varianza específica del indicador, la cual se asume estable en el tiempo. Esta última idea es la que diferencia a la invarianza longitudinal de la invarianza factorial entre grupos (Brown, 2006).

b. *Invarianza métrica*: las cargas factoriales de cada indicador se especifican como invariantes a lo largo del tiempo ($\Lambda_k^g = \Lambda_k^{g'}$). Esta igualdad de pesos entre momentos temporales es necesaria para garantizar relaciones equivalentes entre los factores y sus indicadores en el modelo, permitiendo, así, su comparación (Dimitrov, 2010).

c. *Invarianza escalar*: se asume la invarianza métrica y se aplica una restricción adicional consistente en que los interceptos de cada indicador se especifican como iguales en cada momento temporal ($\tau_k^g = \tau_k^{g'}$).

d. *Invarianza estricta*: en este paso, además de los pesos factoriales y los interceptos, las varianzas error (unicidades) se especifican para ser iguales en el tiempo en cada uno de los indicadores ($\theta_\delta^g = \theta_\delta^{g'}$). Este nivel proporciona evidencia de que los ítems se midieron con la misma precisión en cada punto temporal. Concluyendo que existe este tipo de invarianza, las diferencias encontradas en las puntuaciones se deberían únicamente a las diferencias en las variables latentes (Dimitrov, 2010).

Por su parte, y al contrario que la invarianza de medida –la cual se centra en las variables observadas– los modelos que conciernen solo a las variables latentes son los puestos a prueba con la invarianza estructural (Milfont y Fisher, 2010; Dimitrov, 2010). Hay tres niveles dentro de esta categoría: invarianza de medias, invarianza de las varianzas e invarianza de las covarianzas del factor a lo largo del tiempo. Estos modelos no están anidados entre sí, por lo que a diferencia de los modelos de medida, no son jerárquicos, sino secuenciales.

Los aspectos de la invarianza factorial que se evalúan dependen de los fines específicos de cada trabajo. Así, la decisión con respecto a si debería probarse la invarianza estructural dependerá de si la variabilidad de los constructos y/o las relaciones entre ellos se consideran relevantes. En el contexto de este estudio, atender al modelo estructural permitiría extraer conclusiones acerca del significado que adquiere la CV para la población en su proceso de

envejecimiento, así como comprobar si las relaciones entre factores se mantienen invariantes a lo largo del tiempo; sin embargo, esto es algo que excede los propósitos de este trabajo.

Por último, es importante considerar la identificación del modelo, puesto que cuando ésta no se da, no es posible estimar los parámetros. Las restricciones concretas realizadas en cada paso de invarianza se exponen en el apartado de método.

Objetivo

El objetivo general del presente trabajo fue medir la invarianza longitudinal del WHOQOL-AGE, con el fin de dar soporte metodológico a la comparación y relación entre las puntuaciones de CV de una muestra española en dos momentos temporales.

Como objetivos específicos, se realizó un estudio completo de la dimensionalidad del cuestionario en busca de la estructura más adecuada, a través de un análisis factorial exploratorio (AFE), probando el modelo seleccionado mediante un AFC. Por otra parte, se analizó el sesgo de atrición para arrojar luz acerca de la pérdida de datos en el seguimiento.

Método

Diseño

Estudio longitudinal observacional sobre el estado de salud, calidad de vida y bienestar de una muestra probabilística representativa de población adulta en España.

Muestra y procedimiento

La muestra empleada para este trabajo –adultos no institucionalizados, con sobrerrepresentación de la población mayor de 50 años– forma parte de un estudio de seguimiento que se inició en España hace 7 años dentro del proyecto COURAGE in Europe, en el que se reclutaron varias muestras representativas de distintos países entre 2011 y 2012 (Leonardi et al., 2014; Miret et al., 2014). Para su obtención, se realizó un muestreo aleatorio polietápico estratificado por ámbito geográfico y edad. El periodo de la segunda medición (primer momento del seguimiento) comprendió desde diciembre de 2014 hasta junio de 2015. Actualmente, está en marcha la recolección de una tercera ola de datos y el WHOQOL-AGE está de nuevo incluido en ella, con el fin de evaluar la calidad de vida de los participantes y los posibles cambios en sus trayectorias a lo largo del tiempo.

La muestra inicial estuvo compuesta por todos aquellos españoles que respondieron al cuestionario en el primer momento de medición, haciendo un total de 4.271 participantes. En la segunda ola de evaluaciones, tan solo 2.397 completaron el WHOQOL-AGE. Finalmente, y tras eliminar aquellos casos para los que no se disponía de información completa relativa a las variables objeto de estudio, la muestra seleccionada estuvo formada por las 2.031 personas que contaban con la medida basal (T1) y la de seguimiento (T2).

Los datos de dicha muestra fueron recogidos mediante una entrevista individual de 60 minutos de duración aproximadamente, cara a cara y en el domicilio de los participantes, a través de un programa informático (CAPI, *Computer-Assisted Personal Interviewing*). Los entrevistadores participaron en un curso de entrenamiento para la administración del cuestionario completo. Antes de recopilar los datos, se obtuvo el consentimiento informado. Esta investigación fue aprobada por el comité ético del Hospital Universitario de La Princesa (Madrid) y del Parc Sanitari Sant Joan de Déu (Barcelona).

Variables e instrumentos

El cuestionario WHOQOL-AGE está compuesto por 13 ítems positivos, 8 de ellos derivados del EUROHIS-QOL (Schmidt, Quinn y Power, 2005) y 5 de la versión corta del WHOQOL-OLD (Fang et al., 2012), evaluados en una escala Likert de 5 puntos (ver Tabla 1).

El formato de respuesta combina elementos bipolares y unipolares. Así, el ítem Q1 (desde 1 = “muy mala”, hasta 5 = “muy buena”) y los ítems Q2 al Q8 (desde 1 = “muy insatisfecho”, hasta 5 = “muy satisfecho”) cuentan con un formato bipolar, frente a los ítems unipolares del Q9 al Q13 (1 = “ninguna”, hasta 5 = “completamente”). Investigaciones previas han sugerido una estructura de segundo orden con dos factores de primer orden (Caballero et al., 2013) y un modelo bifactor con un factor general y dos factores específicos (Santos et al., 2018).

Análisis estadísticos

Variables sociodemográficas y estadística descriptiva

Se llevaron a cabo análisis de frecuencias para describir las características de la muestra empleada. Las variables sociodemográficas estudiadas fueron la edad, el sexo, los años de educación, el entorno residencial y el estado civil. Se calcularon algunos estadísticos descriptivos para los ítems de la escala, incluidos los índices de asimetría y curtosis. Lo

concluido con respecto a la normalidad de los datos, se utilizaría posteriormente para elegir el método de estimación de parámetros más adecuado.

Tabla 1

Preguntas del cuestionario WHOQOL-AGE

-
- Q1. ¿Cómo calificaría usted su calidad de vida?*
- Q2. ¿Cómo de satisfecho está con su audición, visión y demás sentidos?*
- Q3. ¿Cómo de satisfecho está con su salud?**
- Q4. ¿Cómo de satisfecho está consigo mismo?*
- Q5. ¿Cómo de satisfecho está con su capacidad para realizar sus actividades de la vida diaria (cotidianas)?*
- Q6. ¿Cómo de satisfecho está con sus relaciones personales?*
- Q7. ¿Cómo de satisfecho está con las condiciones del lugar donde vive (su casa)?*
- Q8. ¿Cómo de satisfecho está con el modo en que utiliza su tiempo?**
- Q9. ¿Tiene usted suficiente energía para su vida diaria?*
- Q10. ¿Cuánto control tiene sobre las cosas que le gusta hacer?**
- Q11. ¿Hasta qué punto está usted satisfecho con sus oportunidades para continuar alcanzando sus objetivos (logrando sus metas) en la vida?**
- Q12. ¿Tiene usted suficiente dinero para cubrir sus necesidades?*
- Q13. ¿Cómo de satisfecho está con sus relaciones íntimas en su vida?**
-

Nota: Los ítems en **negrita** fueron extraídos de la versión corta del WHOQOL-OLD. El resto de ítems, del EUROHIS-QOL. En *cursiva*, los ítems con un formato de respuesta bipolar. El resto, compartían formato unipolar.

Se realizó un análisis del sesgo por atrición estudiando las diferencias en las variables sociodemográficas ya nombradas entre aquellas personas que respondieron al WHOQOL-AGE en ambas olas y las que sólo lo hicieron en la primera. Solo para este fin, se empleó la muestra inicial completa en España ($N = 4271$), dividida en esos grupos. El sesgo de atrición presenta dos problemas: por una parte, un menor número de observaciones disminuye la eficiencia de las estimaciones; por otra, si las características de los que no vuelven a ser encuestados difieren sistemáticamente de las características de los que sí vuelven a serlo, entonces la atrición es selectiva y las estimaciones pueden resultar sesgadas (Bendezú, Denis y Zubizareta, 2007). Se empleó con este fin la prueba χ^2 de Pearson cuando las variables eran categóricas. Para determinar el tamaño del efecto se utilizó la V de Cramer. Para interpretar la magnitud del efecto encontrado se utilizaron los valores propuestos por Cohen (1988): 0,10;

0,30 y 0,50 que se interpretan como pequeño, moderado y grande respectivamente. En el caso de la asociación con variables cuantitativas, se utilizó la prueba t de Student para muestras independientes. Para evaluar el tamaño del efecto se calculó la d de Cohen. Los valores 0,20; 0,50 y 0,80 se interpretan como efectos pequeños, moderados y grandes, respectivamente (Cohen, 1992).

Análisis factorial

En primer lugar, se realizó un AFC sobre los datos de la muestra ya seleccionada en el primer momento temporal ($N = 2031$) para determinar el ajuste del modelo bifactor de dos factores específicos propuesto recientemente en Santos et al. (2018).

En segundo lugar, y tras observar que el modelo encontrado en este estudio no ofrecía una solución admisible, se procedió a explorar otras estructuras subyacentes a partir de un análisis factorial exploratorio (AFE). En un paso previo, se dividió la muestra aleatoriamente en dos, utilizándose para este primer propósito el 70% de los datos de la primera ola. Sobre el 30% restante ($N = 609$), se realizó AFC. Por último, se puso a prueba en la segunda ola completa ($N = 2031$), también con un AFC, la estructura ya seleccionada. Este modo de proceder responde a la recomendación de realizar un uso secuencial de ambos análisis, siempre que el tamaño muestral lo permita (MacCallum et al., 1992; Anderson y Gerbing, 1998).

Análisis factorial exploratorio

En este punto, cabe comentar que en trabajos anteriores sobre el estudio de la dimensionalidad del WHOQOL-AGE se optó por el criterio MAP –*Minimum Average Partial-correlation*– de Velicer (Velicer, 1976) como único método de extracción de factores. En el presente trabajo, y siguiendo las recomendaciones de investigaciones relacionadas (Garrido, Abad y Ponsoda, 2011; Zwick y Velicer, 1986), se decidió no utilizar este criterio ya que su precisión depende del tamaño de los pesos y del número de ítems por factor, siendo mayor cuando las saturaciones son muy altas y el promedio de indicadores por factor es de 8 o más; en caso contrario, el número de dimensiones seleccionadas queda subestimado. Más aún, la implementación de un solo método para decidir el número de factores a extraer resulta insuficiente, motivo por el cual se aconseja el uso de varios criterios objetivos, así como considerar la interpretabilidad de la solución obtenida y el número de factores según la teoría de partida (Lloret-Segura, Ferreres-Traver, Hernández-Baeza y Tomás-Marco, 2014). De este

modo, se llevó a cabo un análisis paralelo, AP (Horn, 1965), ya que de entre diversos métodos ha demostrado ser el más preciso y robusto en distintas condiciones (Wells y Faulkner-Bond, 2016).

Posteriormente, se realizó un AFE explorando las soluciones que oscilaban entre $n-1$ y $n+1$ factores (siendo n el número de factores obtenidos mediante el AP), a través de distintos índices de bondad ajuste (ver apartado con el mismo nombre). El método de estimación de parámetros empleado fue máxima verosimilitud (ML) usando la matriz de correlaciones de Pearson. Un punto importante en la aplicación de ML es comprobar que la distribución de las variables es multivariada normal. Para ello, se ha propuesto estudiar los índices univariados de asimetría y curtosis (Curran, West y Finch, 1996; Abad, Olea, Ponsoda y García, 2011), demostrándose que, para valores por debajo de 2 y 7 respectivamente, los métodos de estimación son sólidos –como ya se ha comentado más arriba, esto se analizó–.

Se utilizó un método de rotación oblicua (oblimin) puesto que, bajo el marco teórico previo, se podía asumir que los factores extraídos estarían correlacionados. Pesos factoriales superiores a 0,30 se consideraron indicativos de la asociación de un ítem con un determinado factor (McDonald, 1985; Bandalos y Finney, 2010).

Análisis factorial confirmatorio

Para determinar el ajuste del modelo seleccionado en el AFE en base a los criterios estadísticos ya mencionados, se realizó un AFC (considerando algunos pesos cruzados y correlación entre factores) en el 30% restante de la primera ola y en el 100% de la segunda ola. El método de estimación empleado en este caso fue máxima MLM (*Maximum Likelihood Mean-adjusted*; Finney y DiStefano, 2006; Santos et al., 2018).

Índices de bondad de ajuste

Entre los índices más comúnmente utilizados para valorar el ajuste de un modelo se encuentra χ^2 (Cochran, 1952); sin embargo, muchos autores desaconsejan su uso debido a su dependencia del tamaño muestral (Hooper, Coughlan y Mullen, 2008). Así, si la muestra es demasiado grande, las decisiones basadas en este índice podrían llevar a rechazar modelos razonables. Debido al tamaño muestral con el que se trabajó en este estudio –y aunque se reporta el valor de χ^2 siguiendo las recomendaciones de Steiger (2007)–, se utilizaron otros índices, a saber: CFI (*índice de ajuste comparativo*; Bentler, 1990), TLI (*índice de Tucker-Lewis*; Tucker y Lewis, 1973) y RMSEA (*raíz del error cuadrático medio de aproximación*;

Steiger, 1990). En el caso de los dos primeros, valores por encima de 0,90 se consideran adecuados, y buenos si son superiores a 0,95 (Hu y Bentler, 1999); por su parte, valores de RMSEA por debajo de 0,08 denotan un ajuste razonable, y por debajo de 0,05, un buen ajuste (Van de Schoot, Lugtig y Hox, 2012).

Análisis de la invarianza longitudinal (AFC longitudinal)

En función del modelo seleccionado a partir del AFE y AFC previos, se llevó a cabo un análisis factorial confirmatorio longitudinal con tres factores correlacionados utilizando MLM (Santos et al., 2018). La versatilidad del AFC hace posible su uso para la comparación de puntuaciones de una misma muestra en distintos puntos en el tiempo.

La invarianza longitudinal de medida se analizó mediante la imposición secuencial de restricciones en las dos olas disponibles: estructura factorial similar en ambos momentos temporales (invarianza configural), invarianza de las cargas factoriales (métrica), invarianza de las cargas factoriales y los interceptos (escalar) e invarianza de las cargas factoriales, interceptos, varianzas y covarianzas de los errores (estricta).

Con el fin de que los modelos estuviesen identificados, se impusieron determinadas restricciones en los sucesivos niveles de invarianza testados. En el modelo 1 (invarianza configural) se fijaron las medias de las todas las variables latentes a cero y sus varianzas a cero. En el modelo 2 (invarianza métrica), se fijaron todos los pesos factoriales como iguales, la media de los factores a cero en ambas olas y las varianzas de las variables latentes a uno en el momento T1, estimándose libremente en el momento T2. Para el modelo 3 (invarianza escalar), se fijaron, en T1, las medias y varianzas de los factores a cero y uno respectivamente, y los interceptos como iguales en ambos tiempos. En el último paso de la invarianza de medida (modelo 4, estricta), además de las restricciones del modelo anterior, se impusieron varianzas error iguales en ambas olas. Para todos estos modelos, se asumieron correlaciones entre los factores y los errores de los distintos momentos temporales.

Los índices de bondad de ajuste se utilizan para determinar el grado en que un modelo teórico es consistente con los datos empíricos, siendo también útil en el caso de la invarianza factorial. Las limitaciones de esta forma de estudiar la invarianza están asociadas a las limitaciones de la metodología de los modelos de ecuaciones estructurales. Lo deseable es contar con datos continuos y un tamaño muestral determinado (Martínez, 2009). En el caso del WHOQOL-AGE, la escala cuenta con 5 opciones de respuesta, algo que, bajo una

distribución aproximadamente normal de los datos, permite tratar las variables como continuas (Finney y DiStefano, 2006). Generalmente, se ha recomendado el test de diferencias de chi-cuadrado para analizar la invarianza. Sin embargo, su sensibilidad a tamaños muestrales grandes (como el caso que nos ocupa) y su suposición subyacente de que el modelo se ajusta perfectamente a los datos, han sido reconocidas desde hace tiempo como problemáticas (Milfont y Fischer, 2010). Es por ello que en estos casos se recomienda utilizar un cambio en CFI mayor que 0,01; sumado a un cambio en RMSEA de 0,015 para poder afirmar que existe un decremento en el ajuste entre modelos anidados, o bien, rechazar el modelo más restrictivo (Chen, 2007; Cheung y Rensvold, 2002).

Software estadístico

El manejo de la base de datos y los análisis descriptivos se llevaron a cabo con el programa Stata 12 (StataCorp, 2011). Tanto para el análisis factorial exploratorio y confirmatorio, como para el estudio de la invarianza longitudinal, se utilizó el programa Mplus 7 (Muthén y Muthén, 2011).

Resultados

Estadística descriptiva

La muestra final estuvo formada por 2.031 sujetos. La edad media en el momento en que se llevó a cabo la primera ola de mediciones fue de 57,70 años (D.T. = 0,32), contando con un 51,8% de mujeres.

El motivo para estudiar la atrición fue verificar la aleatoriedad de la no respuesta en el seguimiento. En la Tabla 2 se exponen las características sociodemográficas de la muestra que solo respondió al WHOQOL-AGE en la primera ola y de la que respondió en ambos momentos temporales; se informa de aquellos casos en los que las diferencias en dichas variables fueron estadísticamente significativas para ambos grupos de participantes y una medida del tamaño del efecto asociado.

Se encontraron diferencias significativas en las variables edad, estado civil y años de educación; sin embargo, el tamaño del efecto encontrado fue pequeño en todos los casos.

Tabla 2.

Características sociodemográficas según participación en el seguimiento (variables correspondientes al primer momento temporal).

	Solo línea base (primera ola)	Completan seguimiento (ambas olas)	T.E.
N° participantes (%)	1.874 (47,99)	2.031 (52,01)	
Edad media (D.T.)	61,03 (0,40)	57,70 (0,32)	0,207*
Sexo (%)			-0,040
Mujer	1.047 (55,9)	1.052 (51,8)	
Hombre	827 (44,1)	979 (48,2)	
Estado civil (%)			0,132*
Casado/vive en pareja	1.061 (56,6)	1.340 (66)	
No casado	813 (43,4)	690 (34)	
Entorno residencial (%)			0,026
Urbano	1.634 (87,1)	1734 (85,4)	
Rural	240 (12,9)	296 (14,6)	
Años educación media (D.T)	10,67 (0,15)	11,60 (0,13)	-0,148*

Nota: T.E.: tamaño del efecto. D.T.: desviación típica. *: Diferencias significativas nivel de confianza 95%.

Los estadísticos descriptivos (separados por olas) para todos los ítems incluidos en el modelo se presentan en la Tabla 3. Siguiendo el criterio descrito en el método, podemos concluir que no hay violación severa de la normalidad puesto que todos los valores absolutos de asimetría y curtosis encontrados son menores que 2 y 7 respectivamente.

Análisis factorial

En primer lugar, se llevó a cabo un AFC probando la estructura bifactor de dos factores. Los resultados mostraron valores buenos de los índices de bondad de ajuste de este modelo (CFI = 0,950; TLI = 0,926; RMSEA = 0,063). Sin embargo, al fijarnos en los pesos estandarizados, se observó que la solución para el primer factor específico no era admisible, encontrándose que las saturaciones de los ítems Q1, Q2, Q3 y Q5 eran negativas, sugiriendo la existencia de un tercer factor específico que incluyera esos elementos (ver ítems señalados de la Tabla 4). Este patrón coincide con el que se encontró en Santos et al (2018), en este caso, con pesos prácticamente iguales a cero.

Tabla 3

Estadísticos descriptivos de los ítems del WHOQOL-AGE en las dos olas

Ítems	Ola-I en 2012				Ola-II en 2015			
	Media	D.T.	Asim.	Curt.	Media	D.T.	Asim.	Curt.
Q1	3,73	,750	-,717	1,143	3,74	,729	-,701	,992
Q2	3,98	,786	-1,010	1,741	3,99	,722	-,984	2,173
Q3	3,84	,878	-,857	,702	3,81	,882	-,967	,962
Q4	4,10	,729	1,051	2,343	4,05	,713	-1,051	2,495
Q5	3,97	,817	-1,044	1,535	3,90	,874	-1,205	1,677
Q6	4,17	,653	-,865	2,404	4,13	,668	-1,020	3,021
Q7	4,24	,680	-1,116	3,129	4,16	,669	-1,056	3,133
Q8	3,95	,803	-,990	1,498	3,95	,781	-1,136	2,074
Q9	4,07	,901	-,880	,487	3,96	,926	-,922	,697
Q10	4,00	,902	-,788	,313	3,97	,898	-,939	,920
Q11	3,80	,948	-,733	,356	3,74	,952	-,766	,385
Q12	3,67	1,096	-,521	-,503	3,66	1,043	-,517	-,320
Q13	4,01	1,004	-1,137	1,115	3,81	1,050	-,987	,700

Nota: D.T.: desviación típica. Asim.: asimetría. Curt.: curtosis

Con permiso de los autores, se reproducen los pesos estandarizados que se hallaron en el modelo de invarianza configural, tanto para el factor general como para los dos factores específicos, comparados con los encontrados en este estudio (Tabla 4).

Tabla 4

Comparación de pesos estandarizados del factor general y de los específicos.

Ítems	F. general	F. general	F. específico	F. específico
	(Santos et al., 2018)	(este estudio)	(Santos et al., 2018)	(este estudio)
Q1	0,580	0,383	0,044	-0,033
Q2	0,661	0,446	0,000	-0,167
Q3	0,819	0,696	0,000	-0,388
Q4	0,722	0,597	0,449	0,154
Q5	0,831	0,646	0,041	-0,246
Q6	0,624	0,501	0,577	0,182
Q7	0,427	0,332	0,400	0,167
Q8	0,629	0,556	0,369	0,107
Q9	0,671	0,543	0,331	0,507
Q10	0,638	0,524	0,342	0,585
Q11	0,591	0,522	0,519	0,490
Q12	0,314	0,321	0,540	0,396
Q13	0,466	0,467	0,403	0,329

Nota: en negrita, los pesos de los ítems correspondientes al primer factor específico. La muestra empleada en este estudio es una submuestra de la utilizada en la investigación de Santos et al (2018), donde se analizaron los datos únicamente de la primera ola, comparándolos con los de Polonia y Finlandia.

En la discusión se retomarán estos resultados. Puesto que la solución obtenida al testar el modelo bifactor de de dos factores específicos no era admisible ni interpretable, se procedió a explorar nuevas soluciones factoriales.

Análisis factorial exploratorio

El análisis paralelo indicó que el número de factores a extraer era 2. Al explorar esta solución, se observó que los índices de bondad de ajuste no eran adecuados (ver Tabla 5). En este modelo, los pesos de la matriz rotada no permitían interpretar dichos factores como indicadores de dos formatos de respuesta (varios ítems pesaban en ambos factores con saturaciones por encima de 0,20 y por debajo de 0,50). Por esta razón, se procedió a explorar las estructuras de uno y tres factores (ver Tabla 5). De acuerdo con los valores de los índices de bondad de ajuste, la solución de tres factores correlacionados resultó la más adecuada (CFI = 0,963; RMSEA = 0,074 [90% I.C. = (0,067- 0,081)]; TLI = 0,934).

La correlación entre factores fue alta: 0,646 entre el primer y el segundo factor; 0,575 entre el primero y el tercero; y 0,596 entre el segundo y el tercer factor.

Tabla 5
Índices de bondad de ajuste de los modelos estudiados en el AFE

Nº factores	χ^2 (g.l.)	<i>p</i>	RMSEA (90% I.C.)	CFI	TLI
1	2018,94 (65)	<0,000	0,146 (0,141-0,152)	0,778	0,734
2	1053,43 (53)	<0,000	0,116 (0,110-0,122)	0,886	0,833
3	364,06 (42)	<0,000	0,074 (0,067-0,081)	0,963	0,934

Nota: I.C.: intervalo de confianza. g.l.: grados de libertad. **En negrita:** la estructura seleccionada

Con respecto a los pesos de la solución rotada (ver Tabla 6), encontramos que todos los ítems muestran saturaciones por encima de 0,30; a excepción del ítem Q1, el cual pesaba en los tres factores con cargas más pequeñas por debajo de 0,30. Por su parte, el ítem Q13 saturó tanto en el segundo como en el tercer factor, teniendo esto sentido sustantivo, que comentaremos más adelante en el apartado de discusión.

Con respecto a los tres factores resultantes, el primero estaba formado por el ítem Q1, de contenido general, y los ítems Q2, Q3 y Q5, correspondientes a preguntas de habilidades funcionales o físicas; el segundo de ellos, incluyendo también el primer ítem, comprendió los

ítems Q4, Q6, Q7 y Q8, además del Q13, todos ellos relativos a la realización personal en distintas facetas; el tercer factor incluyó los ítems Q1 y del Q9 al Q13, estos últimos con el mismo formato de respuesta y diferente al del resto.

Tabla 6

Pesos estandarizados (AFE de tres factores) con rotación oblimin

Ítems	Factor 1	Factor 2	Factor 3
Q1	0,221	0,202	0,184
Q2	0,537	0,093	0,030
Q3	0,894	-0,008	0,022
Q4	0,154	0,746	-0,029
Q5	0,760	0,116	0,048
Q6	0,023	0,820	-0,006
Q7	-0,109	0,570	0,085
Q8	0,108	0,515	0,149
Q9	0,165	-0,062	0,756
Q10	0,040	-0,040	0,859
Q11	-0,048	0,120	0,709
Q12	-0,144	0,086	0,458
Q13	-0,143	0,309	0,085

Análisis factorial confirmatorio

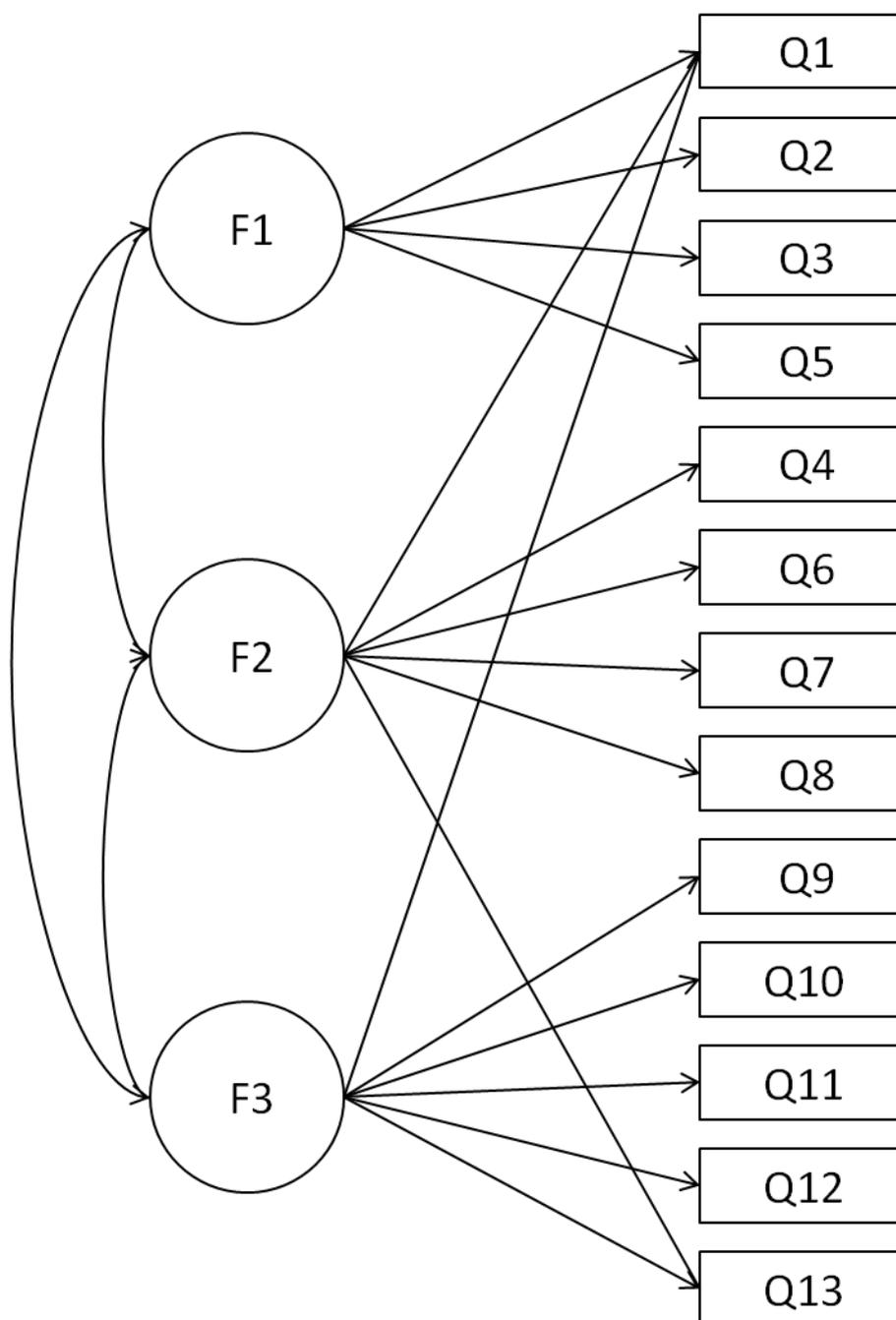
El modelo puesto a prueba aquí se muestra en la Figura 1. Para el 30% de la primera ola, los índices de bondad de ajuste mostraron valores buenos bajo esta estructura (RMSEA = 0,065 [90% I.C. = (0,055-0,074)]; CFI = 0,947; TLI = 0,929). En cuanto al AFC llevado a cabo en la muestra completa de la segunda ola de mediciones, los valores de los índices mostraron, en este caso, un ajuste aceptable (RMSEA = 0,075 [90% I.C. = (0,070- 0,080)]; CFI = 0,926; TLI = 0,902).

Los pesos estandarizados para la solución de tres factores correlacionados de primer orden se encontraron, en ambos análisis, por encima de 0,30; exceptuando el ítem Q1.

Dados estos resultados y la interpretación de los mismos, no se atendió a los índices de modificación. Además, cuando las muestras son grandes, un índice de modificación también grande puede estar sugiriendo la necesidad de agregar un parámetro cuya magnitud, estimado libremente, sería bastante trivial. Este problema es abordado por los software estadísticos

informando sobre los valores esperados de cambio para cada índice. Sin embargo, la inclusión de dichos parámetros no debe estar justificada por la búsqueda de un incremento del ajuste, sino que queda razonada solo cuando existe una base teórica previa convincente; y en el caso que nos ocupa, ésta no tiene lugar.

Figura 1. *Representación gráfica del modelo factorial confirmatorio de tres factores correlacionados (se han omitido los parámetros del modelo por simplicidad).*



Se encontraron correlaciones bastante altas: 0,777 entre el primer y el segundo factor; 0,746 entre el primero y el tercero; y 0,741 entre el segundo y el tercer factor en el caso de la primera ola; en la segunda, la correlación entre el primer y el segundo factor fue 0,776; entre el primero y el tercero 0,751; y 0,746 entre el segundo y el tercer factor.

Análisis de la invarianza longitudinal

El grado en el que el modelo de calidad de vida planteado (trece indicadores observables y tres factores correlacionados) mostró invarianza de medición se estudió en el presente trabajo empleando un AFC longitudinal. Inicialmente se especificó un modelo de invarianza configural con las especificaciones descritas en el apartado de método. Como se muestra en la Tabla 7, este modelo mostró un buen ajuste (RMSEA = 0,049; CFI = 0,938; TLI = 0,924).

En un siguiente paso, se puso a prueba el modelo de igualdad de cargas factoriales a lo largo del tiempo (invarianza métrica). Este modelo reveló unos índices de ajuste adecuados (RMSEA = 0,048; CFI = 0,938; TLI = 0,927), encontrándose que no ajustaba peor que el modelo de invarianza configural ($|\Delta\text{CFI}| < 0,001$; $|\Delta\text{RMSEA}| = 0,001$). Probado este nivel del invarianza, se garantizan relaciones equivalentes entre los factores y sus indicadores, permitiendo la comparación entre momentos.

Posteriormente, se examinó la igualdad a lo largo del tiempo de los interceptos de los indicadores con el modelo de invarianza escalar. Los resultados sugirieron un buen ajuste (RMSEA = 0,047; CFI = 0,938; TLI = 0,930) y no se dio ningún decremento en el índice CFI, pudiéndose concluir la invarianza factorial escalar. En este punto, podría hablarse de que ambos momentos temporales cuentan con las mismas puntuaciones esperadas en cada indicador cuando el nivel absoluto de rasgo es el mismo, en otras palabras, que las diferencias observadas en los indicadores a lo largo del tiempo se deben únicamente a diferencias en las variables latentes.

Con respecto al último nivel de invarianza de medida, se puso a prueba el modelo de invarianza estricta o de las varianzas error en los dos momentos temporales. De nuevo, teniendo en cuenta los índices de bondad de ajuste y la regla del cambio en $\text{CFI} < 0,01$, se satisfizo la invarianza estricta. Lo concluido hasta aquí puede interpretarse como que la cantidad de varianza de los indicadores no contabilizada por las variables latentes fue la misma en ambos momentos de medición.

Tabla 7

Modelos de invarianza longitudinal de medida estudiados.

Modelo	χ^2 (g.l.)	<i>p</i>	TLI	CFI	ΔCFI	RMSEA (90% I.C.)	ΔRMSEA
Modelo 1: Configural	1577,18 (265)	<0,000	0,924	0,938		0,049 (0,047-0,052)	
Modelo 2: Métrica	1595,09 (278)	<0,000	0,927	0,938	0,000	0,048 (0,046-0,051)	-0,001
Modelo 3: Escalar	1597,70 (287)	<0,000	0,930	0,938	0,000	0,047 (0,045-0,050)	-0,001
Modelo 4: Estricta	1686,99 (301)	<0,000	0,929	0,934	0,004	0,048 (0,045-0,050)	0,001

Nota: I.C.: intervalo de confianza. g.l.: grados de libertad.

Discusión y conclusiones

La posibilidad de comparar las medidas obtenidas en diferentes situaciones o momentos es un prerrequisito de toda ciencia, el cual exige el cumplimiento de la invarianza factorial (Elosua, 2005).

El objetivo principal de este trabajo fue analizar la invarianza de medida del cuestionario WHOQOL-AGE a lo largo del tiempo en una muestra de adultos españoles. Con esta finalidad, se llevaron a cabo pruebas de invarianza en el contexto del AFC longitudinal; previamente, se realizó un estudio completo de la dimensionalidad de los datos mediante un AFE y posterior AFC.

En un primer momento, se llevó a cabo un análisis del sesgo por atrición estudiando algunas variables sociodemográficas básicas, a saber: edad, sexo, estado civil, entorno residencial y años de educación. Se encontraron diferencias significativas en las variables edad, estado civil y años de educación entre aquellos participantes que únicamente contestaron a la primera ola y los que lo hicieron en ambas; sin embargo, los tamaños del efecto asociados fueron pequeños (Cohen, 1988). No parece haber, por tanto, una pérdida selectiva de casos que pudiera estar afectando a la representatividad de la muestra inicial; no obstante, sería interesante realizar un análisis de la atrición más completo, incluyendo otras variables que pudieran ser relevantes en el envejecimiento.

Partiendo de un estudio previo, se procedió a testar una estructura bifactor con dos factores específicos, la cual no mostró resultados aceptables. A pesar de que el factor general presentaba pesos elevados tanto en la anterior investigación como en ésta, tal y como se pudo ver en la comparación de ambos estudios, los pesos de los factores específicos sugerían la

escisión en dos dimensiones del primer factor de formato encontrado en Santos et al. (2018). Más aún, el AFE realizado posteriormente sugirió una estructura de tres factores que se correspondía con este patrón y que fue confirmado con índices de ajuste adecuados en ambas olas mediante un AFC.

Si nos fijamos en estos 8 primeros ítems (ver ítems en cursiva de la Tabla 1) que presentan el mismo formato de respuesta, tiene sentido una agrupación de los elementos Q2, Q3 y Q5, por un lado y Q4, Q6, Q7 y Q8, por otro. Los tres primeros ítems señalan una dimensión que ha sido identificada de forma recurrente en la evaluación de la CV: las habilidades funcionales relacionadas con el estado físico y la salud. Los cuatro ítems restantes, aunque más dispares, comparten un significante de realización personal concretado en lo individual, lo social y lo material a través de la comparación entre un estado ideal y uno real (Felce y Perry, 1995). Esto último es consistente con que el ítem Q13 (“*¿Cómo de satisfecho está con sus relaciones íntimas en su vida?*”) con un formato de respuesta diferente, muestre un peso superior a 0,30 también en ese factor, siendo una pregunta muy similar en contenido al ítem Q6 (“*¿Cómo de satisfecho está con sus relaciones personales?*”). El Q1, al tratarse de un ítem general (“*¿Cómo calificaría usted su calidad de vida?*”), presentó pesos bajos repartidos en los tres factores, algo con lo que también se toparon Caballero et al. (2013).

Así, una conclusión que se sigue de los resultados arrojados en este estudio es que los distintos formatos de respuesta crean ruido a la hora de identificar factores de contenido; factores que, al medir la CV, aportarían una interpretación de las puntuaciones más útil de lo que la división por formato ofrece.

En cuanto a las pruebas de invarianza factorial, se han proporcionado evidencias de que se midió el mismo constructo en ambos momentos temporales, demostrando que el WHOQOL-AGE es un instrumento invariante a través del tiempo. Tradicionalmente, encontrar estudios que analizaran la invarianza de medida a través del tiempo no ha sido tarea sencilla. Sin embargo, para garantizar que estamos modelando el cambio de un mismo constructo en distintos momentos es necesario utilizar este tipo de pruebas psicométricas, superando la mera suposición. Con lo encontrado aquí, es posible analizar y comparar las diferencias observadas en las puntuaciones en CV a lo largo del seguimiento, pudiéndose atribuir dichos cambios a diferencias en las variables latentes (cambio *alfa*) y no a inconsistencias en las propiedades de medida de los ítems (cambio *beta*). No fue necesario

aplicar procedimientos de invarianza parcial, por lo que se pudo concluir una invarianza de medida completa (modelo de invarianza estricta).

En los estudios anteriores sobre este cuestionario (Caballero et al., 2013; Santos et al., 2018) se encontraron, de una forma u otra, estructuras alternativas a la de factores correlacionados, que permitieron apoyar la existencia de un factor general de calidad de vida. En este estudio, todas las correlaciones entre factores fueron bastante altas, lo cual estaría indicando que un modelo bifactor sería lo más adecuado (Reise, Moore y Haviland, 2010). Sin embargo, aplicar un modelo bifactor a cualquier estructura es un tema no resuelto a día de hoy. Así, Abad et al. (2017) pusieron a prueba varios métodos de rotación bifactor bajo distintas condiciones, encontrando que en un escenario de complejidad estructural (p. ej., pesos cruzados) y de saturaciones bajas en los factores específicos (por debajo de 0,50), la recuperación de las estructuras propuestas no era adecuada.

Otro punto a comentar, a colación del párrafo anterior, está relacionado con la viabilidad de obtener una puntuación global de la escala. Estudios de validación de otros instrumentos de la familia de este cuestionario han encontrado apoyo para la unidimensionalidad, incluidos los dos estudios sobre el WHOQOL-AGE (Caballero et al., 2013; Santos et al., 2018). Los argumentos sobre si una medida debe puntuarse como reflejo de un solo constructo o dividirse en subescalas son muy comunes en contextos de evaluación cognitiva y no cognitiva (Reise, Moore y Haviland, 2010), aunque el primero de los casos es el más apoyado cuando los factores están muy correlacionados. Este estudio, con sus conclusiones sobre la estructura factorial de la escala, no da respuesta a la utilización de una puntuación global sobre CV, lo cual es una de sus principales limitaciones. Sería apropiado, pues, utilizar distintos métodos de cálculo de puntuaciones teniendo en cuenta el funcionamiento similar de los ítems que se agrupan formando los factores obtenidos.

Otra limitación tiene que ver con el único ítem que presentó respuestas en blanco, el Q13, sobre *relaciones íntimas*. Se contabilizaron un total de 366 valores perdidos en alguna de las dos olas. Para la realización de los análisis factoriales y pruebas de invarianza se prescindió de todos estos casos, que sin duda es una pérdida no aleatoria (MNAR, *missing not at random*). Al encontrarnos con estos datos, habría dos estrategias posibles a aplicar: la primera, suponer que la pérdida de casos es al azar (*missing at random*, MAR) –no confundir con MCAR o pérdida completamente al azar– (Raghunathan, 2004); la segunda, modelar la

probabilidad de los valores perdidos por MNAR a partir de la información de la propia variable y de otras relacionadas ella (Holman y Glas, 2005).

Por otra parte, este estudio presenta varias fortalezas. Una de ellas es que la muestra empleada presenta un gran tamaño, permitiendo análisis con mayor potencia estadística; además, procede de una muestra representativa de la población española. Otra, tiene que ver con los procedimientos utilizados en el análisis de la estructura factorial de los datos, y es que se han utilizado distintos criterios estadísticos y sustantivos para valorar cuál fue la mejor solución factorial.

Este trabajo se llevó a cabo dentro del marco de un plan financiado por la Comisión Europea cuyo principal objetivo es desarrollar instrumentos adecuados para evaluar la CV en población que está envejeciendo. Este proceso natural –estudiado a través de un diseño longitudinal– solo puede ser evaluado adecuadamente si los instrumentos empleados prueban su invarianza factorial (entre grupos y a lo largo del tiempo); y estando prevista la recogida de una tercera ola de datos, lo encontrado aquí da respaldo a las futuras comparaciones entre momentos en el tiempo.

Por otra parte, se proponen algunas líneas de investigación futuras a colación de los resultados de este trabajo. Una de ellas tiene que ver con el estudio de la invarianza estructural, no estudiada aquí. Habiéndose confirmado la invarianza de medida, comparar momentos sobre aspectos de las variables latentes (medias, varianzas y covarianzas) es posible. Dentro de este proyecto de diseño longitudinal, sería interesante poner a prueba la invarianza de las medias de los factores con el propósito de ver qué ocurre a lo largo del seguimiento (el cual va en paralelo al envejecimiento de la muestra) de cara a proponer medidas concretas que mejoren la CV de estas personas. La invarianza de las varianzas y covarianzas de los factores será útil cuando se quiera comprobar si la relación entre distintas facetas del constructo se mantiene o cambia con el paso de tiempo, pudiéndolo también relacionar con el proceso de envejecimiento.

Finalmente, sería interesante estudiar si otras medidas relacionadas con la CV –y que forman parte del cuestionario completo de seguimiento–, muestran el criterio de invarianza longitudinal, lo que supondría anticiparse a la extracción de conclusiones erróneas al comparar o relacionar puntuaciones en distintos momentos temporales.

Financiación

Los datos empleados en la realización del presente Trabajo Fin de Máster, fueron obtenidos con la financiación recibida por parte del VII Programa Marco de la Unión Europea (FP7/2007-2013) (referencia 223071: proyecto COURAGE in Europe), así como del Ministerio Español de Ciencia e Innovación (programa ACI-Promociona 2009-1010) y del Instituto de Salud Carlos III (a través de los proyectos FISPS09/00295, PS09/01845, PI12/01490 y PI13/00059), Los proyectos PI12/01490 y PI13/00059 han sido cofinanciados por el Fondo de Desarrollo Regional Europeo “Una manera de construir Europa”.

Referencias bibliográficas

- Abad, J.F., García-Garzón, E., Garrido, L. y Barrada, J.R. (2017). Iteration of Partially Specified Target Matrices: Application to the Bi-Factor Case. *Multivariate Behavioral Research*, 52, 416-429. doi: 10.1080/00273171.2017.1301244
- Abad, J. F., Olea, J., Ponsoda, V. y García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid: Editorial Síntesis.
- Anderson, J. C., y Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411-423.
- Bandalos, D. L. y Finney, S. J. (2010). Factor Analysis: Exploratory and Confirmatory. En G. R. Hancock y R. O. Mueller (Eds.), *Reviewer's guide to quantitative methods*. Routledge: New York.
- Bendezú, L., Denis, A. y Zubizareta, J.R. (2007). Análisis de la Atrición de la Muestra en la Encuesta Panel CASEN. Observatorio Social Universidad Alberto Hurtado.
- Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY, US: Guilford Press.
- Caballero, F. F., Miret, M., Power, M., Chatterji, S., Tobiasz- Adamczyk, B., Koskinen, S., et al. (2013). Validation of an instrument to evaluate quality of life in the aging population: WHOQOL-AGE. *Health and Quality of Life Outcomes*, 11, 177.
- Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods*, 1, 421-483. doi:10.1177/109442819814004
- Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464-504.
- Cheung, G. W., y Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233-255.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cochran, W. G. (1952). The X^2 test of goodness of fit. *Annals of Mathematical Statistics*, 23, 315-345.
- Dimitrov, D.M. (2010). Testing for factorial invariance in the context of construct validation. *Methods Plainly Speaking*, 43, 121-149. doi: 10.1177/0748175610373459
- Elosua, P. (2005). Evaluación progresiva de la invarianza factorial entre las versiones original y adaptada de una escala de autoconcepto. *Psicothema*, 17, 356-362
- Fang J, Power M, Lin Y, Zhang J, Hao, Y. y Chatterji, S. (2012). Development of short versions for the WHOQOL-OLD module. *Gerontologist*, 52, 66-78.
- Felce, D., y Perry, J. (1995). Quality of life: Its definition and measurement. *Research in developmental disabilities*, 16, 51-74.
- Finney, S. J., y DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G.R. Hancock y R.O. Mueller (Eds.), *Structural equation modeling: a second course* (269-314). Greenwich, CT: Information Age Publishing.
- Garrido, L.E., Abad, F.J. y Ponsoda, V. (2011). Performance of Velicer's Minimum Average Partial Factor Retention Method With Categorical Variables. *Educational and Psychological Measurement*, 71, 551-570. doi: 10.1177/0013164410389489
- Holman, R. y Glas, C.A. (2005). Modelling non-ignorable missing-data mechanisms with item response theory models. *The British Journal of Mathematical and Statistical Psychology*, 58, 1-17
- Horn, J. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185.
- Hooper, D., Coughlan, J. y Mullen, R.M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6, 53-60.

- Hu, L.T., y Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1-55.
- Koch, T., Schultze, M., Eid, M. y Geiser, C. (2014). A longitudinal multilevel CFA-MTMM model for interchangeable and structurally different methods. *Frontiers in psychology*, 5, 1-19. doi: 10.3389/fpsyg.2014.00311
- Leonardi, M., Chatterji, S., Koskinen, S., Ayuso-Mateos, J.L., Haro, J.M., et al. (2014). Determinants of health and disability in ageing population: The COURAGE in Europe Project (Collaborative research on ageing in Europe). *Clinical Psychology and Psychotherapy*, 21, 193-198.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A. y Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de psicología*, 30, 1151-1169.
- MacCallum, R., Roznowski, M. y Necowitz, L. (1992). Model modifications in covariance structure analysis: the problem of capitalization on chance. *Psychological Bulletin*, 3, 490-504.
- Martínez, J.A. (2009). Analysing scale invariance through integral calculus when measuring perceived quality in sports services. *International Journal of Sport Science*, 15, 17-35.
- Meredith, W. M. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525-543.
- Milfont, T.L. y Fischer, R. (2010). Testing measurement invariance across groups: Applications in crosscultural research. *International Journal of Psychological Research*, 3, 111-121.
- Miret, M., Caballero, F. F., Chatterji, S., Olaya, B., Tobiasz- Adamczyk, B., Koskinen, S., et al. (2014). Health and happiness: cross-sectional household surveys in Finland, Poland and Spain. *Bulletin of the World Health Organization*, 92, 716-725.
- McDonald, R. P. (1985). Factor Analysis and Related Methods. *Hillsdale, NJ: LEA*.

- Moreno-Jiménez, B. y Ximénez, C. (1996). La evaluación de la calidad de vida. En G. Buela., V. Caballo y J.C. Sierra (Eds.) *Manual de Evaluación en psicología Clínica y de la salud*. Madrid: Ed Siglo XXI.
- Muthén L. y Muthén B. (2011). Mplus (version 6.12) [software].
- Muthén, L. y Muthén, B. (2017). Mplus User's Guide. 8th ed, Los Ángeles, CA
- Pearson, R., Mundfrom, D. y Piccone, A. (2013). A Comparison of Ten Methods for Determining the Number of Factors in Exploratory Factor Analysis. *Multiple Linear Regression Viewpoints* 39, 1-15.
- Power, M., Quinn, K., y Schmidt, S. (2005). Development of the WHOQOL-old module. *Quality of Life Research*, 14, 2197-2214.
- Raghunathan, T.E. (2004). What do we do with missing data? Some options for analysis of incomplete data. *Annual Review of Public Health*, 25, 99-117.
- Schalock, R. L. (1996). *The quality of children's lives*. Springfield, IL: Charles C. Thomas.
- Schmidt, S., Muhlan, H. y Power, M. (2006). The EUROHIS-QOL 8-item index: psychometric results of a cross-cultural field study. *European Journal of Public Health*, 16, 420-428.
- Smith, K., Avis, N., y Assman, S. (1999). Distinguishing between quality of life and health status in quality of life research: A meta-analysis. *Quality of life research*, 8, 447-459.
- Santos, D., Abad, F.J., Chatterji, S., Olaya, B., Zawisza, K., Koskinen, S., Leonardi, M., Haro, J.M., Ayuso-Mateos, J.L., Caballero, F.F. (2017). Measurement invariance of the WHOQOL-AGE questionnaire across three European countries. *Quality of Life Research*, 27, 1015-1025.
- StataCorp. (2011). *Stata Statistical Software: Release 12*. College Station, TX: StataCorp LP.
- Steiger, J.H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173-180.
- Steiger, J.H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences*, 42, 893-898.
- Taillefer, M.C., Dupuis, G., Roberge, M.A., y Le May, S. (2003). Health related quality of life models: Systematic review of the literature. *Social Indicators Research*, 64, 293-323.

- The WHOQOL Group. (1994). Development of the WHOQOL: Rationale and current status. *International Journal of Mental Health, 23*, 24-56.
- The WHOQOL Group. (1995). The World Health Organization quality of life assessment (WHOQOL): Position paper from the World Health Organization. *Social Science & Medicine, 41*, 1403–1409.
- The WHOQOL Group (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine, 28*, 551–558.
- Tucker, L. y Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1–10.
- Urzúa, A y Caqueo-Urizar, A. (2012). Calidad de vida: una revisión teórica del concepto. *Terapia Psicológica, 30*, 61-71.
- Van de Schoot, R., Lugtig, P. y Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology, 9*, 486-492.
- Velicer, W. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 4*, 321–337.
- Wells, C. S. y Faulkner-Bond, M. (2016). *Educational Measurement. From Foundations to Future*. New York, NY: The Guilford Press.
- Widaman, K. F., y Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). Washington, DC: American Psychological Association.
- Widaman, F., Ferrer, E. y Conger, R. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*, 10-18
- Zwick, W. R., y Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432.

Anexos

Anexo I. Sintaxis Mplus AFE 70% primera ola

```
TITLE: Analisis factorial exploratorio
      DATA: FILE IS "C:/Users/HP1/Desktop/TFM
MPLUS/primer_ola_70_EFA.dat";
      VARIABLE:
NAMES ARE  Q1 Q2 Q3 Q5 Q4 Q6 Q8 Q8 Q9 Q10 Q11 Q12 Q13;

      usevariables Q1-Q13;

      ANALYSIS:
      TYPE = EFA 1 3;
      PARALLEL = 50;
      ESTIMATOR = ML;
      ROTATION = OBLIMIN;
      OUTPUT:
      SAMPSTAT;
      RESIDUAL;
```

Anexo II. Sintaxis Mplus AFC ambas olas

```
TITLE: Analisis factorial confirmatorio
      !DATA: FILE IS "C:/Users/HP1/Desktop/TFM
MPLUS/DATOS/primer_ola_30_CFA.dat";
      !DATA: FILE IS "C:/Users/HP1/Desktop/TFM
MPLUS/DATOS/segunda_ola_CFA_toda.dat";
      VARIABLE:

      NAMES ARE  Q1 Q2 Q3 Q5 Q4 Q6 Q8 Q8 Q9 Q10 Q11 Q12 Q13;

      usevariables Q1-Q13;

      ANALYSIS: ESTIMATOR = MLM;
      MODEL:

      F1 BY Q1-Q5;
      F2 BY Q1 Q4-Q7 Q13;
      F3 BY Q9-Q13;
      F1-F3 WITH F1-F3;

      OUTPUT: MOD(3.84); STANDARDIZED; RESIDUAL;
```

Anexo III. Sintaxis Mplus niveles de invarianza.

```
TITLE: Modelo 1 invarianza configural longitudinal
      DATA: FILE IS "C:/Users/HP1/Desktop/TFM
MPLUS/dos_olas_sin_valores_perdidos.dat";
      VARIABLE:

      NAMES ARE  Q1 Q2 Q3 Q5 Q4 Q6 Q8 Q8 Q9 Q10 Q11 Q12 Q13 Q1_2 Q2_2 Q3_2
      Q5_2 Q4_2 Q6_2 Q8_2 Q7_2 Q9_2 Q10_2 Q11_2 Q12_2 Q13_;

      usevariables Q1-Q13_2;

      ANALYSIS: ESTIMATOR = MLM;
      MODEL:

      !ola 1:
```

```

F1_1 BY Q1* Q2* Q3* Q5*;
F2_1 BY Q1* Q4* Q6* Q8* Q7* Q13*;
F3_1 BY Q1* Q9* Q10* Q11* Q12* Q13*;

[F1_1-F3_1@0];
[Q1-Q13*];

F1_1-F3_1@1;
F1_1-F3_1 WITH F1_1-F3_1*;

!ola 2:
F1_2 BY Q1_2* Q2_2* Q3_2* Q5_2*;
F2_2 BY Q1_2* Q4_2* Q6_2* Q8_2* Q7_2* Q13_2*;
F3_2 BY Q1_2* Q9_2* Q10_2* Q11_2* Q12_2* Q13_2*;

[F1_2-F3_2@0];
[Q1_2-Q13_2*];

F1_2-F3_2@1;
F1_2-F3_2 WITH F1_2-F3_2*;

Q1 WITH Q1_2*;
Q2 WITH Q2_2*;
Q3 WITH Q3_2*;
Q4 WITH Q4_2*;
Q5 WITH Q5_2*;
Q6 WITH Q6_2*;
Q7 WITH Q7_2*;
Q8 WITH Q8_2*;
Q9 WITH Q9_2*;
Q10 WITH Q10_2*;
Q11 WITH Q11_2*;
Q12 WITH Q12_2*;
Q13 WITH Q13_2*;

Q1-Q13;
Q1_2-Q13_2;

OUTPUT: MOD(3.84); STANDARDIZED; RESIDUAL;

TITLE: Modelo 2 invarianza metrica longitudinal
DATA: FILE IS "C:/Users/HP1/Desktop/TFM
MPLUS/dos_olas_sin_valores_perdidos.dat";
VARIABLE:

NAMES ARE Q1 Q2 Q3 Q5 Q4 Q6 Q8 Q8 Q9 Q10 Q11 Q12 Q13 Q1_2 Q2_2 Q3_2
Q5_2 Q4_2 Q6_2 Q8_2 Q7_2 Q9_2 Q10_2 Q11_2 Q12_2 Q13_;

usevariables Q1-Q13_2;

ANALYSIS: ESTIMATOR = MLM;
MODEL:

!ola 1:
F1_1 BY Q1* Q2* Q3* Q5* (11-14);
F2_1 BY Q1* Q4* Q6* Q8* Q7* Q13* (15-110);
F3_1 BY Q1* Q9* Q10* Q11* Q12* Q13* (111-116);

[F1_1-F3_1@0];

```

```

[Q1-Q13*];

F1_1-F3_1@1;
F1_1-F3_1 WITH F1_1-F3_1*;

!ola 2:
F1_2 BY Q1_2* Q2_2* Q3_2* Q5_2* (11-14);
F2_2 BY Q1_2* Q4_2 Q6_2* Q8_2* Q7_2* Q13_2* (15-110);
F3_2 BY Q1_2* Q9_2* Q10_2* Q11_2* Q12_2* Q13_2* (111-116);

[F1_2-F3_2@0];
[Q1_2-Q13_2*];

F1_2-F3_2*;
F1_2-F3_2 WITH F1_2-F3_2*;

Q1 WITH Q1_2*;
Q2 WITH Q2_2*;
Q3 WITH Q3_2*;
Q4 WITH Q4_2*;
Q5 WITH Q5_2*;
Q6 WITH Q6_2*;
Q7 WITH Q7_2*;
Q8 WITH Q8_2*;
Q9 WITH Q9_2*;
Q10 WITH Q10_2*;
Q11 WITH Q11_2*;
Q12 WITH Q12_2*;
Q13 WITH Q13_2*;

Q1-Q13;
Q1_2-Q13_2;

OUTPUT: MOD(3.84); STANDARDIZED; RESIDUAL;

```

TITLE: Modelo 3 invarianza escalar longitudinal

```

DATA: FILE IS "C:/Users/HP1/Desktop/TFM
MPLUS/dos_olas_sin_valores_perdidos.dat";
VARIABLE:

```

```

NAMES ARE Q1 Q2 Q3 Q5 Q4 Q6 Q8 Q8 Q9 Q10 Q11 Q12 Q13 Q1_2 Q2_2 Q3_2
Q5_2 Q4_2 Q6_2 Q8_2 Q7_2 Q9_2 Q10_2 Q11_2 Q12_2 Q13_;

```

```

usevariables Q1-Q13_2;

```

```

ANALYSIS: ESTIMATOR = MLM;
MODEL:

```

```

!ola 1:
F1_1 BY Q1* Q2* Q3* Q5* (11-14);
F2_1 BY Q1* Q4* Q6* Q8* Q7* Q13* (15-110);
F3_1 BY Q1* Q9* Q10* Q11* Q12* Q13* (111-116);

```

```

[F1_1-F3_1@0];
[Q1-Q13*] (I1-I13);

```

```

F1_1-F3_1@1;
F1_1-F3_1 WITH F1_1-F3_1*;

```

```

!ola 2:
F1_2 BY Q1_2* Q2_2* Q3_2* Q5_2* (11-14);
F2_2 BY Q1_2* Q4_2* Q6_2* Q8_2* Q7_2* Q13_2* (15-110);
F3_2 BY Q1_2* Q9_2* Q10_2* Q11_2* Q12_2* Q13_2* (111-116);

[F1_2-F3_2*];
[Q1_2-Q13_2*] (I1-I13);

F1_2-F3_2*;
F1_2-F3_2 WITH F1_2-F3_2*;

Q1 WITH Q1_2*;
Q2 WITH Q2_2*;
Q3 WITH Q3_2*;
Q4 WITH Q4_2*;
Q5 WITH Q5_2*;
Q6 WITH Q6_2*;
Q7 WITH Q7_2*;
Q8 WITH Q8_2*;
Q9 WITH Q9_2*;
Q10 WITH Q10_2*;
Q11 WITH Q11_2*;
Q12 WITH Q12_2*;
Q13 WITH Q13_2*;
Q13_2 WITH Q6_2;

Q1-Q13;
Q1_2-Q13_2;

OUTPUT: MOD(3.84); STANDARDIZED; RESIDUAL;

TITLE: Modelo 4 invarianza estricta longitudinal
DATA: FILE IS "C:/Users/HP1/Desktop/TFM
MPLUS/dos_olas_sin_valores_perdidos.dat";
VARIABLE:

NAMES ARE Q1 Q2 Q3 Q5 Q4 Q6 Q8 Q8 Q9 Q10 Q11 Q12 Q13 Q1_2 Q2_2 Q3_2
Q5_2 Q4_2 Q6_2 Q8_2 Q7_2 Q9_2 Q10_2 Q11_2 Q12_2 Q13_;

usevariables Q1-Q13_2;

ANALYSIS: ESTIMATOR = MLM;
MODEL:

!ola 1:
F1_1 BY Q1* Q2* Q3* Q5* (11-14);
F2_1 BY Q1* Q4* Q6* Q8* Q7* Q13* (15-110);
F3_1 BY Q1* Q9* Q10* Q11* Q12* Q13* (111-116);

[F1_1-F3_1@0];
[Q1-Q13*] (i1-i13);

F1_1-F3_1@1;
F1_1-F3_1 WITH F1_1-F3_1*;

!ola 2:
F1_2 BY Q1_2* Q2_2* Q3_2* Q5_2* (11-14);
F2_2 BY Q1_2* Q4_2* Q6_2* Q8_2* Q7_2* Q13_2* (15-110);

```

```

F3_2 BY Q1_2* Q9_2* Q10_2* Q11_2* Q12_2* Q13_2* (111-116);

[F1_2-F3_2*];
[Q1_2-Q13_2*] (i1-i13);

F1_2-F3_2* ;
F1_2-F3_2 WITH F1_2-F3_2*;

Q1 WITH Q1_2*;
Q2 WITH Q2_2*;
Q3 WITH Q3_2*;
Q4 WITH Q4_2*;
Q5 WITH Q5_2*;
Q6 WITH Q6_2*;
Q7 WITH Q7_2*;
Q8 WITH Q8_2*;
Q9 WITH Q9_2*;
Q10 WITH Q10_2*;
Q11 WITH Q11_2*;
Q12 WITH Q12_2*;
Q13 WITH Q13_2*;

Q1-Q13 (v1-v13);
Q1_2-Q13_2 (v1-v13);

OUTPUT: MOD(3.84); STANDARDIZED; RESIDUAL;

```

Anexo IV. Sintaxis STATA análisis de variables sociodemográficas

```

ssc install cohend

ttest q1011_age, by(attrition)
cohend q101_age1 attrition

ttest q1017_years, by(attrition)
cohend q1017_years attrition

tab q0104_ur attrition_group, chi V

tab q1009_sex attrition_group, chi V

tab q1012_smar_st attrition_group, chi V

```