

UNIVERSIDAD DE SALAMANCA

Departamento de Estadística



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

Tesis Doctoral

Doctorado en Estadística Multivariante Aplicada

**CONTRIBUCIONES A LA DETECCIÓN Y EVALUACIÓN DEL
CAMBIO EN CALIDAD DE VIDA (RESPONSE SHIFT) EN
PRESENCIA DE EVENTOS POSITIVOS BASADAS EN
MODELOS DE TRES VÍAS**

Autora: ESTELINA ORTEGA JALIL

Directoras: Dra. M^a Purificación Vicente-Galindo

Dra. M^a Purificación Galindo-Villardón

2022



Departamento de Estadística
Universidad de Salamanca

**CONTRIBUCIONES A LA DETECCIÓN Y EVALUACIÓN DEL CAMBIO EN
CALIDAD DE VIDA (RESPONSE SHIFT) EN PRESENCIA DE EVENTOS
POSITIVOS BASADAS EN MODELOS DE TRES VÍAS**

Trabajo de Investigación para optar al Grado de
Doctor en Estadística Multivariante Aplicada
por el Departamento de Estadística de la
Universidad de Salamanca

Estelina Ortega Jalil

Salamanca

2022



Departamento de Estadística
Universidad de Salamanca

DRA. M^a PURIFICACIÓN VICENTE-GALINDO

Profesora Titular del Área de Estadística e Investigación Operativa de la Universidad
de Salamanca

y

DRA. M^a PURIFICACIÓN GALINDO-VILLARDON

Catedrática del Área de Estadística e Investigación Operativa de la Universidad de
Salamanca

Certifican que **Dña. ESTELINA ORTEGA JALIL** ha realizado en el Departamento de Estadística de la Universidad de Salamanca, bajo su dirección, el trabajo de investigación para optar al título de Doctor en Estadística Multivariante Aplicada que presenta con el título “**CONTRIBUCIONES A LA DETECCIÓN Y EVALUACIÓN DEL CAMBIO EN CALIDAD DE VIDA (RESPONSE SHIFT) EN PRESENCIA DE EVENTOS POSITIVOS BASADAS EN MODELOS DE TRES VÍAS**”, autorizando expresamente su lectura y defensa. Para ello, firman el presente certificado en Salamanca el 20 de mayo de 2022.

Dra. M^a Purificación Vicente-Galindo

Dra. M^a Purificación Galindo-Villardón

*“Nadie tiene éxitos sin esfuerzos.
Aquellos que tienen éxito
se lo deben
a la perseverancia”*

Ramana Maharshi

A ti, mi siempre Carlos Gómez, y a mi dos hermosos hijos Andrea y Carlos Manuel...

Los amo por completo

AGRADECIMIENTOS

Primeramente, quiero agradecer a Dios por permitirme culminar esta meta de formación académica y acompañarme durante todo este proceso.

A mi directora, Dra. M^a Purificación Vicente-Galindo por la siempre confianza depositada en mí y sus continuas recomendaciones. A mi tutora, Dra. M^a Purificación Galindo-Villardón por siempre transferirme esa fuerza y motivación para no desistir y poder alcanzar este objetivo. Gracias Puri.

Al Dr. Francisco Perdomo-Argüello por facilitarme los datos y resultados de su investigación como insumos para la elaboración de este estudio.

A la Dra. Carmen Rodríguez por brindarme su apoyo incondicional a lo largo de los años de estudio, durante nuestra estancia en conjunto, y todo el soporte mientras elaboraba este trabajo final. Gracias amiga. A la Dra. Mitzi Cubilla por ser esa voz de estímulo para nunca desmayar y haberme considerado desde el principio.

A la Dra. Ana Belén Nieto, a la Dra. Nerea González y a la Dra. Helena Martín-Rodero mi profundo agradecimiento por toda su disposición, facilitación y recomendaciones para lograr los avances significativos que contribuyeron a esta investigación. Aprecio inmensamente su amistad.

A mis compañeros Aurora, Adela, Gonzalo, Elena y Parker por sus palabras de aliento para finalizar este proyecto.

A Fulvia mi bella madre y en especial a mis hermanas Gisell y Bebo, por siempre estar para mí en todo momento.

RESUMEN

La calidad de vida es un concepto que supone la condición de un individuo en una sociedad en un momento determinado y el cual está condicionado por aspectos propios de la persona o por circunstancias del entorno. Cuando se considera el estado de salud de dicho individuo, se habla de calidad de vida relacionada con la salud.

La estabilidad se refiere a las condiciones de respuestas similares de un individuo a lo largo del tiempo; sin embargo, cuando éste manifiesta variaciones, se ha producido un cambio. El nuevo enfoque que estudia la condición de respuesta de una persona sujeta a diferentes eventos durante un período de tiempo se conoce como *Response Shift*.

La calidad de vida relacionada con la salud, la detección del *Response Shift* y sus diferentes tipos ha sido sujeto a la aplicación de diversas metodologías en múltiples áreas de estudio. La literatura señala que entre las técnicas de mayor difusión se encuentran el *Then-Test* y los métodos multivariantes.

Una revisión sistemática realizada para esta investigación, y que abarcó los años 2010 a 2020, evidenció que el método *Then-Test* y los Modelos de Ecuaciones Estructurales de Oort son los métodos más utilizados para la detección de este fenómeno. Otros modelos utilizados fueron la Regresión Lineal Múltiple, Regresión de Efectos Mixtos, Análisis de Trayectoria Latente, Teoría de Respuesta al Ítem, Regresión Logística, Regresión y Clasificación de Árbol, y los Métodos de Importancia Relativa. Mientras que otros métodos se aplicaron en un número menor de estudios.

Debido a la estructura de los datos de Calidad de Vida Relacionada con la Salud que son datos en los que se evalúa la autopercepción de la calidad de vida de los pacientes, a través de los ítems de cuestionarios genéricos y/o específicos y su evaluación a lo largo del tiempo, los métodos para analizar Tablas de Tres Vías representan una alternativa más potente que las utilizadas actualmente para el análisis del *Response Shift* y sus diferentes tipos, ya que permite comparar las estructuras factoriales de covariación de los ítems del cuestionario utilizado para la evaluación del constructo latente y analizar su evolución en el tiempo. [Vicente-Galindo \(2003\)](#) demostró el STATIS-Dual ([Escoufier, 1976a](#); [L'Hermier des Plantes, 1976](#)) puede ser utilizado

para estudiar Reconceptualización a través de la imagen euclídea que representa las estructuras en los diferentes tiempos, en el espacio estructurado por la métrica de Hilbert Schmit, Recalibración evaluando las normas de los vectores que representan las matrices de operadores en ese espacio euclídeo y Repriorización analizando la inercia absorbida por los diferentes ejes factoriales en los análisis individuales de cada tiempo.

Teniendo en cuenta que el análisis STATIS-Dual declara congruentes las estructuras latentes que comparten la primera dimensión latente y el hecho de que los cuestionarios de Calidad de Vida son siempre (o casi siempre) multifactoriales, es evidente que esta línea de investigación abierta por [Vicente-Galindo \(2003\)](#) puede y debe seguir avanzando a utilizar técnicas de integración de estructuras que capturen la naturaleza multidimensional del constructo de Calidad de Vida Relacionada con la Salud. El trabajo de investigación de esta tesis doctoral supone un avance para el conocimiento basado en el MetaBiplot propuesto por [Martín-Rodríguez \(1996\)](#) y [Martín-Rodríguez et al.\(2002\)](#)

Nuestra propuesta de investigación es la aplicación del método MetaBiplot que representa en un subespacio consenso los elementos que permiten evaluar la condición de cambio en diferentes momentos. Tras la revisión teórica de los fundamentos del MetaBiplot, se establecieron criterios que permiten la detección del Response Shift e identificar sus tipos en datos de calidad de vida relacionado con la salud.

Una vez realizado los análisis mediante la aplicación del método MetaBiplot, los resultados del estudio permitieron evidenciar la detección del *Response Shift*, la identificación de *Response Shift* de tipo Recalibración, y la no existencia de *Response Shift* del tipo Reconceptualización.

El trabajo termina poniendo de manifiesto que el MetaBiplot es una herramienta multivariante de integración de Tablas de Tres Vías que proporciona excelentes resultados en la detección del *Response Shift*.

INDICE GENERAL

AGRADECIMIENTOS	xii
RESUMEN	xiii
INDICE GENERAL	xv
INDICE DE TABLAS	xix
INDICE DE FIGURAS	xxi
INTRODUCCIÓN	xxiii
1. CALIDAD DE VIDA RELACIONADA CON LA SALUD Y <i>RESPONSE SHIFT</i> : UNA REVISIÓN CONCEPTUAL.....	2
1.1. Calidad de vida.....	2
1.1.1. Calidad de vida: una revisión conceptual	2
1.1.2. Calidad de vida relacionada con la salud.....	12
1.2. Estabilidad y cambio (<i>Response Shift</i>) en la calidad de vida relacionada con la salud	27
1.2.1. Estabilidad y cambio.....	27
1.2.2. <i>Response Shift</i> : un nuevo enfoque sobre el cambio	31
1.2.3. Modelo teórico del <i>Response Shift</i>	39
1.2.3.1. El modelo Sprangers y Schwartz.....	39
1.2.3.2. El modelo Lepore y Eton.....	40
1.2.3.3. El modelo Wilson y Cleary	41
1.2.3.4. El modelo Rapkin y Schwartz	42
2. MÉTODOS PARA LA DETECCIÓN DE <i>RESPONSE SHIFT</i> EN CALIDAD DE VIDA RELACIONADA CON LA SALUD: UNA REVISIÓN SISTEMÁTICA	45
2.1. Introducción	45
2.2. Revisión sistemática: Una metodología de búsqueda	46

2.2.1.	Conceptualización.....	46
2.2.2.	Fuentes de datos y estrategia de búsqueda.....	48
2.2.2.1.	Fuentes de datos	48
2.2.2.2.	Estrategias de búsqueda.....	49
2.2.3.	Selección de los artículos.....	50
2.2.4.	Calidad de los artículos seleccionados.....	50
2.3.	Una revisión sistemática sobre los métodos para detectar <i>Response Shift</i> en calidad de vida relacionada con la salud	52
2.3.1.	Fuente de datos	53
2.3.2.	Estrategia de búsqueda.....	53
2.3.3.	Selección de artículos	54
2.3.4.	Calidad del estudio.....	56
2.3.5.	Características descriptivas del estudio	57
2.3.6.	Sobre los métodos de detección de <i>Response Shift</i>	66
2.3.7.	Análisis crítico de la revisión sistemática.....	70
2.3.8.	Consideraciones finales	73
3.	MÉTODOS ESTADÍSTICOS PARA LA DETECCIÓN DEL <i>RESPONSE SHIFT</i>	75
3.1.	Características de los estudios de calidad de vida relacionado con la salud	75
3.2.	Métodos tradicionales para el análisis de calidad de vida relacionada con la salud y detectar <i>Response Shift</i>	77
3.2.1.	<i>Then- Test</i> : Técnica para detectar <i>Response Shift</i> en CVRS	78
3.3.	Métodos multivariantes para detectar <i>Response Shift</i> en calidad de vida relacionada con la salud	83
3.3.1.	El método de Ahmavaara.....	84
3.3.2.	El modelo de Zwinderman.....	85
3.3.3.	El modelo Mayo et al.	87

3.3.4.	Los Modelos de Ecuaciones Estructurales.....	89
3.3.4.1.	Los modelos de Curvas de Crecimiento Latente.....	100
3.3.4.2.	El modelo Zee.....	104
3.3.4.3.	El modelo Schmitt.....	106
3.3.4.4.	El planteamiento de Taris.....	108
3.3.4.5.	El modelo Oort.....	110
3.4.	Nuevos métodos multivariantes propuestos para la detección del <i>Response Shift</i> en calidad de vida relacionada con la salud.....	117
3.4.1.	El modelo de la Teoría de Respuesta al Ítem.....	117
3.4.1.1.	Características generales y supuestos de los modelos de la Teoría de Respuesta al Ítem.....	118
3.4.1.2.	Modelo logístico de dos parámetros.....	119
3.4.1.3.	Aplicación de los modelos de la Teoría de Respuesta al Ítem en la detección de <i>Response Shift</i>	120
3.4.2.	Modelos de Efectos Mixtos.....	123
3.4.3.	Medidas de Importancia Relativa.....	125
3.4.3.1.	Métodos 1: Método basado en la prueba de cambio en las ponderaciones.....	126
3.4.3.2.	Método 2: Prueba de cambio en los rangos del dominio.....	127
3.4.4.	Método de Clasificación y Regresión de Árbol.....	128
3.4.5.	Método STATIS Dual como herramienta para evaluar <i>Response Shift</i>	131
4.	METABILOT: UNA PROPUESTA PARA LA DETECCIÓN DEL <i>RESPONSE SHIFT</i>	136
4.1.	Métodos para detectar cambios en espacios consenso.....	136
4.2.	MetaBiplot en la detección de <i>Response Shift</i>	137
4.2.1.	MetaBiplot en el contexto de Calidad de Vida.....	138

4.2.2.	Método MetaBiplot.....	141
4.2.3.	Evaluación del subespacio consenso	143
4.2.4.	Evaluación de las trayectorias de los ítems	146
4.2.5.	MetaBiplot y el <i>Response Shift</i> : criterios para su detección.....	147
5.	EVALUACIÓN DEL <i>RESPONSE SHIFT</i> EN LA PERCEPCIÓN DE CALIDAD DE VIDA DE PACIENTES MIOPESES SOMETIDOS A CIRUGÍA REFRACTIVA CON LÁSER EXCIMER UTILIZANDO EL METABILOT.....	150
5.1.	Descripción del caso de estudio	150
5.2.	Análisis sobre la calidad de las escalas de medida	155
5.3.	El MetaBiplot como herramienta de inspección de <i>Response Shift</i> en pacientes operados de miopía.....	158
5.4.	Consideraciones de los resultados.....	167
	CONCLUSIONES	170
	REFERENCIAS BIBLIOGRÁFICAS	174
	ANEXOS	206
	ANEXO 1. Publicación de artículo	206
	ANEXO 2. Trabajos en proceso de publicar.....	217
	ANEXO 3. Participación en congresos nacionales e internacionales	218

INDICE DE TABLAS

Tabla 1. Dimensiones de la calidad de vida.....	9
Tabla 2. Cuatro aspectos de la Calidad de Vida	12
Tabla 3. Preguntas de investigación desde la perspectiva de la calidad de vida relacionadas con la salud	14
Tabla 4. Requerimientos de validación de la escala de desempeño de la calidad de vida según propósito del análisis	16
Tabla 5. Supuestos requeridos para la medición de la calidad de vida relacionada con la salud y problemas identificados.....	17
Tabla 6. Instrumentos utilizados para la medición de la calidad de vida relacionada con la salud	19
Tabla 7. Cuestionarios genéricos para evaluar la calidad de vida relacionada con la salud.....	20
Tabla 8. Cuestionarios específicos para evaluar la calidad de vida relacionada con la salud.....	23
Tabla 9. Comparación de los componentes asociados al Response Shift por los diferentes autores	33
Tabla 10. Preguntas utilizadas como lista de verificación.....	51
Tabla 11. Artículos incluidos en la revisión sistemática: Enero 2010 – Julio 2020 (1/6).....	58
Tabla 12. Características de los artículos incluidos en la revisión (n=107).....	64
Tabla 13. Instrumentos más utilizados en los estudios evaluados	66
Tabla 14. Número de veces que fueron utilizados los métodos para la detección de Response Shift	68
Tabla 15. Número de artículos según método de detección y tipo de Response Shift	69
Tabla 16. Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift utilizando el método Then- Test (2010-2020)	80

Tabla 17. Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift utilizando los modelos de Curvas de Crecimiento Latente (2010-2020).....	104
Tabla 18. Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift, utilizando el Modelo Oort (2010-2020).....	115
Tabla 19. Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift, utilizando los Modelos de la Teoría de Respuesta al Ítem (2010-2020)	122
Tabla 20. Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift, utilizando el Modelo de Efectos Mixtos (2010-2020)	124
Tabla 21. Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift, utilizando las Medidas de Importancia Relativa (2010-2020)	128
Tabla 22. Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift, utilizando el Método de Clasificación y Regresión de Árbol (2010-2020)	131
Tabla 23. Características Demográficas de los pacientes sometidos a la Cirugía Excimer Laser Corneal. Agosto 2012 a septiembre 2013	152
Tabla 24. Descripción de las Variables del NEI-VFQ-25	153
Tabla 25. Estadísticos de cada ítem del cuestionario NEI-VFQ-25 en cada uno de los momentos evaluados.....	155
Tabla 26. Análisis de las medias de las dimensiones en los diferentes momentos (preoperatorio, al día siguiente, al mes siguiente, tres meses después)	157
Tabla 27. Valores propios del MetaBiplot.....	159
Tabla 28. Cosenos cuadrados del subespacio consenso según los momentos.....	159
Tabla 29. Porcentaje de inercia absorbida por el consenso según momento	160
Tabla 30. Cosenos de los ángulos entre componentes comunes según momento ..	160

INDICE DE FIGURAS

Figura 1. Dominios de la calidad de vida	3
Figura 2. Modelo de Sprangers & Schwartz de Response Shift de calidad de vida .	40
Figura 3. Fases de la revisión sistemática utilizando la metodología PRISMA	47
Figura 4. Fuentes alternativas para la búsqueda de literatura en una revisión sistemática.....	48
Figura 5. Pasos para la selección de los artículos incluidos en la síntesis cualitativa	50
Figura 6. Diagrama de flujo PRISMA de la estrategia de búsqueda y los artículos incluidos en la revisión	56
Figura 7. Número de publicaciones por año: 2010-2020	57
Figura 8. Especialidades evaluadas en los diferentes artículos seleccionados	65
Figura 9. Número de artículos (n=107), detección y tipo de detección de Response Shift.....	66
Figura 10. Artículos según número de Métodos utilizados para detectar Response Shift.....	67
Figura 11. Frecuencia de los métodos utilizados para la detección de Response Shift por año	70
Figura 12. Esquema de relaciones en el método Then-Test	80
Figura 13. Esquema del método Ahmavaara	85
Figura 14. Metodología del desarrollo de un modelo teórico.....	90
Figura 15. Representación de las relaciones causales a través del diagrama de secuencia.....	91
Figura 16. Representación del modelo de la Teoría de Respuesta al Ítem	117
Figura 17. Representación de los parámetros de la Curva Característica del Ítem	119
Figura 18. Modelo de Clasificación y Regresión de Árbol	129

Figura 19. Representación de los datos de Calidad de Vida: Pacientes (I), Ítems (J) y momentos (T) como datos de Tres Vías	139
Figura 20. Proceso de generación del subespacio consenso, basado en el MetaBiplot	142
Figura 21. Caso 1: Existencia de una estructura común	144
Figura 22. Caso 2: Estructura común con normas diferente.....	144
Figura 23. Caso 3: Vectores mal representados.....	145
Figura 24. Caso 4: Ausencia de una estructura común.....	145
Figura 25. Trayectorias envolventes de un ítem	146
Figura 26. Trayectorias excéntricas de un ítem	147
Figura 27. Valores de las medias de los ítems en los diferentes momentos (preoperatorio, al día siguiente, al mes siguiente, tres meses después)	156
Figura 28. Representación de las variables (ítems) en el subespacio consenso (MetaBiplot).....	161
Figura 29. Ítems con trayectorias envolventes	165
Figura 30. Ítems con trayectorias excéntricas.....	166

INTRODUCCIÓN

La calidad de vida es un concepto que supone la condición de un individuo en una sociedad en un momento determinado y el cual está condicionado por aspectos propios de la persona o por circunstancias del entorno. Cuando se considera el estado de salud de dicho individuo, se habla de calidad de vida relacionada con la salud (CVRS).

De acuerdo con lo planteado por [Howard y Dailey \(1979\)](#), la estabilidad se refiere a las condiciones de respuestas similares de un individuo a lo largo del tiempo; sin embargo, cuando éste manifiesta variaciones, se ha producido un cambio. El nuevo enfoque que estudia la condición de respuesta de una persona sujeta a diferentes eventos durante un período de tiempo se conoce como *Response Shift*

Los estudios de cambios en la calidad de vida fueron expandidos al contexto de la psicología de la salud por [Schwartz y Sprangers \(1999\)](#). Al examinar el comportamiento de los individuos desde una perspectiva clínica, estos autores precisan que el *Response Shift* se genera como el cambio en el significado de una autoevaluación aplicada a un sujeto. Por medio de este enfoque se generan tres premisas fundamentales: la existencia de un cambio en los estándares internos de la medida (Recalibración), la existencia de un cambio en los valores (Repriorización) y la existencia de un cambio en la definición del constructo (Reconceptualización).

La CVRS, la detección del *Response Shift* y sus diferentes tipos ha sido sujeto a la aplicación de diversas metodologías en múltiples áreas de estudio. La literatura señala que entre las técnicas de mayor uso se encuentran el *Then-Test* y los métodos multivariantes. De este último grupo, los Modelos de Ecuaciones Estructurales han sido los más empleados por su fortaleza metodológica. Sin embargo, otros modelos se han explorado para encontrar nuevas metodologías para examinar el fenómeno: Teoría de Respuesta al Ítem, Efectos Mixtos, Medidas de Importancia Relativa y el Método de Clasificación y Regresión de Árbol.

La metodología del *Then-Test* es ampliamente utilizada por especialistas del área de la salud, quienes realizan estudios sobre la calidad de vida en este contexto, y aplican una técnica estadística basada en las diferencias de los promedios de los resultados de

los datos recolectados, y con ellos logran emitir un juicio ponderado sobre el comportamiento de las variables, en este caso, los pacientes. Debido a que esta técnica se basa en valores promedios, la misma no captura la esencia específica de las variables, es sensible al sesgo, sólo se puede utilizar en la evaluación de datos primarios y no permite la identificación de *Response Shift* de tipo Repriorización y Reconceptualización (Sajobi et al., 2018). Ante esta situación, el *Then-Test* representa una metodología aún considerada, pero limita la capacidad de explicación del fenómeno de *Response Shift*.

Con el interés de reducir la probabilidad de sesgos en los estudios de CVRS, se han utilizados métodos estadísticos multivariantes con el propósito de detectar de manera más precisa los cambios que se puedan producir en la condición de vida de un individuo o un grupo de individuos ante la exposición a un evento relacionado con su salud.

Sin embargo, la aplicación de métodos estadísticos multivariantes demanda del conocimiento profundo e implementación de estas técnicas, mientras que sus principales usuarios no son necesariamente estadísticos. Ante esta situación surge la oportunidad de seguir explorando otras técnicas multivariantes que permitan, mediante metodologías igualmente rigurosas, pero de fácil interpretación, poder evaluar el fenómeno del *Response Shift* e identificar sus diferentes tipos.

Debido a la estructura de los datos de CVRS que son datos en los que se evalúa la autopercepción de la calidad de vida de los pacientes, a través de los ítems de cuestionarios genéricos y/o específicos y su evolución a lo largo del tiempo, los métodos para analizar Tablas de Tres Vías representan una alternativa más potente que las utilizadas actualmente para el análisis del *Response Shift* y sus diferentes tipos, ya que permite comparar las estructuras factoriales de covariación de los ítems del cuestionario utilizado para la evaluación del constructo latente y analizar su evolución en el tiempo. Vicente-Galindo (2003) demostró que el STATIS-Dual (Escoufier, 1976a; L'Hermier des Plantes, 1976) puede ser utilizado para estudiar Reconceptualización a través de la imagen euclídea que representa las estructuras en los diferentes tiempos, en el espacio estructurado por la métrica de Hilbert Schmit, Recalibración evaluando las normas de los vectores que representan las matrices de

operadores en ese espacio euclídeo, y Reconceptualización analizando la inercia absorbida por los diferentes ejes factoriales en los análisis individuales de cada tiempo.

Teniendo en cuenta que el análisis STATIS-Dual declara congruentes las estructuras latentes que comparten la primera dimensión latente y el hecho de que los cuestionarios de Calidad de Vida son siempre (o casi siempre) multifactoriales, es evidente que esta línea de investigación abierta por [Vicente-Galindo \(2003\)](#) puede y debe seguir avanzando a utilizar técnicas de integración de estructuras que capturen la naturaleza multidimensional del constructo de CVRS. El trabajo de investigación de esta tesis doctoral supone un avance para el conocimiento basado en el MetaBiplot propuesto por [Martín-Rodríguez \(1996\)](#) y [Martín-Rodríguez et al.\(2002\)](#).

Ante esta nueva línea de investigación propuesta, surge la siguiente pregunta: ¿Existen otros métodos multivariantes de Tablas de Tres Vías que permitan detectar el fenómeno de *Response Shift* y lograr identificar sus diferentes tipos?

Luego de examinar diferentes técnicas multivariantes de Tablas de Tres Vías, nuestra investigación tiene como propósito explorar la aplicación del método MetaBiplot como una técnica alternativa para la detección del *Response Shift* y la identificación de sus diferentes tipos. La propuesta surge al reconocer que el MetaBiplot se fundamenta en la obtención de un subespacio consenso que permite evaluar los elementos (individuos y variables) en los diferentes momentos, siendo estas características similares al comportamiento de los datos de calidad de vida relacionada con la salud.

El desarrollo de esta investigación se conforma por cinco capítulos. El capítulo I, Calidad de vida relacionada con la salud y *Response Shift*: una revisión conceptual, presenta una revisión de la literatura que permite enmarcar el contexto de calidad de vida, su relación con la salud, ¿qué es el *Response Shift*? y bajo qué tipos se manifiesta, así como entender el modelo teórico desarrollado para explicar el fenómeno.

El capítulo II, Métodos para la detección de *Response Shift* en calidad de vida relacionada con la salud: una revisión sistemática, corresponde a una revisión sistemática de la literatura científica realizada con base a los años 2010 a 2020, y cuyo propósito es identificar las metodologías estadísticas utilizadas para la detección del

cambio, tipos de cambios que lograron identificar cada método y los contextos en que se utilizaron estas técnicas.

Con base a los resultados de esta revisión sistemática, se desarrolla en el capítulo III, Métodos estadísticos para la detección del *Response Shift*, el cual presenta los fundamentos teóricos de cada método estadístico resultante de la revisión y aquellos otros de alta relevancia que surgen de la revisión de la literatura.

El capítulo IV, MetaBiplot: una propuesta para la detección del *Response Shift*, presenta los elementos conceptuales y metodológicos de los métodos Biplot y sus diferentes tipos, las características y procedimientos del MetaBiplot, y finalmente se evalúan y establecen criterios teóricos para el análisis de la detección del *Response Shift* mediante esta técnica.

El capítulo V, Aplicación del MetaBiplot en datos de calidad de vida relacionados con la salud, toma como base un caso de estudio de datos correspondiente a una muestra de pacientes que fueron sometidos a cirugía ocular y a su vez evaluada su autopercepción en cuanto a su calidad de vida relacionada con la visión en diferentes momentos ([Perdomo-Arguello, 2015](#)).

Seguido se realiza un análisis sobre la calidad de las escalas de medida con el propósito de examinar si los datos poseen escalas de medidas diferentes que requieran ser transformados para lograr su estandarización general, y luego poder ser interpretados. Posteriormente, se someten los datos del caso de estudio al análisis del MetaBiplot para contrastar los resultados obtenidos con respecto a los criterios teóricos establecidos. Los resultados finales serán evaluados y se emitirán conclusiones sobre lo alcanzado con esta nueva técnica aplicada a los datos de CVRS.

Se espera que luego de realizar los análisis y experimentaciones correspondientes, y comprobar la adecuación de la funcionalidad de esta técnica multivariante a los datos de calidad de vida, se permita disponer de un método alternativo para la detección de *Response Shift* y sus diferentes tipos.

El trabajo termina poniendo de manifiesto que el MetaBiplot es una herramienta Multivariante de Integración de Tablas de Tres Vías que proporciona excelentes resultados en la detección del *Response Shift*.

CAPÍTULO I

CALIDAD DE VIDA RELACIONADA CON LA SALUD Y RESPONSE SHIFT: UNA REVISIÓN CONCEPTUAL

1. CALIDAD DE VIDA RELACIONADA CON LA SALUD Y *RESPONSE SHIFT*: UNA REVISIÓN CONCEPTUAL

La calidad de vida es un concepto que supone la condición de un individuo o una sociedad en un momento determinado y el cual está condicionado por aspectos propios de la persona o por circunstancias del entorno. Consecuentemente, bajo este concepto se puede inferir cuál es el nivel de bienestar de una persona en el tiempo, siendo lo último de carácter dinámico.

La revisión bibliográfica sobre este concepto lleva a entender que la calidad de vida posee un alto componente subjetivo que no necesariamente es similar entre pares, o sea entre individuos. Mientras para algunos, la calidad de vida puede mejorar cuando se logra obtener alimentos, vivienda o educación; para otros, la calidad de vida cambia favorablemente cuando existen lujos o posiciones sociales de alto impacto.

Como este criterio es tan abstracto como los puntos de referencia necesarios para establecer el nivel de calidad de vida de un caso en particular, en este capítulo se realizará una revisión bibliográfica que permita profundizar la conceptualización de este término, el cual abarca muchos contextos como los de tipo sociales, económicos e incluso clínicos.

1.1. Calidad de vida

1.1.1. Calidad de vida: una revisión conceptual

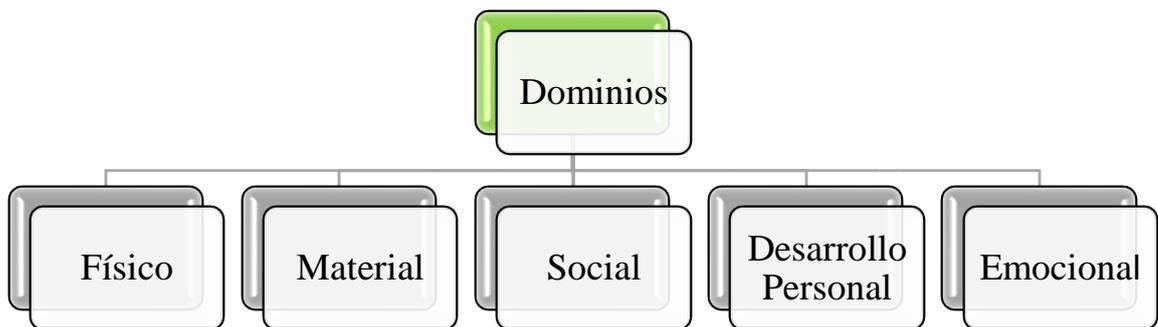
La calidad de vida es un concepto que hace alusión a varios niveles de la generalidad. Considera el bienestar de la colectividad y aspectos específicos de carácter individual. Por lo tanto, calidad de vida se conceptualiza desde una visión tan amplia como el enfoque que incida en el ser humano.

La Organización Mundial de la Salud (WHOQOL, 1998) define la calidad de vida como la percepción de los individuos sobre su condición de vida en el contexto cultural y sus sistemas de valores en los cuales viven y en relación con sus objetivos, expectativas, estándares y motivaciones. Es un concepto de amplio espectro afectado de manera compleja por la salud física de las personas, condiciones psicológicas, nivel de independencia, relaciones sociales, creencias personales y las interacciones con su medio ambiente.

La calidad de vida es un término del cual se ha intentado establecer una definición que abarque todas las áreas que integre los componentes subjetivos y objetivos considerando que el punto en común es el bienestar individual. Desde una óptica general, se pueden derivar cinco 5 dominios:

- Físico (como salud, seguridad física)
- Material (privacidad, alimentos, vivienda, transporte, posesiones)
- Social (relaciones interpersonales con la familia, amistades)
- Desarrollo personal (educación, productividad, contribución)
- Emocional (autoestima, estado respecto a los demás, religión)

Figura 1. *Dominios de la calidad de vida*



Fuente: Organización Mundial de la Salud (1998)

Sin embargo, es importante comprender que la respuesta a cada uno de estos dominios es subjetiva y tan variable gracias a la influencia de factores sociales, materiales, la edad misma, la situación de empleo o a las políticas en salud (INCMSZ, 2013). Pese a esto, no existe una definición única del concepto ni una completa diferenciación con otros conceptos similares, siendo frecuentemente mal utilizada.

Urzúa y Caqueo-Urizar (2012), en su artículo “Calidad de vida: una revisión teórica del concepto”, señala que dicho concepto siempre ha sido vinculado con variables psicológicas que incorporan en sí mismo el concepto de bienestar y que aún muchos investigadores no diferencian claramente en sus estudios cada concepto o lo utilicen de manera indistinta. Ambos citan a Meeberg (1993) quien indica que se utilizan otros términos para designar a la calidad de vida como satisfacción con la vida, bienestar subjetivo, bienestar, auto-reporte en salud, estado de salud, salud mental, felicidad, ajuste, estado funcional y valores vitales, mientras que otros autores agregan más confusión al no definir lo que para ellos representa calidad de vida en sus investigaciones.

Fernández-Ballesteros (1998) define que las principales líneas de discusión sobre el concepto de calidad de vida son dos: (a) Si esta corresponde netamente a una percepción subjetiva de la persona sobre ciertas condiciones, o bien incluye además de esta percepción, la evaluación objetiva de las mismas condiciones y (b) si se refiere a un concepto ideográfico (el propio sujeto establece sus componentes) o bien nomotético (estableciendo un criterio general para todos los sujetos).

Otros problemas recurrentes en la investigación de calidad de vida son:

- La inexistencia de un modelo o marco teórico que avale la definición presentada
- La no diferenciación con factores que influyen la calidad de vida
- La validez de los instrumentos de medición utilizados

El concepto de calidad de vida representa un término multidimensional de las políticas sociales que significa tener buenas condiciones de vida ‘objetivas’ y un alto grado de bienestar ‘subjetivo’, también incluye la satisfacción colectiva de necesidades a través

de políticas sociales en adición a la satisfacción individual de necesidades (Palomba, 2002). Esta autora reitera el carácter objetivista y subjetivista del concepto, toda vez que los elementos que influyen en la excelsa condición de vida del individuo son diversos en características y significancia, por tanto, multidimensional.

Ardila (2003) formaliza una nueva definición de carácter integradora donde señala que calidad de vida es un estado de satisfacción general, derivado de la realización de las potencialidades de la persona que posee aspectos subjetivos y objetivos. Es una sensación subjetiva de bienestar físico, psicológico y social que incluye la intimidad, la expresión emocional, la seguridad percibida, la productividad personal y la salud percibida. Como aspectos objetivos considera el bienestar material, las relaciones armónicas con el ambiente físico y social con la comunidad, y la salud objetivamente percibida.

El concepto calidad de vida que es utilizado bajo un contexto de la salud, la justicia y la ética, ha sido estructurado teniendo en cuenta los aportes de varias disciplinas de contenidos muy diferentes. Según González (2002), este concepto inicia con el carácter valorativo, contextual e histórico al que debe atenderse, así como su fidelidad a los valores humanos que constituyan expresión de progreso social y respeto a una individualidad en que se armonicen necesidades individuales y sociales. La utilización de este concepto permite valorar las condiciones de vida de las personas y comunidades para poder estimar el grado de progreso alcanzado y seleccionar las formas de interacción humana y con el medio ambiente más adecuado para acceder a una existencia digna, saludable, libre, con equidad, moral y feliz.

Vicente-Galindo (2003) realiza en su trabajo doctoral una revisión bibliográfica exhaustiva tomando en consideración las publicaciones más destacadas de los años noventa. La revisión destaca que la mayoría de los autores definen la calidad de vida conceptualmente y, salvo la literatura teórica, no existe una definición explícita. En la investigación señala que Haas (1999) identificó cinco atributos para la definición y usos:

- Es una evaluación de las circunstancias de la vida de un individuo
- Tiene carácter multidimensional

- Está en relación con la escala de valores del individuo y es dinámica
- Comprende indicadores objetivos y subjetivos
- Es mucho más fiable si es evaluada mediante indicadores subjetivos

En consecuencia, [Vicente-Galindo \(2003\)](#) propone utilizar la definición de calidad de vida propuesta por [Haas \(1999\)](#) donde señala que calidad de vida es una evaluación subjetiva del bienestar físico, psicológico y social.

Para [Veenhoven \(2001\)](#), la calidad de vida es un concepto paraguas, es decir, a una conceptualización que incorpora un amplio conjunto de definiciones según puntos de vistas y enfoques que pertenece a una familia de conceptos que hacen referencia a los objetivos más relevante de la vida los cuales son alcanzados para el propio bien individual y no como medio para obtener otros objetivos.

Sin embargo, [Moreno y Ximénez \(1996\)](#) presenta una revisión bibliográfica para lograr una aproximación al concepto de calidad de vida desde una perspectiva psicológica con el objetivo de identificar los principales instrumentos para la evaluación de este concepto. Este señala la naturaleza compleja de este concepto lo cual hace difícil su propia definición conceptual. Desde el campo de la psicología, la calidad de vida es referida como la respuesta individual y social de las personas ante el conjunto de situaciones reales de la vida diaria centrándose en la percepción y estimación de ese bienestar, en el análisis de los procesos que conducen a esa satisfacción y en los elementos integrantes de la misma.

Existen amplias discusiones en la literatura sobre la calidad de vida, y si la misma es objetiva o subjetiva. [Cummins \(2000\)](#) concluye que ambos tipos de medidas han demostrado ser indicadores útiles, y aun cuando son independientes, el grado de dependencia aumenta cuando las condiciones objetivas de vida son más pobres. En este sentido, [Wrosch y Scheier \(2003\)](#) y [Huppert y Whittington \(2003\)](#) han investigado la calidad de vida relacionándola con indicadores objetivos (condiciones de vida, estado de salud) e indicadores subjetivos (satisfacción con la vida, bienestar). Ambos asumen que estos indicadores explican parte de la varianza de la calidad de vida

reportada por las personas, y que éstas pueden valorar más una cosa que otra, dependiendo de la significación que le otorguen y los estándares de comparación.

El proceso de medición y evaluación de la calidad de vida es igual de difícil que su aproximación conceptual dada la naturaleza múltiple, compleja y la bipolaridad objetiva-subjetiva que impone a sus medidas múltiples criterios en función de la opción efectuada (Moreno & Ximenez, 1996). Los dos enfoques generalmente utilizados para la medición de esta condición suponen que el primero está centrado en las variables objetivas y externas al sujeto, mientras que el segundo orientado al análisis de los aspectos subjetivos de la calidad de vida.

Ambos enfoques suponen, necesitan y requieren una revisión rigurosa para cuantificar la realidad tanto objetiva como subjetiva, mientras que igualmente requieren de una elaboración teórica y empírica de modelos bajo metodologías estrictas y rigurosas que se complementan mutuamente.

El primero de los enfoques, fundamentalmente sociológico y económico, ha tratado de establecer el bienestar social de una población a partir de los datos cuantitativos y objetivos, utilizando preferentemente la metodología de los indicadores sociales. Este enfoque no deja de lado los aspectos cualitativos de la vida real, sino que trata de elaborar modelos teóricos correspondientes que sirvan para captar los hechos reales. Bauer (1966), Bloom (1978) y Blanco-Abarca y Chacón (1985) citado por Moreno y Ximénez (1996) señala que el uso de los indicadores sociales en áreas como salud, educación, bienestar social y seguridad ciudadana permitiría establecer la calidad de vida de una población en un momento dado y la incidencia en estas de programas sociales y políticos, o bien efectuar comparaciones entre diferentes estratos de una misma población.

Este primer enfoque tiene ventajas y limitaciones, ya que los indicadores sociales expresan unos datos y unos hechos vinculados al bienestar social de una población, pero no reflejan necesariamente el grado de satisfacción y felicidad que esa población puede gozar. Los indicadores materiales y objetivos pueden ser elementos necesarios, pero no suficientes para dar cuenta del bienestar subjetivo de la población en su conjunto y de los individuos, pues son limitados e insuficientes.

El segundo de los enfoques propone la elaboración de metodologías e instrumentos que permitan la apreciación subjetiva del nivel de satisfacción global o de áreas parciales como la satisfacción política o familiar. En este sentido, [Moreno y Ximénez \(1996\)](#) continúa señalando que el procedimiento implica algunos riesgos donde se reporte una baja calidad de vida de carácter objetivo (estabilidad económica o desempeño laboral) pero un aceptable nivel de satisfacción personal. A pesar de existir una tendencia de manifestar una mayor satisfacción con aspectos de carácter general, con relación a los aspectos concretos o específicos, se manifiestan altos niveles de insatisfacción individual.

En esta línea de ideas, [Albouy et al. \(2010\)](#) examina la condición de bienestar que experimentan las personas o la satisfacción general en sus vidas en un momento dado. En los estudios de felicidad, el bienestar se logra medir cuando los participantes responden su nivel de satisfacción con la siguiente pregunta: “En términos generales, ¿se encuentra Usted muy satisfecho, satisfecho, medianamente satisfecho, no muy satisfecho, para nada satisfecho con su vida cotidiana?”

Estos autores señalan que otro acercamiento para medir el bienestar es la medición de la calidad de vida de una persona, lo que resulta de una evaluación de la situación de un individuo en términos de las diferentes dimensiones que influyen en su condición de vida (en términos materiales, en términos de salud, condiciones de vivienda, seguridad, etc.) y restarle la posibilidad de tener una calidad de vida satisfactoria.

Para medir bienestar de la población se utiliza una lista de preguntas relacionadas con aspectos muy puntuales como las dificultades económicas, pagos atrasados, restricciones en el consumo y posibles dificultados en el hogar. Sin embargo, cuando se examina la calidad de vida, se debe intentar medir la situación de las personas en términos de condiciones de trabajo, posibilidades de descanso y diversión, su nivel de integración social, vulnerabilidad económica o física, entre otras. Estas dimensiones se presentan en la Tabla 1.

Tabla 1. Dimensiones de la calidad de vida

Dimensiones		
Educación	Condiciones materiales de vida	Contacto con otros
Salud	Restricciones financieras	Inseguridad económica
Condiciones de trabajo	Participación en la vida pública	Inseguridad física

Fuente: [Albouy, Godefroy & Lollivier \(2010\)](#)

El interés de continuar investigando y comparar las condiciones de vida de los individuos en los países ha llevado a la proliferación de una amplia cantidad de indicadores sobre calidad de vida. Este es un enfoque a nivel macro donde se evalúan las condiciones generales de las personas de una región con respecto a otra y, de forma comparativa, los mecanismos adoptados por países para elevar dichas condiciones. Esta premisa es señalada por [Rojas \(2014\)](#) quien indica que algunos estudios orientan el énfasis de las investigaciones en la comparación entre países y la necesidad de políticas intervencionistas. Este esfuerzo ha llevado a la selección y agrupación de una amplia lista de variables, donde académicos han desarrollado una gran cantidad de índices que implícitamente conceptualizan la calidad de vida sin esbozar las razones fundamentales de una definición específica.

Todo esfuerzo a nivel macro parte por metodologías más individualizadas examinan objetiva y subjetivamente la calidad de vida de los individuos. [Velarde-Jurado y Ávila-Figueroa \(2002\)](#) indican que las mediciones de calidad de vida pueden estar basadas en varios sistemas. Uno de ellos son las encuestas directas que generan referencias iniciales, diagnósticas y cambios a través del tiempo, pero que sus resultados se basan en mediciones blandas con una carga variable de subjetividad y se requiere de métodos de evaluación válidos, reproducibles y confiables. Se cuentan con métodos objetivos que mediante cuestionarios generan escalas e índices que permiten medir las dimensiones que conforman el estado de salud de un individuo.

De igual manera, señalan que existen dos tipos de instrumentos para medir la calidad de vida: genéricos y específicos. Los primeros son útiles para comparar diferentes poblaciones y padecimientos, pero tienen el riesgo de ser poco sensibles a los cambios clínicos, por lo cual su finalidad es meramente descriptiva. Los instrumentos específicos se basan en las características especiales de una determinada situación,

sobre todo para evaluar cambios físicos y efectos a través del tiempo, y permitiendo una mayor capacidad de discriminación y predicción.

Los autores señalan que “aunque es muy atractivo elaborar un instrumento nuevo, es importante recordar que la elaboración y validación consumen mucho tiempo y no se tiene la certeza de que será útil. Por lo que se recomienda usar instrumentos ya existentes, muchos de ellos son adecuados y pueden aplicarse a los propósitos del estudio” (p.449). Bajo esta perspectiva, los autores realizaron una revisión de los instrumentos publicados y que son utilizados para medir la calidad de vida en el cual describen, entre genéricos y específicos, su objetivo, validez, confiabilidad, dimensión y función.

En general, [Velarde-Jurado y Ávila-Figueroa \(2002\)](#) sugieren una amplia gama de enfoques las cuales están centradas en la perspectiva de condición satisfactoria del propio individuo. [Veenhoven \(2006\)](#)¹ señala que la calidad de vida es un concepto utilizado para referirse al valor de la vida de una persona o que tan buena es esa vida. Pero que, al no poder obligar a un uso estricto de este concepto, se propone esclarecer su uso por medio de un marco de referencia o matriz que ubique y describa los diferentes significados concernientes al término calidad de vida, lo que denomina “El marco de análisis de las Cuatro Calidades de Vida”².

Veenhoven establecía que no es fructífero tratar de hacer una distinción entre indicadores objetivos y subjetivos para el estudio de la calidad de vida, como realizan la mayoría de los investigadores. La mayor diferencia en el tema yace en conocer las cualidades objetivas y subjetivas, y el tipo de disciplina desde la cual se examina la calidad de vida. Con esto propone una matriz de doble entradas (*fourfold matrix*) la cual ofrece mayor ventaja en dicha distinción y mayor sentido teórico.

Las cuatro calidades de vida no pueden ser agregadas y que, en consecuencia, es conveniente trabajar dentro de una matriz. Esto simplifica el trabajo de tratar de construir un único indicador que resuma las condiciones de calidad de vida y permitir

¹ Este artículo es una versión reducida del artículo original publicado en *Journal of Happiness Studies* (2000), 1(1), 1-39, *The Four Qualities of Life: Ordering Concepts and Measures of the Good Life*

² *The Four-Qualities-of-Life framework of analysis*, concepto que Veenhoven incorpora por primera vez en su publicación de 2000.

ubicarlo dentro de uno de los cuatro indicadores que representan la percepción de una persona. La matriz de Veenhoven tiene el propósito de proveer una evaluación acerca de la calidad de vida de las personas en lugar de proveer explicaciones. Esta conceptualización constituye una precondition para cualquier estudio sobre los factores que la explican.

La matriz de Veenhoven sugiere que la evaluación de la calidad de vida puede basarse en los resultados de ella misma, donde estos resultados se refieren a la buena vida y que son los objetivos más relevantes para una persona en el sentido que son alcanzados por su propio interés y no como un mecanismo para alcanzar otros objetivos. Existen dos tipos evaluaciones de los resultados de calidad de vida en la matriz de Veenhoven: primero, una evaluación de la calidad externa de la vida de una persona, la cual se basa en la contribución de este individuo a otros miembros de la sociedad. Segundo, una evaluación de la calidad interna de vida de una persona, la cual se basa en su satisfacción o apreciación de su vida, y cuál es la calidad de su vida en la medida que la experimenta.

La evaluación de la calidad de vida también se basa en la presencia de aquellas condiciones (posibilidades) que son relevantes sobre el estado de una persona y cuáles están dentro de la esfera de las políticas públicas. En este sentido, las posibilidades en la vida se refieren a las oportunidades que una persona tiene para alcanzar una mejor condición, siendo estas posibilidades los medios para lograrla. Las posibilidades de una mejor vida son clasificadas en dos tipos: externas e internas. La primera se refiere a las condiciones del entorno de una persona (seguridad, instituciones políticas, etc.) y las segundas son condiciones de la propia persona (salud física y mental, educación, etc.)

La matriz de cuatro aspectos de la Calidad de Vida, que se presenta en la Tabla 2, distingue cuatro combinaciones derivadas de las dos dicotomías: posibilidades y resultados de vida, las cuales son presentadas de manera vertical, mientras que las diferencias entre calidades externas e internas se muestran horizontalmente.

Tabla 2. Cuatro aspectos de la Calidad de Vida

	Calidad externa	Calidad interna
Posibilidades de vida	Habitabilidad al entorno	Capacidad de vida
Resultados de vida	Utilidad de la vida	Satisfacción con la vida

Fuente: [Veenhoven \(2006\)](#)

Los cuatro escenarios de calidades de vida según las diferentes combinaciones permitidas son:

- Las posibilidades de calidad externa de vida (*Livability of environment*), denominada habitabilidad de una persona a su entorno
- Las posibilidades de calidad interna de vida (*Life-ability of a person*), también llamada capacidad de vida de una persona
- Los resultados de calidad externa de vida (*Utility of a person's life*), conocido como utilidad de la vida de una persona
- Los resultados de calidad interna de vida (*Person's satisfaction with its life*), denominado satisfacción de una persona con su propia vida

Es importante establecer que cada cuadrante constituye una evaluación de la calidad de vida de una persona. Estas evaluaciones pueden estar correlacionadas, dado que se espera que las posibilidades estén altamente correlacionadas con los resultados ya que la definición de las posibilidades está basada en las relaciones positivas previas hacia los resultados.

1.1.2. Calidad de vida relacionada con la salud

El concepto de Calidad de Vida Relacionada con la Salud (CVRS) se incorpora como una faceta del estado de vida de un individuo en la medida que se examina su bienestar como paciente médico. [Urzúa \(2010\)](#) explica que la CVRS, como una medida del estado de salud de las personas, ha sido uno de los conceptos de mayor utilización en el área de la salud y que su estudio ha sido distinto al de calidad de vida. La CVRS es usada indistintamente como estado de salud estado funcional y surge por el interés de

identificar estrategias efectivas para mejorar la condición del paciente como resultado de intervenciones médicas. Este concepto es utilizado en el campo de la medicina en su interés por evaluar la calidad de los cambios como resultado de intervenciones médicas, porque debe limitarse a la experiencia que el paciente tiene de su enfermedad, porque caracterizan la experiencia del paciente resultado de los cuidados médicos o para establecer el impacto de la enfermedad en la vida diaria.

Esta es una premisa sumamente importante dado que otros investigadores sugieren que la CVRS debería enfocarse en características tales como ingreso, libertad y calidad del medioambiente ya que estos indicadores están fuera del contexto de los resultados médicos, y podrían ser vinculados con mayor facilidad a estos que otros indicadores propuestos por las ciencias sociales tales como la felicidad, la satisfacción con la vida o el bienestar subjetivo.

La investigación concluye en definir “la calidad de vida relacionada con la salud como el nivel de bienestar derivado de la evaluación que la persona realiza en diversos dominios de su vida, considerando el impacto que en éstos tiene su estado de salud”

Urzúa señala que la utilización de un claro concepto de CVRS puede ser beneficioso en la práctica clínica, en estudios de eficacia, efectividad, riesgo o como indicador del cuidado, por lo cual establece para cada tipo de estudio un conjunto de preguntas de investigación primaria, tal como muestra en la Tabla 3.

La ventaja del uso de indicadores centrados en la calidad de vida a nivel personal es que, al focalizar su mirada en cada persona, enfatiza la particularidad de cada individuo para responder a su enfermedad y al tratamiento asignado, siendo para esto la medición centrada en el paciente la más adecuada.

La complejidad de estudiar calidad de vida y CVRS obliga a trabajar en una estrecha frontera entre ambos contextos. Desde una perspectiva psicológica, la medida de la calidad de vida tiene como objetivo principal medir los efectos de las intervenciones en el cuidado de la salud entendida de forma global y positiva, evaluar la calidad de tal cuidado, estimar las necesidades de la población, mejorar las decisiones clínicas, y estudiar las causas y consecuencias del estatus de la salud. Pero, todos los que participan en el sistema del cuidado de la salud tienen objetivos comunes,

fundamentalmente consistentes en la extensión de la duración de la vida y la mejora de la calidad de vida (Moreno & Ximenez, 1996).

Tabla 3. Preguntas de investigación desde la perspectiva de la calidad de vida relacionadas con la salud

Tipo de Estudio	Pregunta de Investigación Primaria	Pregunta de Investigación de CVRS
Eficacia	¿Está el grupo asignado a tratamiento asociado con resultados clínicos bajo circunstancias ideales cuando distribuyeron los pacientes seleccionados?	¿Los grupos de tratamiento difieren en su CVRS?
Efectividad	En la práctica, ¿fue la aplicación del tratamiento bajo condiciones ordinarias par un paciente típico?	¿Tratamientos mejoran la CVRS?
Establecimiento de riesgo	¿Qué variables están relacionadas con resultados?	¿La CVRS predice los resultados en los pacientes? ¿Qué variables predicen la CVRS?
Calidad del cuidado	¿Cómo la calidad del cuidado fue comparada con estándares? ¿Cómo la calidad del cuidado varía por grupos?	¿Cómo la CVRS varía por grupos?

Fuente: [Urzúa \(2010\)](#)

[Ebrahim \(1995\)](#) establece que los propósitos específicos de las mediciones en CVRS son monitorear la salud de la población, evaluar el efecto de las políticas sociales y de salud, focalizar los recursos con relación a las necesidades, diagnóstico de la naturaleza, severidad y pronóstico de una enfermedad y evaluar los efectos de un tratamiento.

La medición de la calidad de vida se ha convertido en un componente importante para entender los resultados de la condición de salud de una persona. Para aquellos casos donde la población sufre de enfermedades crónicas, la medición de la calidad de vida proporciona un mecanismo significativo para determinar el impacto de los diferentes medios adoptados en los cuidados médicos cuando la cura no es posible. Ello implica sustentar los enfoques basado en la funcionalidad subjetiva en comparación a las expectativas personales y su definición en términos de las experiencias subjetivas,

estados y percepciones. Por tanto, su naturaleza es propia al individuo, pero significativa y entendible para la mayoría de las personas (Burckhart & Anderson, 2003).

Las premisas presentadas por Burckhardt y Anderson hacen suponer que existen elementos genéricos para todas las personas que son condicionantes de la calidad de vida de un individuo y que, a su vez, estos pueden influir en la CVRS de este. En consecuencia, se establece la necesidad de desarrollar sistemas para medirla con el propósito de definir su nivel influencia exógena, entendiéndose como la influencia de factores externos como los psicológicos, sociales, económicos, entre otros, y los resultados de carácter clínico propios de cada paciente.

Las mediciones en CVRS tienen sus ventajas ya que su uso ha permitido describir una condición o estado de salud, señalar cambios en el funcionamiento del paciente, proveer un pronóstico o establecer normas de referencias. Esto es señalado por Testa (2000b) quien destaca la existencia de una brecha en la interpretación de los resultados y efectividad de los tratamientos médicos como una forma para evaluar la calidad de vida. En este sentido, expresa que puede haber una disfunción al examinar los resultados estrictamente médicos con los relativos a la CVRS.

Testa (2000b) identifica alguna de las razones que generan un serio desafío en la medición e interpretación de la CVRS: la naturaleza subjetiva del constructo calidad de vida, la forma indirecta de cómo es evaluado, las múltiples fuentes para la medición de errores, la heterogeneidad de las propiedades estocásticas de los cambios longitudinales con respecto al rango de la escala, la compleja asociación entre los diferentes resultados y la falta de claridad de los objetivos que se persiguen definidos en términos de cambios en la calidad de vida.

Testa (2000a) logra identificar tres categorías de determinantes que influyen en la CVRS: predictores médicos, predictores actitudinales y predictores demográficos. Al revisar los resultados de las investigaciones de calidad de vida, es importante que se identifiquen las limitaciones y fortalezas de las mediciones y análisis utilizados dentro del contexto de los objetivos iniciales de la investigación.

Las mediciones de CVRS deben alcanzar los estándares de desempeño correspondiente al propósito del análisis. Esto supone que los indicadores u objetivos perseguidos para medir esta condición serán propios del análisis que se desea realizar y que la escala de desempeño de calidad de vida necesitará una validación pertinente a estos objetivos. Testa (2000a) establece los requerimientos de validación, los cuales se presentan en la Tabla 4.

Tabla 4. *Requerimientos de validación de la escala de desempeño de la calidad de vida según propósito del análisis*

Análisis	Objetivo de Investigación Primaria	Análisis de validación de la escala de desempeño
Predictivo	Mejorar la adherencia	Alta correlación positiva entre los determinantes de adherencia
Discriminante	Evaluación de las necesidades de utilización de los servicios de salud	Alto radio de varianza (alta/baja utilización) al comparar entre grupos y dentro del grupo
Evaluativo	Distinción entre dos agentes (sujetos)	Alta capacidad de respuesta y sensibilidad sobre los rangos de escala operativo

Fuente: Testa (2000a)

Un análisis evaluativo debe emplear escalas de calidad de vida que respondan a los cambios y diferencias consecuentes al constructo de calidad de vida que se haya definido, y a su vez posea una escala que permita detectar los cambios y las diferencias de manera significativa. Los grados de diferencias o cambios que puedan ser considerados como significativos dependen de la hipótesis de investigación establecida, los tipos de análisis cuantitativos correspondientes y la intención o perspectiva del investigador (Ibíd).

En este punto se destaca la importancia de validar los indicadores y los resultados obtenidos, así como el orientar estos indicadores a los propios objetivos del análisis. Pero la literatura señala que la utilización de la CVRS como un indicador de evaluación proporciona de información de alto valor y relevancia para el clínico o investigador médico, dada la naturaleza de sus mediciones de carácter clínico y fisiológicas tradicionales. Sin embargo, se señala que esta información es de limitado interés para

el paciente, dado que no describe la capacidad funcional y el bienestar del individuo, lo cual es su principal preocupación. Esta diferencia en intereses puede llevar a que pacientes con similares criterios clínicos presenten respuestas diferentes.

La deficiencia entre los intereses del paciente y el investigador demuestra una debilidad en la evaluación por indicadores netamente clínicos, lo cual puede influir en la evaluación general de la calidad de vida y sus aspectos relacionados a la salud. En la Tabla 5 se describen los supuestos planteados por [Wu \(2000\)](#) los cuales son requeridos para una medición de carácter universal y los problemas enfrentados que impactan en la CVRS.

Tabla 5. *Supuestos requeridos para la medición de la calidad de vida relacionada con la salud y problemas identificados*

	Enfermedad / Tratamiento	Síntoma	Calidad de vida relacionada con la salud
Supuestos	Enfermedad es uniforme	Experiencia de síntomas es uniforme	Impacto en el funcionamiento o bienestar es uniforme
	Efectos del tratamiento son uniformes	Reporte de síntomas es uniforme	
Problemas	Procesos de enfermedad son diferentes	Experiencia difiere	Impacto en el funcionamiento o bienestar difiere
	Fisiopatología difiere	Reportes difieren	Definiciones de calidad de vida relacionada con la salud difiere
	Respuesta al tratamiento difiere		
	Efectos adversos difieren		

Fuente: [Wu \(2000\)](#)

Sobre estos supuestos y problemas planteados, [Urzúa \(2010\)](#) señala que la evaluación de la CVRS debe realizarse a través de:

- Instrumentos genéricos
- Instrumentos específicos (como para enfermedades particulares o grupos de pacientes definidos)
- Áreas o dimensiones de funcionamiento (dolor)

A pesar de que esta evaluación, específicamente la construida por enfermedades específicas, pudiera tener ciertas ventajas como la comprensión particular del impacto de una enfermedad sobre determinados dominios de la vida de una persona en particular, aportando con esto herramientas útiles al clínico y a los terapeutas para intervenciones con miras a aumentar la calidad de vida, se sugiere que centrarse específicamente en la enfermedad.

Y con respecto a las dimensiones que componen la CVRS, tampoco existe un acuerdo en la literatura existente pero sí pareciera haber consenso en la inclusión de dominios específicos a cada enfermedad en el caso de las evaluaciones de determinadas patologías dada las particularidades de cada una de ellas (Ibíd).

[Tuesca-Molina \(2005\)](#) señala que el concepto de salud está fundamentado en un marco biopsicológico, socioeconómico y cultural, teniendo en cuenta los valores positivos y negativos que afectan nuestra vida, nuestra función social y nuestra percepción; por tanto, la redefinición del concepto de salud es de naturaleza dinámica y multidimensional. De ahí deriva la importancia de medir la calidad de vida y su relación con la salud.

Señala que el uso de medidas de CVRS evalúa una amplia gama de dimensiones y han sido diseñadas para ser aplicadas a todo tipo de pacientes con diferentes patologías y en el ámbito poblacional a fin de comparar y evaluar las diferentes variaciones en el estado de salud.

La importancia de esta medición permite destacar el estado funcional del paciente que refleje su salud física, mental y social, tomando en cuenta que la variable de expresión será de tipo subjetiva y puede obtener un valor cuantitativo que represente el estado de salud. [Tuesca](#) identifica (Tabla 6) instrumentos que son ampliamente utilizados para medir la CVRS y describe su función general.

[Ramírez \(2007\)](#) presenta los resultados de una selección y revisión de artículos científicos biomédicos publicados entre 1991 a 2006 el cual se centra en listar y describir instrumentos genéricos y específicos que son utilizados para evaluar la CVRS, su validez, sus propiedades psicométricas y los resultados sobre salud registrados en las bases de datos del Centro Latinoamericano y del Caribe de

Información en Ciencias de la Salud (LILACS) y Medline perteneciente a la National Library of Medicine (Biblioteca Nacional de Medicina) de los Estados Unidos, tal como se muestra en las Tablas 7 y 8.

Tabla 6. *Instrumentos utilizados para la medición de la calidad de vida relacionada con la salud*

Instrumentos Genéricos	Descripción general del instrumento
Sickness Impact Profile (Perfil de las Consecuencias de la Enfermedad)	Estos aportan información descriptiva, predictiva o evaluativa de un individuo o grupo de individuos a manera de puntuación que resume en cada dimensión la puntuación obtenida y al aplicar diferentes estrategias para interpretar, en resumen, a fin de caracterizar el perfil de salud global del sujeto.
Nottingham Health Profile (Perfil de Salud de Nottingham)	
Health Survey SF-36 (Cuestionario de Salud SF-36)	
EuroQol	Estos instrumentos producen valores que reflejan la preferencia de los individuos por los estados de salud, compuestos por distintos ítems de diferentes dimensiones que permiten estimar el cálculo de los años de vida ajustados por calidad. Reflejan las medidas de utilidad basadas en las preferencias o utilidades que los individuos asignan a los diferentes estados de salud.
Quality of Well Being Scale	
Matriz de Rosser y Kind	
Health Utility Index	
15-D	
Instrumento Específico	Descripción general del instrumento
Índice de Calidad de Vida de Spitzer (ICVS)	Instrumento simple, corto y de fácil administración que incluye diferentes dimensiones relacionadas con la calidad de vida y permite discriminar los cambios entre grupos de pacientes. Consta de 5 componente, con tres alternativas de respuesta cada uno, evaluando: actividad física, autonomía, salud percibida, apoyo social y familiar, y control y aceptación de las propias limitaciones y perspectivas de futuro

Fuente: [Tuesca-Molina \(2005\)](#)

Algunos instrumentos fueron excluidos por Ramírez por demostrar criterios de inconsistencia propios del análisis. En la investigación, Ramírez logra identificar y examinar un total de 67 estudios que utilizan diferentes instrumentos o cuestionarios, de los cuales 33 son de carácter genéricos y 34 de carácter específico.

Tabla 7. *Cuestionarios genéricos para evaluar la calidad de vida relacionada con la salud*

Cuestionario	Dominio	Función	Objetivo
The Nottingham Health Profile: NHP Traducción al español	Funcionalidad, físico, emocional, social, nivel actividad, percepción deterioro	Discriminación	Validación
The Dartmouth - COOP Traducción al español	Movilidad, físico, social, emocional, apoyo social, dolor	Discriminación Evaluación	Validación Confiabilidad
The Sickness Impact Profile: SIP	Movilidad, físico, emocional, social	Predicción Evaluación	Validación Confiabilidad
The Functional Status Index: FSI	Funcionalidad, físico, social emocional	Discriminación Evaluación	Validación Confiabilidad
The Kamofsky Performance Status (KPS) Scale	Funcionalidad, dolor	Discriminación Predicción	Validación Confiabilidad Descripción
Quality of life Index: QL - Index	Funcionalidad, físico, emocional, social, vitalidad	Discriminación	Validación Confiabilidad Descripción
The Dartmouth - COOP	Movilidad, físico, social, emocional, apoyo social, dolor	Discriminación Evaluación	Validación Confiabilidad
The Health Assessment Questionnaire: HAQ	Funcionalidad, físico, social, mental, dolor, auto-cuidado	Discriminación Predicción	Validación Confiabilidad Descripción
The Sickness Impact Profile: SIP 68 Short Generic Version	Funcionalidad, físico, emocional, social, autonomía, satisfacción	Discriminación Predicción Evaluación	Validación Confiabilidad
The McMaster Health Index Questionnaire	Movilidad, físico, emocional, social	Discriminación Evaluación	Validación Confiabilidad
The Duke Health Profile: DUKE	Funcionalidad, físico, social, emocional,	Discriminación Predicción	Validación Confiabilidad

	autoestima, dolor, apoyo social		
The general Health Questionnaire: HRQOL	Funcionalidad, social, emocional, ansiedad	Discriminación	Validación Confiabilidad Descripción
Cuestionario Criterio de Calidad de Vida: CCV	Funcionalidad, físico, social, emocional, vómito, apetito, sueño, fatiga, dolor	Discriminación	Validación Confiabilidad
The European Research and Treatment Quality Life-Questionnaire: EORTC	Funcionalidad, físico, social, emocional, fatiga, dolor	Discriminación	Validación Confiabilidad
QLQ - C36	Funcionalidad, físico, social, emocional, vitalidad, mental, dolor	Discriminación	Validación Confiabilidad
The MOS-Short-form General Health Survey	Funcionalidad, físico, social, emocional, mental, dolor	Discriminación	Validación Confiabilidad
The Medical Outcomes Study Short - Form Health Survey: MOS	Funcionalidad, físico, social, emocional, dolor	Discriminación	Validación Confiabilidad Descripción
The McGill Pain Questionnaire: MQOL	Funcionalidad, físico, social, emocional, espiritual, apoyo, social	Discriminación	Validación Confiabilidad Descripción
The Quality of Well-Being Scale: QWB	Funcionalidad, físico, social, emocional, nivel actividad, deterioro	Discriminación Predicción	Validación Confiabilidad Descripción
The Modified Health Assessment Questionnaire: M-HAQ	Funcionalidad, físico, social, emocional, ansiedad	Discriminación Predicción	Validación Confiabilidad
The Multidimensional Index of Quality of Life: MIQL	Funcionalidad, físico, social, emocional, cognición	Evaluación	Validación Confiabilidad

The RAND 36 - Item Health Survey Questionnaire: RAND-SF36	Funcionalidad, físico, social, emocional, dolor, energía, fatiga	Discriminación Predicción	Validación Confiabilidad
The Perceived Quality of Life Scale: PQoL	Funcionalidad, físico, social, emocional, dolor, fatiga, depresión	Discriminación Predicción	Validación Confiabilidad
The European Research Questionnaire Quality of Life: EUROQOL 5D	Funcionalidad, físico, autocuidado, ansiedad, limitación	Discriminación	Validación Confiabilidad
Encuesta: SF-36	Funcionalidad, físico, social, emocional, dolor	Discriminación Evaluación	Validación Confiabilidad
The General Health Perception: GHP	Funcionalidad, físico, social, emocional, vitalidad, mental, dolor	Predicción	Validación Confiabilidad
The Wisconsin Quality of Life Index Canadian version: CaW-QLI	Funcionalidad, físico, social, emocional, satisfacción, apoyo social	Discriminación Predicción	Validación, Confiabilidad
The Medical Outcomes Study Short-form Health	Funcionalidad, físico, social, emocional, dolor	Discriminación Predicción	Validación Confiabilidad
Survey: MOS 6A The QL-Index LASA Scale	Funcionalidad, físico, social, emocional, satisfacción, apoyo social	Discriminación	Validación Confiabilidad
The World Health Organization Quality of Life: WHOQOL - 100	Funcionalidad, físico, social, emocional, fatiga, dolor	Discriminación	Validación Confiabilidad
The World Health Organization Quality of Life: WHOQOL Instrument	Funcionalidad, físico, social, emocional, autonomía	Discriminación	Validación Confiabilidad
The Health and Activities Limitations Index: HALEX	Funcionalidad, físico, social, emocional	Discriminación Evaluación	Validación Confiabilidad
The Index of Well-Being: IWB	Funcionalidad, físico, social, emocional	Discriminación Evaluación	Validación Confiabilidad

Fuente: [Ramírez \(2007\)](#)

Tabla 8. *Cuestionarios específicos para evaluar la calidad de vida relacionada con la salud*

Cuestionario	Dominio	Función	Objetivos
The Chronic Venous Insufficiency Questionnaire: CMQ	Funcionalidad, física, dolor, limitación, difusión	Discriminación Predicción	Validación Confiabilidad
The Diabetes Quality of Life Measure: DQOL	Satisfacción, Metas, estrés, depresión	Discriminación Predicción	Validación
The Epilepsy Surgery Inventory - 55: ESI-55	Deterioro Neurológico, Cognición, limitación física	Discriminación Evaluación	Validación Confiabilidad
The Seattle Questionnaire: SAQ	Funcionalidad, física, social, emocional, dolor	Discriminación Predicción	Validación Confiabilidad
The Mild Hypertension Vital Signs Quality of Life Questionnaire: VSQOLQ	Dolor, limitación	Discriminación Predicción Evaluación	Validación Confiabilidad
The Diabetes Specific Quality of Life Scale for patients With Type Diabetes: DSQOLS	Satisfacción, actividad física, estrés, depresión	Discriminación Predicción	Validación Confiabilidad
Quality of Life Questionnaire for Graves Ophthalmopathy: GO-QOL	Cambios apariencia física, disminución agudeza visual, deterioro psicológico	Discriminación Predicción	Validación Confiabilidad
The Quality of Life Questionnaire for Multiple Sclerosis: QOLQ for MS (93)	Funcionalidad física, limitación emocional, fatiga	Discriminación Predicción	Validación Confiabilidad
The Goldman Specific Activity Scale (94)	Funcionalidad, físico, dolor, limitación	Discriminación Predicción Evaluación	Validación
Quality of life Parkinson Disease Questionnaire: PDQ-39 (95)	Deterioro, Neurológico, limitación	Discriminación Predicción	Validación Confiabilidad

The Quality of Life en Epilepsy: QOLIE-89	Funcionalidad, físico, social, mental, cognición, status, post-crisis, relación de pareja	Discriminación	Validación Confiabilidad
The Quality of Life en Epilepsy: QOLIE-31	Funcionalidad, físico, social, mental, cognición, status, post-crisis, relación de pareja	Discriminación	Validación Confiabilidad
The Quality of Life en Epilepsy: QOLIE-10	Funcionalidad, físico, social, mental, cognición, status, post-crisis, relación de pareja	Discriminación Prueba repetida	Validación Confiabilidad
The Chronic Ear Survey: CES	Funcionalidad, físico, social, mental, dolor	Discriminación Evaluación	Validación Confiabilidad
The Medical Outcomes Study (MOS) Short form Health Survey: MOS-HIV	Funcionalidad, físico, social, mental, cognitivo, síntomas, fiebre, fatiga, energía, dolor, estrés	Discriminación	Validación Confiabilidad
The oral health-related quality of life questionnaire: OHRQOL	Funcionalidad, física, social, emocional	Discriminación Evaluación	Validación Confiabilidad
The Dermatology Life Quality Index	Funcionalidad, física, social, emocional	Predicción	Validación Confiabilidad
The Chronic Respiratory Disease Questionnaire: CRQ	Funcionalidad, físico, social, mental, dolor, disnea, fatiga	Discriminación Predicción Evaluación	Validación Confiabilidad
The Arthritis Impact Measurement Scales: ALMS	Funcionalidad, física, social, emocional, dolor, nivel actividad	Discriminación Predicción Evaluación	Validación Confiabilidad
The Measure Yourself Medical Outcome Profile	Funcionalidad, físico, social, emocional	Predicción	Validación Confiabilidad
The Marks Asthma Quality of Life Questionnaire: AQLQ	Funcionalidad, físico, social, emocional, síntomas, dependencia de tratamiento	Discriminación Evaluación	Validación Confiabilidad

The Inflammatory Bowel Disease Questionnaires: IBDQ	Funcionalidad, físico, social, emocional, función intestinal y sistémica	Predicción	Validación Confiabilidad
The Functional Disability Index of the Health Assessment Questionnaire: HAQFDI	Funcionalidad, física, social, emocional, dolor, nivel actividad	Discriminación , Predicción Evaluación	Validación Confiabilidad
The Prostate Cancer Specific Quality of Life Instrument: PROSQOLI	Funcionalidad, físico, social, emocional, fatiga, limitación	Predicción Evaluación	Validación Confiabilidad
The Fatigue Severity Scale: FSS	Funcionalidad, físico, social, vitalidad, fatiga	Discriminación	Validación Confiabilidad
The Self report HIV-Specific Quality of Life: HOPES	Funcionalidad, físico, social, mental, cognición	Discriminación	Validación Confiabilidad
The Health Assessment Questionnaire Disability Index: Spanish HAD-DI	Funcionalidad, física, social, emocional, dolor, nivel actividad	Discriminación Predicción Evaluación	Validación Confiabilidad
The Arthritis Impact Measurements Scales Spanish AIMS	Funcionalidad, física, social, emocional, dolor, nivel actividad	Discriminación Predicción Evaluación	Validación Confiabilidad
The Chronic Respiratory Disease Questionnaire: COPD	Funcionalidad, física, social, mental, energía, fatiga	Discriminación Evaluación	Validación Confiabilidad
The Medical Outcomes Study Health Survey: Spanish MOS-HIV	Funcionalidad, físico, social, mental, cognitivo, síntomas fiebre, fatiga, energía	Discriminación	Validación Confiabilidad
The Multidimensional Quality of Life Questionnaire for HIV: MQOL-HIV	Funcionalidad, físico, social, emocional, espiritual, apoyo social	Discriminación	Validación Confiabilidad

The HIV-AIDS - Targeted Quality of Life HAT-QoL	Funcionalidad, físico, Predicción temor control de HIV, satisfacción	Predicción	Validación Confiabilidad
The Medical Outcomes Study Health Survey: MOS - HIV	Funcionalidad, salud general, física, social, mental cognición, energía, fatiga	Predicción	Validación Confiabilidad
The Schwarts Cancer Fatigue Scale: SCFS	Funcionalidad, física, social, vitalidad, fatiga	Discriminación	Validación Confiabilidad

Fuente: [Ramírez \(2007\)](#)

La medición de la CVRS se convierte en un elemento determinante para establecer el estado de un individuo o paciente que padezca de alguna afección en una etapa inicial o crónica. Este padecimiento incide, de manera indiscutible, en la calidad de vida general de la persona debido a la influencia de otros factores que son propios del individuo o de su medio entorno. Los resultados de carácter clínico por situaciones de salud pueden llevar al individuo a tener variaciones en su calidad de vida, los cuales lo impulsan a nuevos desafíos y a enfrentarse con la necesidad de realizar cambios ante su nueva condición.

La CVRS ha sido un tema sujeto de profundos estudios para definir los elementos que influyen en la condición de los pacientes durante y al final de sus tratamientos. Durante la revisión de la literatura que examina los factores que inciden en la CVRS se han destacado diversos aspectos de carácter físicos, económicos, sociales, desarrollo personal y emocionales, entre otros, los cuales han direccionado los estudios a la necesidad de formular sistemas de medición que permitan establecer las condiciones finales de la calidad de vida de los individuos luego de atravesar por los procesos clínicos.

Sin embargo, los procesos de medición de la CVRS dependen ampliamente de la percepción del paciente al responder ante un instrumento, por lo que se reitera la existencia de un alto componente subjetivo en los resultados logrados. Aunado a ello, surge una fuerte interrogante: ¿Qué sucede cuando los criterios propios adoptados por

los pacientes varían al momento de responder a los instrumentos que miden la calidad de vida?, ¿Cómo pueden ser los resultados de un paciente cuando su percepción sobre su condición de calidad de vida difiere a lo largo del tiempo?

Estas preguntas permiten incursionar en los nuevos enfoques para medir la CVRS según el paciente, la posibilidad de cambios en su definición personal sobre su calidad de vida, así como las metodologías que se han desarrollado para evaluar las condiciones de los individuos, pero asumiendo las variaciones de percepción a lo largo de la evaluación.

Nunes (2009) examina la relación que existe entre los cambios en el estado de salud y los cambios en los estándares internos subyacentes a la evaluación de CVRS, y que al registrarse cambios en el constructo que se está evaluando, las respuestas a las mismas preguntas dejan de ser comparables. En consecuencia, la CVRS no puede ser medida estrictamente en términos de puntuaciones verdaderas y puntuaciones latentes, y las diferencias individuales en el proceso de evaluación no pueden ser vistas como fuentes de error, sino como un proceso intrínseco al propio proceso de evaluación de CVRS, Nunes (2009) citando a Rapkin y Schwartz (2004).

1.2. Estabilidad y cambio (*Response Shift*) en la calidad de vida relacionada con la salud

1.2.1. Estabilidad y cambio

El análisis del cambio a lo largo del tiempo es de suma importancia en estudios de CVRS toda vez que se examina variaciones en las respuestas de los individuos a través del período de estudio. El estudio del cambio en las personas puede realizarse bajo un enfoque de cambios cualitativos y cambios cuantitativos. (Nunes, 2009)

Los cambios de carácter cualitativo son relativamente sencillos de evaluar ya que utilizan respuestas controladas y cerradas, generalmente dicotómicas, mientras que los

cambios de carácter cuantitativo sufren de mayor complejidad producto de la comparación de valores medios de las respuestas en diferentes períodos en el tiempo. Más aún, cuando los criterios o parámetros no son estables, el proceso de medición no será valedero, ya que pueden manifestarse diferentes dimensiones a lo largo del estudio.

El estudio del cambio parte del principio de que exista estabilidad, lo cual se define como la propiedad de un cuerpo de recuperar su equilibrio inicial; constancia o permanencia en un estado o duración en el tiempo³. [Taris \(2000\)](#) hace una distinción entre las diferentes formas de presentarse la estabilidad: invarianza estructural, estabilidad normativa, estabilidad cuantitativa y consistencia intraindividual.

La invarianza estructural representa el grado de continuidad en la naturaleza de un determinado fenómeno sujeto de investigación ([Mortimer et al., 1982](#)). Bajo este criterio, un constructo se caracteriza por las mismas dimensiones y cuando persisten los patrones de relación entre sus componentes a lo largo del tiempo; es decir, no ha cambiado. [Golembiewski et al. \(1976\)](#) afirman que dos o más estructuras factoriales son iguales cuando tienen el mismo patrón de los factores de carga, pero no es necesario que la magnitud de los factores de carga se mantenga. La invarianza estructural implica que se mantenga constante la estructura factorial en cada momento del tiempo y cada situación experimental, y de mantenerse el constructo sin cambio es posible evaluar si los resultados son estables o consistentes. En consecuencia, se evalúa la existencia de estabilidad normativa y/o cuantitativa.

La estabilidad normativa se relaciona con la estabilidad observada en los rangos de orden de las preferencias de los individuos, o sus diferencias, sobre los atributos de interés a lo largo del tiempo ([Kagan, 1980](#)). Se evalúa como la correlación entre las medidas de un atributo a un período de tiempo determinado. Cuando las autocorrelaciones son elevadas entre las medidas en cuestión, se manifiesta un alto nivel de estabilidad normativa, mientras que al existir autocorrelaciones bajas se manifiestan cambios significativos en el orden de los sujetos (pacientes) durante el estudio.

³ <http://www.wordreference.com/definicion/estabilidad>

La estabilidad cuantitativa consiste a la estabilidad que existe en la magnitud o cantidad media de un fenómeno a lo largo del tiempo (Baltes & Nesselroade, 1973). Este tipo de estabilidad se evalúa a través del cambio en los valores medios asociados a los grupos en el tiempo.

La consistencia intraindividual se refiere a la consistencia en el orden de los atributos que establece el mismo individuo (Emmerich, 1968), o sea la estabilidad de las preferencias de una persona. Se evalúa calculando el coeficiente de correlación de los rangos de orden de los atributos que ha dado el individuo en dos momentos del tiempo.

Según Nunes (2009), uno de los criterios determinantes para evaluar la condición de estabilidad de los pacientes es el hecho de que el constructo sea estructuralmente invariante, y es bajo esta condición de continuidad del fenómeno que tendrá sentido evaluar los consecuentes niveles de estabilidad normativa, estabilidad cuantitativa y consistencia intraindividual. Por tanto, es imperativo comprobar la condición de invarianza estructural.

Lo anterior se deriva dado que los individuos pueden cambiar sus estándares personales, valores, prioridades o la percepción sobre qué es para ellos la CVRS luego de sufrir alternaciones en sus estados de salud. Estos cambios se pueden manifestar a lo largo de un padecimiento, los cuales son efectos propios al proceso de acomodación de la enfermedad según lo planteado por Schwartz y Sprangers (2000).

Entre los estudios pioneros relacionados al cambio se encuentra el trabajo de Golembiewski et al. (1976) que establece tres tipos diferentes de cambios los cuales pueden surgir como respuesta a una intervención. En su tipología, se consideran tres tipos de cambios diferentes: cambio Alpha, cambio Beta y cambio Gamma.

El Cambio Alpha consiste en una variación en el nivel de algún estado existencial relacionado con un dominio conceptual constante, el cual deberá ser medido con un instrumento constantemente calibrado. Este primer tipo representa el verdadero cambio y es una conceptualización de carácter convencional, toda vez que se manifiesta como una variación de una entidad considerando que la escala de medición permanece constante. El concepto es opuesto al de estabilidad cuantitativa o estabilidad de nivel, presentado por Taris (2000) citando a Baltes y Nesselroade (1973)

El Cambio Beta se manifiesta como la variación en los intervalos de medida de las dimensiones de una realidad estable debido a los cambios en la escala de medida de un sujeto de estudio. Por consiguiente, esta segunda clasificación se refiere a una recalibración en la escala de medida aplicada al encuestado. Es común que en los análisis de calidad de vida se manifieste un ajuste en las medidas utilizadas para evaluar los cambios dado que los pacientes pueden establecer diferentes estimaciones o percepciones de la realidad a través del tiempo, particularmente asociado a los procesos propios de adaptación del sujeto.

El Cambio Gamma representa cambio profundo en el constructo, ya sea por una reconceptualización de su significado o una repriorización de las dimensiones de la realidad. Esto significa que la condición preexistente es completamente diferente a la condición presente o futura por los diversos factores que influyen o caracterizan las nuevas instancias de un individuo.

Para [Golembiewski et al.\(1976\)](#), el cambio Gamma es precisamente el opuesto de invarianza estructural. En ausencia de invarianza estructural, puede haberse producido un cambio Beta o un cambio Gamma, lo que supone la necesidad de examinar los niveles de estabilidad del constructo en estudio y la existencia de cambios en las unidades de medida del instrumento o si el constructo ha cambiado y posee diferentes unidades de medidas. Por tanto, para establecer una clasificación de cambio deben conocerse dos aspectos: que el constructo estudiado sea constante a lo largo del tiempo y que la calibración del instrumento de medida sea constante en los diferentes períodos de evaluación ([Nunes, 2009](#); [Vicente-Galindo, 2003](#)). De igual forma, previo a la evaluación sobre un cambio Alpha, debe realizarse una revisión y descarte de la existencia de cambios Beta y Gamma; así como previo a definir un cambio Beta, habrá que asegurarse que no preexista un cambio Gamma que influya las variables en estudio (Ibíd).

1.2.2. *Response Shift*: un nuevo enfoque sobre el cambio

Cuando se trata el tema del cambio, la literatura examinada pone de manifiesto los diferentes aspectos que supone una condición de estabilidad y entender que el cambio se manifiesta como la postura del sujeto de estudio que es influenciada por elementos endógenos y exógenos propios de cada uno, y que estos elementos pueden ser comunes o diferentes para más de un sujeto dependiendo de las condiciones existenciales previas y posteriores a la evaluación.

Un concepto que ha estado examinando el cambio y su influencia sobre los individuos es señalado como *Response Shift*. *Response Shift* es una terminología en inglés cuya traducción sería Cambio en la Respuesta, pero dicha traducción no recoge el verdadero significado funcional del término⁴, por lo que las investigaciones actuales mantienen el vocablo en su idioma original.

El concepto *Response Shift* tiene sus primeros antecedentes en las investigaciones en el ámbito educativo de [Howard y Dailey \(1979\)](#) donde se refieren al cambio en los estándares internos de medida o en el marco de referencia de un individuo que incide en el establecimiento de una medida y otra. Estos autores señalan que el *Response Shift* es un sesgo manifestado entre una intervención del sujeto y otra, lo cual genera una afectación en los resultados finales de un cambio. Al examinar estas premisas sobre el sesgo de Howard y Dailey, se identifican ciertas similitudes generales con la tipología de Golembiewski donde señala que el cambio Beta es una variación en los intervalos de medición. Para el primero, la condición general no era muy clara por lo que lo estableció la existencia de un sesgo, mientras que el segundo autor identificó la influencia de la escala de medida en las respuestas obtenidas. Particularmente, Howard y Dailey no establecieron condiciones similares a lo denominado cambio Gamma; probablemente, por considerar que la condición del sujeto de estudio era un nuevo caso y diferente, o que no iniciaba en una etapa previa que sufrió cambios profundos del constructo.

⁴ Exposición dialéctica sobre el concepto presentada por [Vicente- Galindo \(2003\)](#)

Los estudios de cambios en la calidad de vida fueron expandidos al contexto de la psicología de la salud por [Schwartz y Sprangers \(1999\)](#). Al examinar el comportamiento de los individuos desde una perspectiva clínica, estos autores precisan que el *Response Shift* se genera como el cambio en el significado de una autoevaluación aplicada a un sujeto. Por medio de este enfoque se generan tres premisas fundamentales: la existencia de un cambio en los estándares internos de la medida, la existencia de un cambio en los valores y la existencia de un cambio en la definición del constructo.

Para los autores, estas premisas establecen que el cambio puede darse por uno de tres componentes:

- Recalibración, el cual tiene un enfoque psicológico y supone una variación interna en la escala de medición
- Repriorización, supone un cambio en la valorización del individuo a sus intereses y, en consecuencia, hay una variación en los niveles de importancia de constructo
- Reconceptualización, supone una redefinición completa del constructo

Con este nuevo enfoque de *Response Shift*, se examina como los individuos pueden sufrir cambios en su calidad de vida tomando en cuenta variaciones en la escala de medición, sus prioridades, y de replanteamientos en el nuevo significado de calidad de vida por parte de estas personas. Por tanto, se puede señalar que el concepto de *Response Shift* ha sufrido un proceso de evolución y, en la literatura examinada, se han presentado adecuaciones progresivas en la medida que se extrapola y aplica en otras áreas de estudio. La Tabla 9 presenta una evolución comparativa de este concepto.

[Vicente-Galindo \(2003\)](#) explica la caracterización de los tres componentes presentados por [Schwartz y Sprangers \(1999\)](#) al señalar que, inicialmente, la calidad de vida del individuo se evalúa por medio de instrumentos, que recopilan información, que pueden ser no objetivos, dado que la información es suministrada por el mismo encuestado convirtiendo cualquier resultado en un estado típico de *Response Shift*. Los estándares individuales de medida de funcionamiento son establecidos por las personas con base a su propia interpretación de la realidad tomando en consideración sus experiencias y

percepciones, las cuales pueden variar y establecer nuevas escalas a lo largo del tiempo como un efecto de las nuevas experiencias y puntos de vista. Como tal, el individuo quien dentro de los estudios de CVRS es considerado como el paciente, ha generado una nueva escala para medir su propio estado de su calidad de vida, llevando el cambio hacia el primer componente: la recalibración.

Tabla 9. Comparación de los componentes asociados al Response Shift por los diferentes autores

Autor	Golembiewski et al. (1976)	Howard y Dailey (1979)	Schwartz y Sprangers (1999)
Área de Estudio	Organizaciones	Educación	Psicología y Salud
Componentes	Alpha		Variación dentro de la misma escala
	Beta	Cambios en los estándares internos de medida (sesgo)	Recalibración
	Gamma		Repriorización y Reconceptualización

Cuando el paciente determina cambiar sus valores o prioridades, se incorpora hacia el segundo componente que es la repriorización. Este tipo de cambio se manifiesta en aspectos asociados a la CVRS, toda vez que la literatura señala que de los tres elementos que toman vital importancia para las personas en determinar su condición de vida (salud física, salud social y salud psicológica), el individuo puede darle prioridad a uno de estos elementos en base a su contexto de vida presente. Pero cuando la persona sufre alguna situación en este contexto de vida que le disponen perspectivas diferentes en base a sus intereses particulares o porque posee alguna afectación emocional o clínica, el enfoque de vida puede cambiar y por tanto sus nuevas prioridades. Esto es repriorización.

El tercer componente según esta tipología es la reconceptualización, el cual se formaliza cuando los aspectos que constituyen el constructo en estudio cambian en el tiempo. Para un individuo, puede darse una repriorización de los intereses, pero cuando existe un cambio total de los objetivos de vida, aquellos aspectos que le dan su

significancia a nivel personal pueden redefinir qué representa para ellos la calidad de vida. En consecuencia, se manifiesta una nueva conceptualización que permite medir la calidad de vida.

La aplicación de los componentes asociados a *Response Shift* es todavía utilizada en la investigación asociada al cambio. [Hamidou et al. \(2014\)](#) realizó un estudio prospectivo multicéntrico para evaluar el impacto del componente de recalibración en el *Response Shift* en el tiempo hasta el deterioro⁵ en las valoraciones de la CVRS en pacientes de cáncer y las influencias en las expectativas de calidad de vida comparadas con las aspiraciones iniciales en el tiempo hasta el deterioro. En sus análisis, que involucraron 381 mujeres con cáncer de mama, se destacó que a raíz de una recalibración de los estándares internos de los pacientes en la evaluación de sus condiciones de calidad de vida se logró evidenciar un deterioro de los individuos a lo largo del tiempo.

En el mismo, se señala que *Response Shift* es un proceso natural que podría distorsionar la interpretación de los cambios en las puntuaciones de CVRS en el tiempo en estudios comparativos de intervención. En consecuencia, la caracterización de *Response Shift* puede ser un requerimiento para obtener una evaluación válida y sensible del cambio a través del tiempo.

[Hamidou et al. \(2014\)](#) reitera que “la recalibración de los estándares internos de los pacientes de cáncer de mama posee un efecto significativo en el tiempo en las valoraciones sobre el deterioro de la calidad de vida” ... “Sin embargo, *Response Shift* ha sido definido como un fenómeno de tratamiento-dependencia, valores de los pre-test, post-test y *then-test* de los sujetos de control que solamente reflejan efectos debidos a su historia, maduración o experimentación” (p.6).

Luego de la revisión anterior, varios autores definen el concepto de *Response Shift* de la siguiente manera:

- *Response Shift* según [Schwartz y Sprangers \(2000\)](#), es el cambio en el significado de la autoevaluación de un individuo de un constructo objetivo

⁵ En el tiempo hasta el deterioro (*on time to deterioration - TTD*) supone examinar cómo los pacientes se deterioran a lo largo del período de estudio.

como resultado en los cambios en los estándares internos de la escala de medición (recalibración), en las escalas de valores (repriorización) o la definición del constructo objetivo (reconceptualización).

- *Response Shift* es un proceso natural de ocurrencias que pueden distorsionar la interpretación del cambio en la valorización de CVRS a lo largo del tiempo en los estudios de intervención. Por tanto, la evaluación de *Response Shift* puede ser necesaria para obtener una evaluación válida y sensitiva del cambio en el tiempo (Hamidou et al., 2011)
- Howard et al. (2011) define el *Response Shift* como el resultado de la recalibración, reconceptualización y repriorización de los estándares internos y referencias utilizadas para la autoevaluación. El hecho de no considerar *Response Shift* puede resultar en un exceso o falta de calidad en el registro de los verdaderos cambios fisiológicos

La definición del concepto puede tener variaciones dependiendo de la motivación del investigador, los antecedentes consultados o la experimentación realizada. Sin embargo, Schwartz y Sprangers (2000) señalan que un elemento clave en esta conceptualización, particularmente para los estudios enfocados a la CVRS, es que *Response Shift* es una adaptación a los cambios en la salud.

La definición más ampliamente utilizada para explicar *Response Shift* está basada en los componentes de Schwartz y Sprangers (2000) donde formula los elementos de recalibración, repriorización y reconceptualización en el significado de la autoevaluación de un individuo. Sin embargo, existe una vertiente de investigadores en los temas de calidad de vida que han iniciado a validar el constructo de *Response Shift*.

Este es el caso de Ubel et al. (2010) quienes presentan el artículo “Abandonando el lenguaje de *response shift*: un alegato a favor de la claridad conceptual en distinguir la recalibración de la escala en un verdadero cambio en la calidad de vida” señalando que esta línea de investigación se ha visto obstaculizada por una confusión conceptual creada por este término (*Response Shift*) que agrupa las fuentes de error de medición (como la escala de recalibración) con verdaderas causas de los cambios de calidad de

vida (como la adaptación). En consecuencia, los autores proponen abandonar el uso del término *Response Shift* a favor de términos menos ambiguos como la escala de recalibración y la adaptación.

La propuesta presentada por este grupo de investigadores se centra en señalar algunas inconsistencias de contexto para cada uno de los tres componentes de Schwartz y Sprangers, los cuales se describen a continuación:

- El primer componente de *Response Shift*, escala de recalibración, ocurre cuando se manifiestan circunstancias cambiantes que motivan a que las personas cambien la forma de cómo ellos interpretan una escala subjetiva de respuesta. Sin embargo, se descubre que mientras las personas envejecían, éstas tendían a redefinir su significado de una salud perfecta. Este planteamiento dificulta comparar las puntuaciones subjetivas de salud entre personas de diferentes grupos de edades, así como personas que experimentan enfermedades crónicas o discapacidades redefinirán su significado de un alto nivel de felicidad.
- Los reportes de calidad de vida también pueden ser influenciados en cómo la escala de valor de las personas cambia en respuesta a sus propias circunstancias, lo que se entiende como repriorización. Cambios en los valores pueden ser un mecanismo con los cuales las personas se adaptan emocionalmente a las enfermedades o discapacidades. [Ubel et al. \(2010\)](#) presentaron el caso de pacientes de cáncer quienes experimentaban los efectos secundarios sus tratamientos y quienes llegaron a cambiar sus valores, replanteando lo que era más importante en sus vidas. A pesar de las complicaciones en sus tratamientos, los pacientes fueron capaces de mantener una alta calidad de vida, mientras quienes no cambiaron sus valores experimentaron una declinación. Con esto se evidenció que las personas pueden manifestar *Response Shift* al cambiar sus escalas de valores, pero esta condición de *Response Shift* no necesariamente invalidaba las mediciones de calidad de vida. En contraste, la recalibración de la escala es por definición una amenaza a la validación de las mediciones de calidad de vida, y una persona

que recalibra una escala dificulta la comparación entre un reporte propio con respecto a otro.

- El tercer componente es la reconceptualización del constructo que es medido. La postura de [Ubel et al. \(2010\)](#) es cuestionar sobre el significado de la reconceptualización para la validación de los reportes de calidad de vida, pues para éstos la respuesta no es clara. Si se desea que las personas interpreten libremente su estado de calidad de vida, la reconceptualización no es una amenaza; sin embargo, si se desea que las personas juzguen sus calidades de vida en cualquier momento utilizando la misma definición de calidad de vida, entonces esto se convertiría en una amenaza. Por tanto, se necesitará encontrar los mecanismos para prevenir esta reconceptualización o desarrollar métodos para descubrir si la reconceptualización influye en los reportes realizados por las mismas personas.

Los cuestionamientos de [Ubel et al. \(2010\)](#) a las posturas tradicionales de Schwartz y Sprangers, aunque no muy claros, concluyen señalando que es momento para que los investigadores en materia de calidad de vida abandonen el término de *Response Shift* y dirijan sus esfuerzos en determinar la confiabilidad en las comparaciones de las medidas de calidad de vida en el tiempo.

Ante esta postura, [Sprangers y Schwartz \(2010\)](#) coinciden con Ubel, Peeters y Smith sobre la existencia de fenómenos diferentes e identificables que son incorporados dentro del término *Response Shift*, y que esta área temática se beneficia con la claridad conceptual. Sin embargo, las autoras refutan las críticas presentadas por [Ubel et al. \(2010\)](#) señalando que no se revisó la literatura sobre el tema en los últimos cinco años y es inadecuado la propuesta de cambiar el término *Response Shift* por recalibración y adaptación. Destacan con una clara explicación sobre los componentes y las implicaciones de *Response Shift* que los estudios hipotéticos utilizados no presentan una solución a los problemas identificados.

Reconocen que *Response Shift* ha generado confusiones conceptuales y operacionales, y que se necesita un lenguaje más preciso y una terminología más específica, mientras que este campo de estudio se beneficiará con una distinción explícita de la

recalibración de *Response Shift*, la repriorización de *Response Shift* y la reconceptualización de *Response Shift*. De igual forma, aceptan que la distinción realizada por Oort (2005) entre la perspectiva de medición y la perspectiva conceptual de *Response Shift* es un gran avance para otras investigaciones en el tema.

Concluyen aceptando la preocupación en torno a la confusión conceptual del término *Response Shift*, el cual necesita ser resuelta, debido a la complejidad del fenómeno que se encuentra en las etapas iniciales de investigación.

Reeve (2010) señala que no existe ventaja en abandonar el término *Response Shift* como sugiere Ubel et al. (2010)⁶, ya que el término es ampliamente conocido en este campo de estudio y ha generado impacto en la forma de medir longitudinalmente calidad de vida. Sin embargo, estos los planteamientos han proporcionado un incentivo para iniciar un diálogo sobre el tema con oportunidades de refinar el lenguaje sobre *Response Shift* y educar a los investigadores. Ante estas críticas, se ha identificado algunas oportunidades para continuar con los estudios: la selección de los conceptos de calidad de vida que serán medidos; los cuestionarios utilizados para evaluar los conceptos de calidad de vida; el diseño de estudios e investigaciones; la caracterización de la población objetivo; y el análisis y reporte de los resultados.

Eton (2010) realiza una evaluación sobre los artículos antes mencionados y destaca la utilidad que se genera en definir una ruta para nuevos avances en esta línea de investigación. A pesar de que existen diferencias de opinión, se identifican algunos puntos clave. Primero, la confusión conceptual en torno a *Response Shift* que necesita ser resuelta. Segundo, la mala utilización del término *Response Shift* como aquellos hallazgos anómalos que son difíciles de explicar. Y tercero, se necesita una aclaración más enfática de que *Response Shift* es más que una recalibración.

A pesar de que todos estos temas merecen ser atendidos, Eton señala que no se abandonará el término *Response Shift* hacia otras instancias.

⁶ Reeve (2010) realiza una crítica parcial a lo expuesto por Ubel et al. (2010) como parte de su artículo.

1.2.3. Modelo teórico del *Response Shift*

A continuación, se presenta la conceptualización de los modelos clásicos que permiten explicar el fenómeno de *Response Shift*.

1.2.3.1. El modelo Sprangers y Schwartz

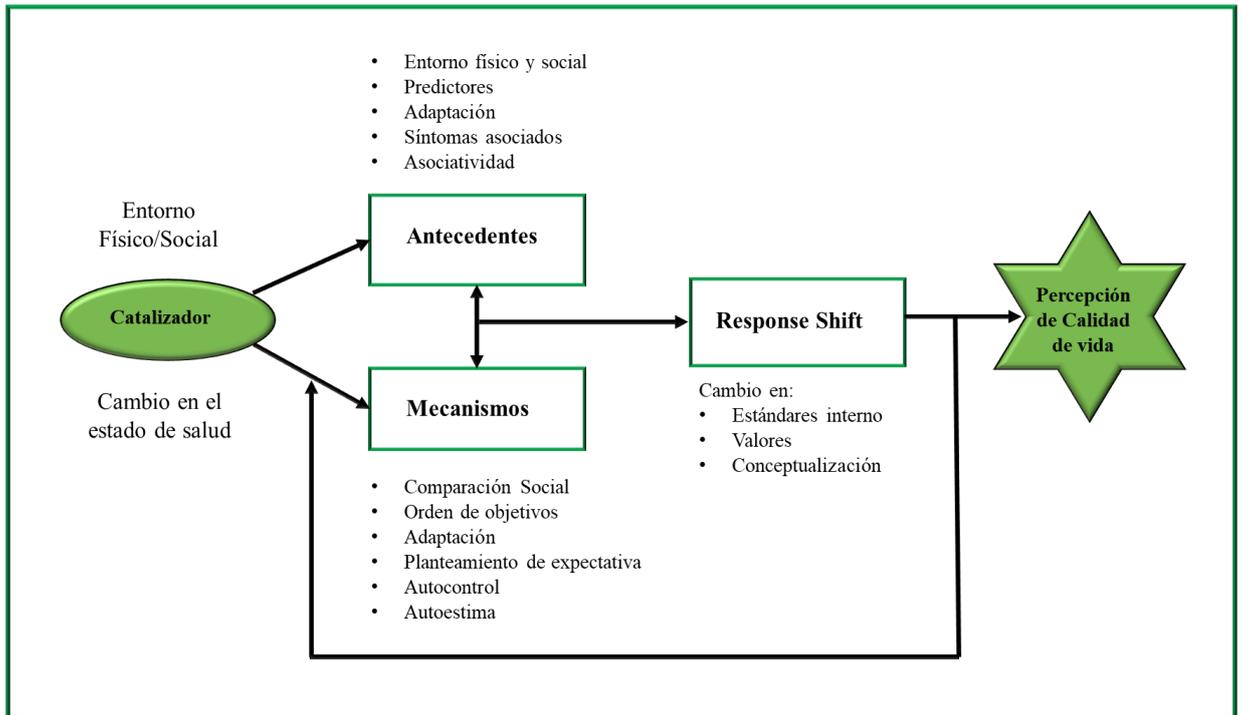
Sprangers y Schwartz (1999) formulan un modelo teórico que explica la presencia o ausencia de *Response Shift* como el efecto de un conjunto de elementos iniciado por los cambios en el estado de salud de una persona hasta la definición de su condición de calidad de vida. El modelo presenta cinco componentes: catalizador, antecedentes, mecanismos, *Response Shift* y calidad de vida percibida.

Según el modelo, el catalizador representa un cambio en el estado de salud, ya sea como el resultado o no de un tratamiento. Los antecedentes corresponden a las características y predisposiciones del individuo, mientras que los mecanismos se refieren a los procesos afectivos de comportamiento y de bienestar que utiliza el individuo para adaptar el catalizador.

El *Response Shift*, como componente del modelo, es el cambio en el significado de la autoevaluación de la calidad de vida como resultado de cambios en los estándares internos, en los valores o en la conceptualización de CVRS; y la calidad de vida percibida es aquel constructo multidimensional que incorpora por lo menos tres dominios: físico, social y psicológico. La Figura 2 presenta un esquema general del modelo.

El esquema muestra una secuencia donde un paciente es afectado por un catalizador (cambio en su estado de salud por una enfermedad o discapacidad) y del cual se desprenden mecanismos que permiten al individuo enfrentar su nueva condición y sobrellevar el padecimiento. Los antecedentes influyen paralelamente como todos aquellos elementos previos que expanden o reducen las condiciones de la enfermedad. Ante estas condiciones, el paciente puede atravesar por un proceso de *Response Shift* en el cual puede generar cambios en sus estándares internos, valores y definición sobre su condición de vida, para posteriormente establecerla.

Figura 2. Modelo de Sprangers & Schwartz de Response Shift de calidad de vida



Fuente: Sprangers y Schwartz (1999)

1.2.3.2. El modelo Lepore y Eton

El modelo desarrollado por Lepore y Eton (2000) establece que *Response Shift* refleja procesos cognitivos que mantienen o mejoran la calidad de vida en personas que enfrentan serios problemas de salud, y quienes se adoptan uno de dos modelos para explicar el papel del *Response Shift* en la percepción de la calidad de vida: el modelo supresor y el modelo amortiguador (*Suppressor and Buffer Models*).

El modelo supresor establece que la presencia de *Response Shift* suprime la relación entre el problema de salud y la percepción de calidad de vida. Este principio surge como explicación a que los problemas físicos de una persona actúan como catalizador para el *Response Shift* motivando automáticamente a una recalibración de sus estándares, un cambio en los valores o una reconceptualización del significado de calidad de vida, los cuales llevan al paciente a limitar la posible afectación de la enfermedad (catalizador).

Por el otro lado, el modelo amortiguador establece que el *Response Shift* puede ser estimulado por mecanismos intra e interpersonales como “estrategias de afrontamiento” y “apoyo social”, permitiendo que algunos individuos desarrollen *Response Shift* y otros no dependiendo de la intensidad de la colaboración externa durante el tratamiento y la valoración de las condiciones de la vida del propio paciente.

1.2.3.3. El modelo Wilson y Cleary

Wilson y Cleary (1995) presentan un modelo para explicar *Response Shift* como un cambio no concordante en dos o más dimensiones de la salud de un individuo. Para lograr identificar estos cambios, los autores establecen cinco tipos diferentes de medidas que permiten evaluar el estado de salud de individuo: biológica y fisiológica, síntomas, funcionamiento, percepción de salud general, y calidad de vida.

Las variables biológicas y fisiológicas sirven para evaluar el verdadero funcionamiento de células y órganos, y como se determinan por medio de un reporte elaborado por el mismo paciente, no son implicadas en el fenómeno de *Response Shift*. Los síntomas representan percepciones de un estado físico o fisiológico anormal, los cuales pueden variar en frecuencia e intensidad dependiendo de la percepción del paciente, el significado y la respuesta a esos síntomas, los cuales pueden ser recalibrado y/o reconceptualizados. Mientras que las variables de funcionamiento evalúan la capacidad para desarrollar tareas específicas. La percepción general de salud evalúa la percepción del paciente de su nivel de salud en general, y la calidad de vida se considera la medida de la satisfacción con la vida. Ambos aspectos son altamente susceptibles al *Response Shift*.

De acuerdo con los proponentes del modelo, si el cambio en dos o más dimensiones sigue trayectorias paralelas, no hay porqué pensar que existe la presencia de esta condición. Pero un cambio en alguna dimensión mientras que las demás permanecen estables sugiere que el paciente se ha enfrentado a un caso de *Response Shift*. Por lo tanto, se puede establecer que *Response Shift* puede estar siempre presente.

1.2.3.4. El modelo Rapkin y Schwartz

Mediante una visión actualizada del abordaje de cómo *Response Shift* influye en la evaluación de calidad de vida, Rapkin y Schwartz (2004) destacan la influencia que posee la respuesta de un individuo a cualquier instrumento de autoevaluación. La valoración por las diferencias de los individuos o los cambios longitudinales pueden afectar cómo las personas responden a los ítems de calidad de vida. Factores como las circunstancias personales, eventos estresantes, progresión de las enfermedades y las intervenciones médicas también influyen en los criterios individuales utilizados para evaluar la propia condición de vida. La valoración es una faceta escondida de todas las mediciones de calidad de vida, y todos los estudios que incluya la autopercepción de los individuos pueden estar influenciadas por su propia valoración.

En el marco de cómo la valoración individual puede afectar la definición y detección del *Response Shift*, los autores proponen cuatro procesos cognitivos diferentes y que complementan el modelo original de Sprangers y Schwartz (1999) de *Response Shift*. Estos procesos se describen de la siguiente manera:

- La evaluación de calidad de vida induce un marco de referencia correspondiente a las experiencias individuales que influyen directamente en las respuestas. Este marco de referencia depende del significado que tenga cada uno de los individuos sobre una pregunta.
- Los individuos seleccionan experiencias dentro de sus marcos de referencias. Por tanto, se propone una estrategia de selección subjetiva que es por lo menos determinado en parte por el mismo ítem, la medición del contexto de calidad de vida y la situación de evaluación.
- Cada experiencia seleccionada es calificada mediante los estándares subjetivos de comparación. Esta comparación puede estar basada en los puntos de referencias personales incluyendo las funciones previas, capacidades perdidas o experiencias extremas. Debe considerarse que los individuos pueden seleccionar estándares con algún tipo de sesgo producto de alguna tendencia, y no necesariamente su verdadera opinión.

- Para obtener las ponderaciones de calidad de vida, los individuos deben establecer algoritmos combinatorios para totalicen sus evaluaciones con base a experiencias relevantes y que permitan formular una respuesta.

La propuesta de [Rapkin y Schwartz \(2004\)](#) permite una retroalimentación dinámica para explicar cómo las puntuaciones de calidad de vida pueden estar impactadas por las percepciones individuales de las personas, las cuales deben ser estabilizadas contabilizando las diferencias interindividuales y temporales a pesar de los cambios en el estado de salud de un individuo. En resumen, cualquier respuesta a los ítems de calidad de vida puede entenderse como una función del proceso de valoración.

A lo largo de esta revisión conceptual, se ha evidenciado la necesidad de entender los elementos asociados para determinar la calidad de vida de un individuo, y particularmente la posibilidad de su medición. Esto con el propósito de poder establecer científicamente qué tanto ha mejorado o se ha deteriorado esta condición personal de cada uno. El contexto de salud ofrece un amplio espectro de oportunidades para desarrollar estos estudios, toda vez que la disponibilidad de datos y su afluencia permiten reducir la brecha entre la conceptualización teórica del fenómeno y la realidad experimentada y comprobable.

Sin embargo, correspondería realizar una revisión sobre los métodos utilizados para la detección del *Response Shift* para así examinar las diferentes metodologías recientes utilizadas para de datos de CVRS.

CAPÍTULO II

MÉTODOS PARA LA DETECCIÓN DE *RESPONSE SHIFT* EN CALIDAD DE VIDA RELACIONADA CON LA SALUD: UNA REVISIÓN SISTEMÁTICA

2. MÉTODOS PARA LA DETECCIÓN DE *RESPONSE SHIFT* EN CALIDAD DE VIDA RELACIONADA CON LA SALUD: UNA REVISIÓN SISTEMÁTICA

2.1. Introducción

[Schwartz \(2010\)](#) en *Applications of Response Shift Theory and Methods to Participation Measurement: A Brief History of a Young Field* señala que *Response Shift*, como una medida para detectar el cambio, continúa siendo un tema estudiado y señala que los avances teóricos deben proporcionar mediciones diferentes o métodos cuantitativos más rigurosos que permitan detectar los cambios en los individuos a lo largo del tiempo.

En un esfuerzo de esclarecer la importancia de los efectos de *Response Shift* en la interpretación de los resultados de Calidad de Vida Relacionada con la Salud (CVRS), [Schwartz \(2010\)](#) realiza un meta-análisis con información hasta el 2005 donde se registró una colección de 494 artículos con la palabra clave *Response Shift* y se examinaron aquellos que utilizaron métodos para detectar y analizar cambios con datos longitudinales.

El estudio destacó importantes señalamientos con respecto a las investigaciones de *Response Shift*:

- Cuestiona la sensibilidad y utilidad del método *Then-Test* ya que la técnica supone la aplicación de un mismo instrumento aplicado en varios momentos con la posibilidad de un sesgo que puede estar dado por el error de memoria.
- Los artículos publicados e incluidos en el meta-análisis no contaron con la información suficiente y adecuada para detectar *Response Shift* en sus métodos, por lo que recomendó estandarizar la metodología en futuras investigaciones.

El estudio de [Schwartz \(2010\)](#) dejó en evidencia la necesidad de desarrollar y utilizar nuevos métodos para detectar *Response Shift*.

El continuo interés de detectar estos cambios motivó a la realización de una revisión sistemática que permitiera cosechar y examinar la literatura reciente y los métodos estadísticos utilizados en la detección de *Response shift* en estudios de CVRS.

Esta revisión sistemática permitió identificar los métodos tradicionalmente utilizados en estudios clínicos específicos vinculados a la CVRS (datos primarios) o en estudio de bases de datos elaboradas por instituciones médicas, académicas o centros de investigación (datos secundarios) sobre este tema.

2.2. Revisión sistemática: Una metodología de búsqueda

2.2.1. Conceptualización

La revisión sistemática es una metodología que permite identificar de forma estructurada diferentes estudios específicos en base a criterios de inclusión y exclusión, y a su vez facilita el proceso de refinamiento de la elegibilidad de las publicaciones relevantes o de interés (Montori et al., 2003). Sáenz (2001) la define como la búsqueda exhaustiva de estudios relevantes sobre un tema, donde una vez identificados, los resultados son sintetizados de acuerdo con un método preestablecido y explícito. Considerando las definiciones anteriores, Kitchenham et al. (2009) establece la revisión sistemática como un tipo de estudio secundario (investigaciones sobre lo investigado) que utiliza una metodología bien definida para identificar, analizar e interpretar toda evidencia disponible relacionada con la pregunta específica de manera que sea imparcial y reproducible. Desde un punto de vista formal, la revisión sistemática sintetiza los resultados de investigación primaria mediante estrategias que limitan el sesgo y error aleatorio (Cook et al., 1997). Su objetivo es identificar, evaluar y combinar la evidencia de estudios originales primarios usando un método riguroso (Carrizo & Moller, 2018).

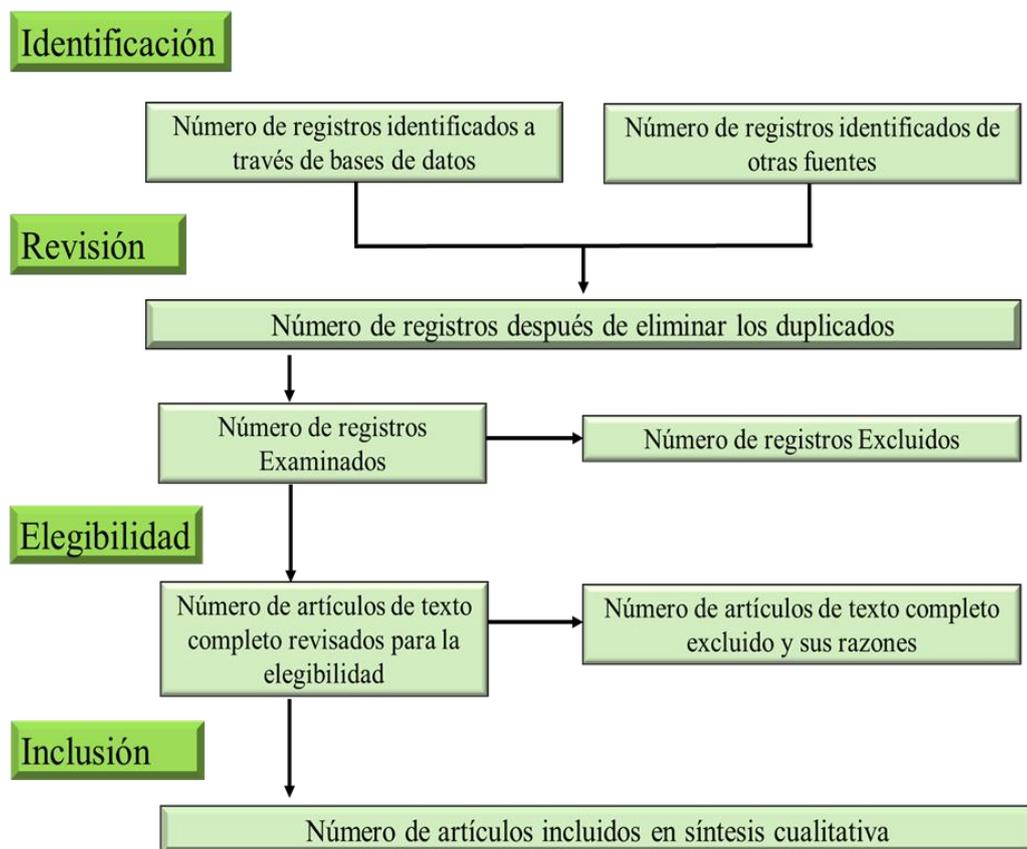
Las revisiones sistemáticas cuentan con un protocolo de investigación que describe los objetivos, diseño y metodología que deberá contener los siguientes elementos:

- Definición de las preguntas de investigación y objetivo
- Definición de los criterios de la inclusión y exclusión de la revisión sistemática

- Identificación de las bases de datos y métodos de búsqueda que se utilizará
- Definición de términos de búsquedas
- Búsqueda de bases de datos científicas y extracción de contenido y datos relevantes
- Evaluación de la calidad de los resultados
- Reunir los resultados más sobresalientes para el análisis

Para desarrollar una revisión sistemática, se recomienda utilizar la metodología PRISMA (Moher et al., 2015) que mediante criterios de elegibilidad preespecificados, permite reducir el sesgo en la identificación, selección, síntesis y resumen de los resultados de estudios previamente publicados (Moher et al., 2009). Esta metodología consiste en una lista de verificación de 27 ítems y un diagrama de flujo de cuatro fases: Identificación, Revisión, Elegibilidad e Inclusión, tal como muestra la Figura 3.

Figura 3. Fases de la revisión sistemática utilizando la metodología PRISMA



Fuente: Moher et al. (2009)

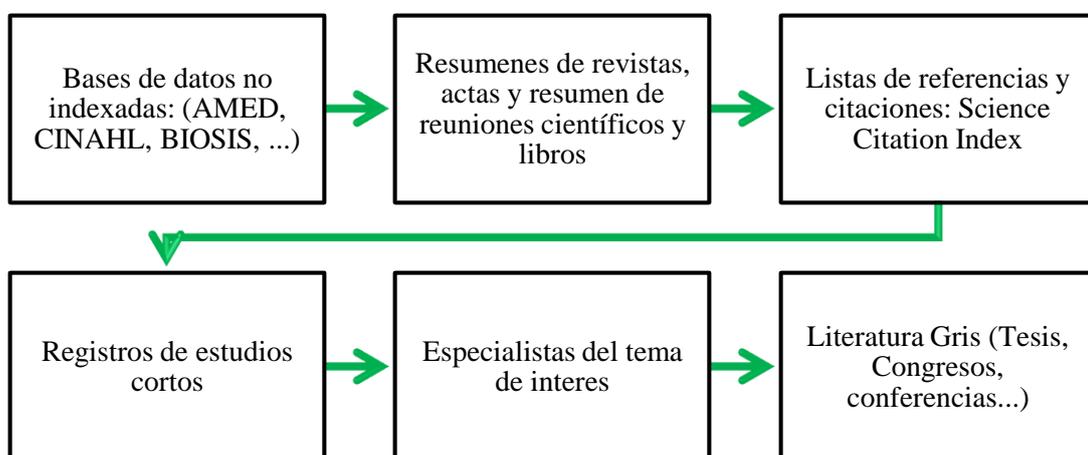
La alta capacidad de precisión y confiabilidad de esta herramienta ha permitido importantes beneficios en investigaciones relacionadas con la salud (Liberati et al., 2009). Publicaciones recientes reflejan la aceptación de esta metodología en estudios de *Response Shift* tales como datos reportados por pacientes (*patient-reported outcomes* - PRO) (Hinds et al., 2018), estudios oncológicos (Ilie et al., 2019b; Nielsen et al., 2017), ortopedia (Powden et al., 2018) y de lesión previa (Scholten et al., 2017).

2.2.2. Fuentes de datos y estrategia de búsqueda

2.2.2.1. Fuentes de datos

Una revisión sistemática requiere de una búsqueda en más de una base de datos, donde la selección de estas bases de datos dependerá del objetivo de la revisión sistemática. De acuerdo con lo recomendado por el Manual Cochrane de Revisiones Sistemáticas de Intervenciones (The Cochrane Collaboration, 2011), las principales bases de datos para estudios primarios en ciencias de la salud son MEDLINE, EMBASE, PsycInfo, Web of Science, Scopus y CENTRAL como bases de datos electrónicas. Es importante señalar que la búsqueda en diferentes bases de datos dará definitivamente resultados diferentes. Otras fuentes que se pueden considerar según el tipo y profundidad de la revisión son bases de datos no indexadas, resúmenes de revistas, estudios cortos y literatura gris, entre otros (Figura 4).

Figura 4. Fuentes alternativas para la búsqueda de literatura en una revisión sistemática



2.2.2.2. Estrategias de búsqueda

A partir del objetivo general y empleando la metodología *Population, Intervention, Comparison, Outcomes, Context*, se deberán definir las palabras clave empleadas en la ecuación de búsqueda. Además, para ampliar la misma, se podrán utilizar sinónimos que sean usualmente empleados en la literatura académicas, generalmente aquellos que se usen en listas de palabras o sean términos controlados.

La búsqueda para cada base de dato será la misma ecuación de búsqueda permitiendo generar resultados equivalentes.

Estas estrategias de búsqueda pueden ser simple o avanzada. Cuando se refiere a una estrategia simple se introducen palabras clave en una única ventana para obtener todos los documentos que contengan dichas palabras clave, ya sea en el título o en el resumen del artículo. En las búsquedas avanzadas, se debe conocer uno o más términos de campos específicos, como autor, título del artículo, título de revistas, año de publicación, etc.

De acuerdo con [Martin-Rodero \(2014\)](#), una ecuación de búsqueda eficaz es aquella formada por descriptores y sus correspondientes clasificadores combinados entre si mediante operadores booleanos:

- *AND*: Relaciona términos (operador de intersección)
- *OR*: Une conceptos relacionados (operador de unión)
- *NOT*: Elimina documentos que contengan términos no deseados (excluye)

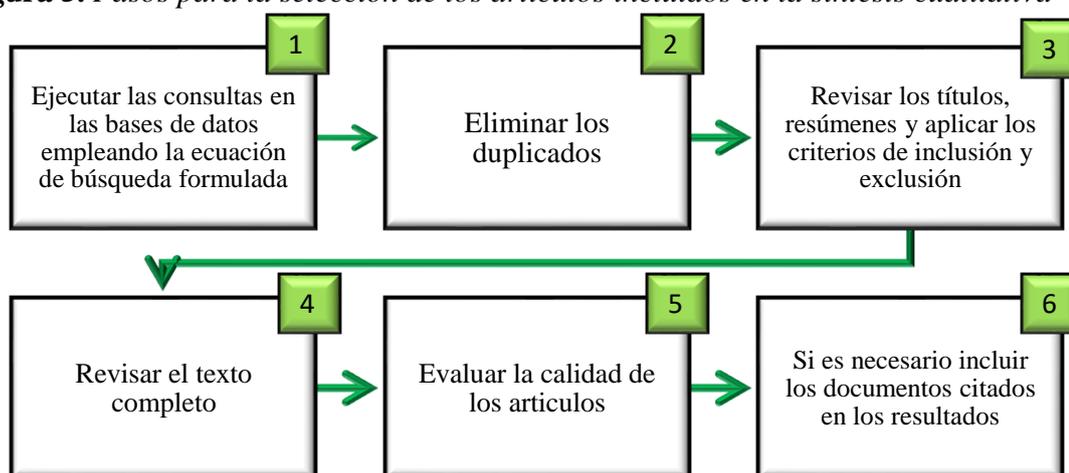
La aplicación de estrategias de búsqueda en diferentes bases de datos electrónicas proporciona un gran número de referencias bibliográficas, considerando las referencias duplicadas que se obtengan entre diferentes bases de datos, por lo que se recomienda la utilización de programas para la gestión automatizada de las citas bibliográficas como ProCite o Reference Manager ([Ferreira-González et al., 2011](#)).

2.2.3. Selección de los artículos

La selección de los artículos potenciales que formarán parte del estudio inicia a partir de la revisión de los títulos y resúmenes identificados, de los cuales se diseña una lista de criterios explícitos y comprensibles (Meade & Richardson, 1997). Esta revisión deberá estar a cargo de, por lo menos, dos revisores de forma independiente, con el objetivo de aumentar la fiabilidad y la seguridad del proceso (Ferreira-González et al., 2011) además de evaluar el grado de acuerdo entre los revisores mediante el cálculo del estadístico Kappa (Altman, 1991; Streiner et al., 1995) para cada uno de los ítems de la hoja de selección.

Los pasos para la selección de los artículos son detallados en la Figura 5.

Figura 5. Pasos para la selección de los artículos incluidos en la síntesis cualitativa



Es importante tener en cuenta todo el proceso de localización y selección de artículos que se han reportado, indicando el número de artículos identificados, así como los duplicados los que se han eliminado y la causa de la eliminación.

2.2.4. Calidad de los artículos seleccionados

Para la evaluación de la calidad de cada uno de los artículos se deberá realizar una lista para verificar los aspectos relevantes de cada artículo seleccionado, esto permite considerar cómo cada uno de los artículos seleccionados aportan a la pregunta de

investigación. Para evaluar la lista de verificación se establece una puntuación para decidir si se incluye o excluye el artículo en base a la calidad de su contenido con relación al tema examinado. Comúnmente se utiliza una escala nominal en base a las categorías *Si Cumple, Parcialmente Cumple, No Cumple*” y se le asignan una puntuación de 1.0, 0.5 y 0.0, respectivamente.

Una vez asignada la puntuación a cada uno de los ítems de la lista de verificación se suman las puntuaciones asignadas y se presentan los resultados en una tabla para así evidenciar la calidad metodológica de los artículos incluidos en la revisión sistemática, donde el investigador establecerá el punto de corte.

[Cruz-Benito \(2016\)](#) propone una lista de preguntas como guía que podrían ser utilizadas como lista de verificación (Véase Tabla 10).

Tabla 10. Preguntas utilizadas como lista de verificación

No.	Preguntas	Escala (Y, P, N)	Puntuación (1.0; 0.5; 0)
1.	¿Están claramente especificado los objetivos de la investigación?		
2.	¿El estudio fue diseñado para lograr estos objetivos?		
3.	¿Se describen claramente las técnicas utilizadas y se justifica?		
4.	¿Se mide adecuadamente las variables consideradas?		
5.	¿Se describe claramente los métodos de recolección de los datos?		
6.	¿Se describe claramente los datos recopilados?		
7.	¿Está claro el propósito del análisis de datos?		
8.	¿Las técnicas estadísticas para analizar los datos están adecuadamente descriptas y su uso justificado?		
9.	¿Se presentan los resultados negativos (Si hubiera)?		
10.	¿Los investigadores discuten algún problema?		
11.	¿Se responde adecuadamente todas las preguntas de investigación?		
12.	¿Qué tan claro son los vínculos entre datos, interpretación y conclusión?		
13.	¿Los hallazgos se basan en proyectos múltiples?		

Fuente: [Cruz-Benito \(2016\)](#)

2.3. Una revisión sistemática sobre los métodos para detectar *Response Shift* en calidad de vida relacionada con la salud

Las investigaciones sobre la CVRS y los procesos para la detección del *Response Shift* han mostrado un continuo interés en su estudio impulsado por una fuerte motivación de establecer métodos estadísticos más robustos que permitan identificar de manera más precisa y con rigor metodológico el nivel de afectación percibido por un paciente a lo largo de un tratamiento médico.

[Sajobi et al. \(2018\)](#) realizaron una revisión sistemática con el propósito de identificar los métodos utilizados para la detección de *Response Shift* y sus diferentes tipos basado en resultados informados por pacientes (*patient-reported outcomes*, por sus siglas en inglés, PRO). Con un número de 1,032 artículos inicialmente identificados entre 2005 a 2016, un total de 101 artículos (9.8%) cumplieron satisfactoriamente los criterios de inclusión y fueron seleccionados para el análisis.

Los resultados destacaron que la metodología más utilizada es el diseño retrospectivo *Then-Test* que permite evaluar el cambio del estándar interno del paciente comparando las puntuaciones de los ítems con otros dos momentos el *Pre-Test* y el *Post-Test*. Entre otros métodos identificados y que fueron muy utilizados en datos primarios como secundarios se encuentran los Modelos de Ecuaciones Estructurales (SEM) de [Schmidt \(1982\)](#) y [Oort \(2005\)](#), Teoría de Respuesta al Ítem (IRT) de [Anota et al. \(2014\)](#) y [Guilleux, et al. \(2015\)](#), Método de Importancia Relativa de [Lix et at. \(2013\)](#), Análisis de Trayectoria Latente de [Ahmed et al. \(2011\)](#), Modelo de Árbol de Regresión de [Li y Schwartz \(2011\)](#).

El marco teórico presentado en el Capítulo II (Métodos estadísticos para la detección del *Response Shift*) y el Capítulo III (Modelos de Tres Vías: Propuesta para la detección del *Response Shift*) examinan todo un conjunto de métodos estadísticos que son viables para el estudio de este fenómeno. Al reconocer que [Sajobi et al. \(2018\)](#) no destacan la utilización de muchos de los métodos estudiados en esta revisión literaria, surge la siguiente pregunta ¿Existen otros métodos estadísticos emergentes que se

estén utilizando actualmente en la detección del *Response Shift* y que puedan otorgar resultados más robustos con respecto a este proceso?

En consecuencia, se realiza una revisión sistemática que permita localizar, analizar y sintetizar sistemáticamente la evidencia existente sobre los métodos estadísticos utilizados en la detección de *Response Shift* en estudios de CVRS en un marco de tiempo actualizado.

2.3.1. Fuente de datos

Tomando en cuenta que la revisión sistemática se realizará bajo el fundamento del *Response Shift* y CVRS, se examinaron diferentes bases de datos que registraran publicaciones orientadas a las ciencias de la salud y psicología. Luego de este examen, se seleccionaron las siguientes bases de datos: MEDLINE/PubMed, Scopus, Web of Science Core Collection, y PsycINFO. Con el propósito de reducir el sesgo en la selección y ser más inclusivo en el número de publicaciones, se incorporó el buscador Google Scholar.

Los términos considerados para la búsqueda en todas las bases de datos y buscador fueron: “Quality of Life”, “Health-Related Quality of Life” y “Response Shift” tanto como Descriptores o Palabras Clave en los campos título y resumen.

2.3.2. Estrategia de búsqueda

La estrategia de búsqueda aplicada a la base de datos MEDLINE vía PubMed se planteó con la siguiente ecuación: “Quality of Life” [Mesh] OR “Quality of Life” [Title/Abstract] OR “Life Quality” [Title/Abstract] OR “Health-Related Quality of Life” [Title/Abstract] OR “Health Related Quality of Life” [Title/Abstract] OR “QoL” [Title/Abstract] OR “HRQoL” [Title/Abstract]) AND “Response Shift” [Title/Abstract].

En la búsqueda se establecieron dos términos fundamentales que son no mutuamente excluyente: la inclusión del concepto CV o CVRS, y la inclusión del concepto

Response Shift. Esto se estableció con el fin de permitir una mayor cosecha de publicaciones que traten sobre el tema y que los resultados preliminares fueran propios y pertinentes al objetivo de esta revisión.

A su vez, se excluyeron aquellos términos que fueran inherentes al concepto *Response Shift* como Recalibración, Repriorización y Reconceptualización, entre otros relacionados, toda vez que al utiliza estos vocablos se está tratando sobre el tema *Response Shift*. A su vez, también se excluyeron términos relacionados a las características de los datos comúnmente utilizados en estos temas, como datos longitudinales, invarianza, datos primarios y secundarios, entre otros.

A diferencia de las revisiones sistemáticas realizadas por [Sajobi et al. \(2018\)](#) y [Ilie et al. \(2019b\)](#) que utilizaron este conjunto de términos en sus búsquedas, nuestra investigación planteó de manera intencional la especificación de los términos con el propósito de evitar cosechar publicaciones con orientaciones en otras áreas como ciencias sociales, administrativas, economía, etc.

Se utilizaron los siguientes filtros: “Humans”, “Journal Article”, “English”, “2010/01/01 - 2020/07/31”. La estrategia consideró que solamente se realizaría el análisis con documentos publicados en idioma inglés y en revistas científicas que tratarán sobre casos aplicados a humanos. La búsqueda se realizó en agosto de 2020 y cuyo período abarcó del 1 de enero de 2010 hasta el 31 de julio de 2020.

Esta estrategia de búsqueda se adaptó a cada una de las otras bases de datos consultadas y se completó con el análisis del listado bibliográfico de los artículos que fueron seleccionados.

2.3.3. Selección de artículos

Se eligieron para el estudio aquellos artículos que cumplieron los siguientes criterios:

- adecuarse a los objetivos de la búsqueda de localización y síntesis de los métodos estadísticos utilizados en la detección de *Response Shift* en estudios de CVRS,
- estar publicados en revistas revisadas por pares,

- poder recuperar el texto completo del trabajo

Se consideró como criterio de elegibilidad que el término *Response Shift* estuviera incluido en el título, resumen y/o palabra clave.

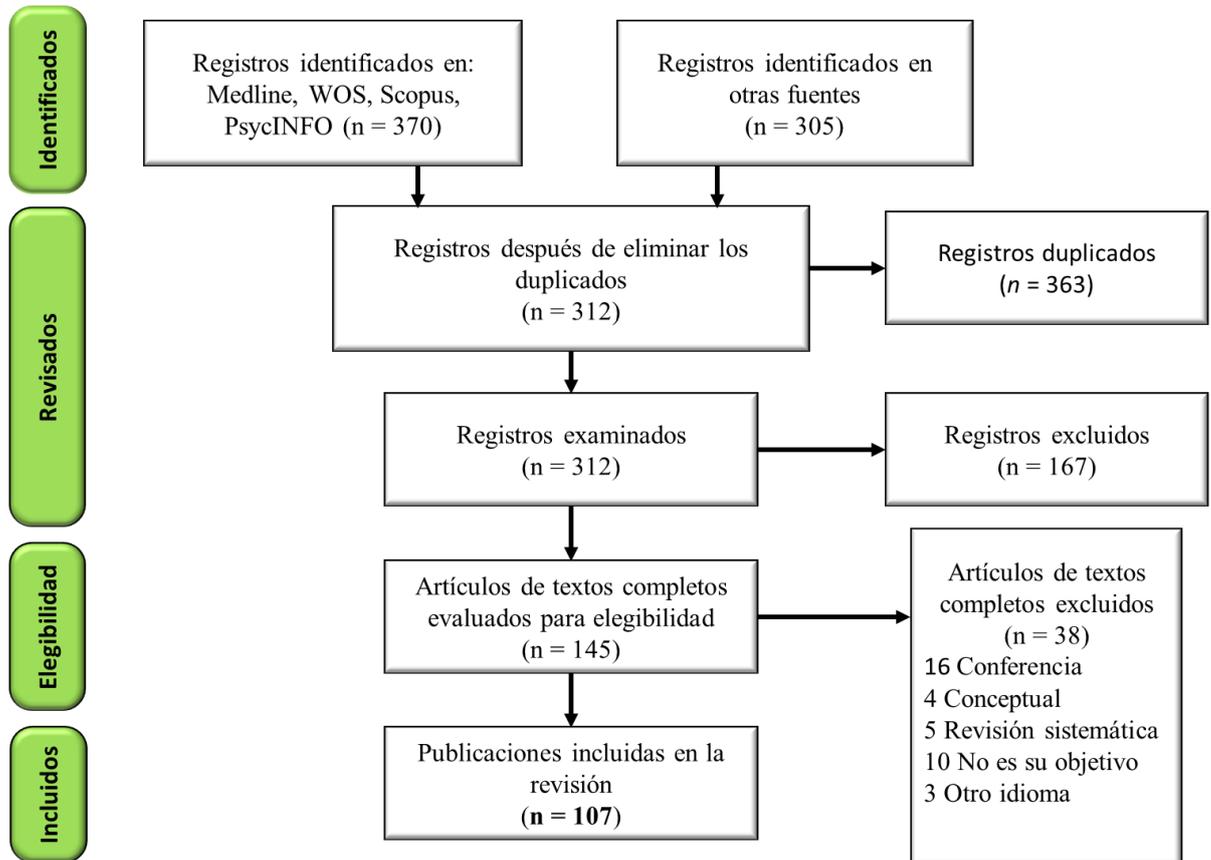
Como criterios de exclusión, se consideró:

- aquellos estudios en otro idioma distinto del inglés,
- que no incluyeran de manera explícita y detallada la utilización de métodos estadísticos para la detección de *Response Shift*,
- estudios cuyo objetivo principal no fuera la detección de *Response Shift* y su clasificación,
- documentos basados en conferencias, notas editoriales, publicaciones de revisiones sistemáticas, evaluaciones conceptuales, tesis y cualquier otra literatura gris.

La selección de los artículos fue llevada a cabo de manera independiente por dos autores (Estelina Ortega-Gómez y Purificación Galindo-Villardón) que procedieron inicialmente a la lectura de los títulos, después de los resúmenes, y finalmente del artículo a texto completo. Para la inclusión de los artículos en el estudio, se estableció que la valoración de la concordancia entre autores debía ser superior al 80% (índice de Kappa). En caso de discrepancias se repitió el proceso hasta que estas fueron resueltas mediante consenso entre todos los autores.

Estos criterios de inclusión y exclusión se aplicaron a todos los artículos en la revisión de elegibilidad. La búsqueda de la literatura y la selección de la publicación es ilustrada en la Figura 6 por el diagrama de flujo PRISMA.

Figura 6. Diagrama de flujo PRISMA de la estrategia de búsqueda y los artículos incluidos en la revisión



2.3.4. Calidad del estudio

Una vez revisado los artículos seleccionados tomando en cuenta los criterios de inclusión y exclusión se procedió a evaluar la calidad de estos. Para evaluar la calidad de los artículos seleccionados se elaboró una lista de preguntas (Cruz-Benito, 2016), véase Tabla 10, sobre los aspectos relevantes que deberá cumplir cada artículo, luego se procede a asignar una puntuación de la evaluación y se decidirá la inclusión o exclusión del artículo para el análisis de aquellos seleccionados. Se repite la aplicación de los criterios, y por último se deberá fijar un punto de corte para la evaluación.

2.3.5. Características descriptivas del estudio

La estrategia de búsqueda identificó un total de 675 artículos (tal como muestra la Figura 6) en las bases de datos MEDLINE/PubMed (141), Scopus (46), Web of Science Core Collection (114), PsycINFO (69) y Google Scholar (305). Se eliminaron 363 registros duplicados, 167 registros que no incluían el término “*Response Shift*” en los campos título, *abstract* y/o palabra clave, y otros 38 artículos fueron excluidos por no cumplir con los criterios de inclusión establecidos, quedando así 107 publicaciones incluidas en el análisis, lo que correspondió al 15.9% del total de los artículos obtenidos de búsqueda inicial ($n=107$). El listado de los artículos incluidos para el análisis se presenta en la Tabla 11.

Teniendo en cuenta el año de publicación, se observó que el mayor número de artículos se publicaron en 2017 con 15 estudios, seguidos por 2014 y 2016 con 12 y 13 aportes, respectivamente. El resto de los años muestra el continuo interés de estas investigaciones (Figura 7). La Tabla 12 muestra que un total de 56 revistas fueron identificadas en la revisión sistemática, donde el 31.8% de los artículos fueron publicados en *Quality of Life Research* (34), seguido por *Health & Quality of Life Outcomes* (10), *Journal of Clinical Epidemiology* (4) and *European Journal of Cancer Care* (4).

Figura 7. Número de publicaciones por año: 2010-2020

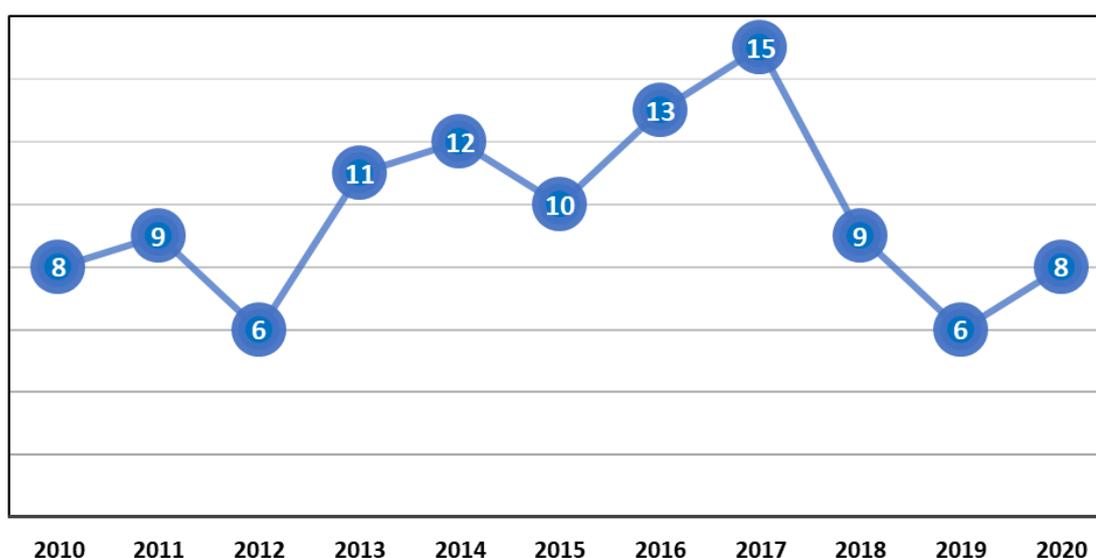


Tabla 11. Artículos incluidos en la revisión sistemática: Enero 2010 – Julio 2020 (1/6)

Autor	País	Año	Especialidad del estudio	Tipo de datos	Muestra (Inicio)	Muestra (Final)	Tipo de población	Instrumento	Response Shift
(Dempster et al., 2010)	Irlanda Norte	2010	Cardiología	Primario	84	57	Mixta	SEIQOL-DW	Si
(Ito et al., 2010)	Japón	2010	Gastroenterostomía	Primario	20	13	Mixta	SF-36	Si
(Kievit et al., 2010)	Países Bajos	2010	Artritis	Primario	212	197	Mixta	DAS28 / VAS	Si
(Kvam et al., 2010)	Noruega	2010	Oncología	Primario	260	239	Mixta	EORTC QLQ-C30	Si
(McPhail & Haines, 2010)	Australia	2010	Medicina General	Primario	103	101	Mixta	EQ-5D / VAS	Si
(Razmjou et al., 2010)	Canadá	2010	Ortopedia	Primario	107	107	Mixta	ASES-36	Si
(Taminiau-Bloem et al., 2010)	Países Bajos	2010	Oncología	Primario	92	50	Mixta	EORTC QLQ-C30	Si
(King-Kallimanis & Oort, 2010)	Países Bajos	2010	Pacientes con HIV	Primario	403	403	Mixta	HAART	Si
(Ahmed et al., 2011)	Canadá	2011	Neurología	Secundarios	1566	1312	Mixta	SF-12	No
(Hinz et al., 2011)	Alemania	2011	Urología	Primario	427	275	Masculino	PHQ-2	No
(King-Kallimanis et al., 2011)	Países Bajos	2011	Neurología	Secundarios	2952	1767	Mixta	NARCOMS/SF-12	No
(Li & Schwartz, 2011)	Estados Unidos	2011	Neurología	Secundarios	3839	3008	Mixta	NARCOMS/SF-12	Si
(Carolyn Schwartz et al., 2011)	Estados Unidos	2011	Neurología	Secundarios	3008	3008	Mixta	NARCOMS/ SF-12	Si
(Shi et al., 2011)	Taiwán	2011	Gastrointestinal	Primario	359	353	Mixta	GIQLI	Si
(Höfer et al., 2011)	Austria	2011	Cardiología	Primario	64	64	Mixta	SEIQoL y MacNew	Si
(Lacey et al., 2011)	Estados Unidos	2011	Obesidad	Primario	1865	1865	Mixta	-	Si
(Barclay-Goddard et al., 2011)	Canadá	2011	Neurología	Secundarios	677	677	Mixta	SF-36	Si
(Galenkamp et al., 2012)	Países Bajos	2012	Gerontología	Primario	1992	1274	Mixta	SRH-T1 y SRH-T2	Si
(Kimura et al., 2012)	Japón	2012	Salud Bucal	Primario	173	138	Mixta	OHRQoL	Si
(Neuman et al., 2012)	Estados Unidos	2012	Oncología	Primario	60	60	Mixta	EORTC -30; EORTC-38 y SQOLS8	Si

Tabla 11. Artículos incluidos en la revisión sistemática: Enero 2010 – Julio 2020 (2/6)

Autor	País	Año	Especialidad del estudio	Tipo de datos	Muestra (Inicio)	Muestra (Final)	Tipo de población	Instrumento	Response Shift
(Reissmann et al., 2012)	Alemania	2012	Ortodoncia	Primario	126	126	Mixta	OHIP-G y OHIP-G49	Si
(X.-H. Zhang et al., 2012)	Singapur	2012	Ortopedia	Primario	74	74	Mixta	SF-36 / EuroQoL	Si
(C. E. Schwartz & Rapkin, 2012)	Estados Unidos	2012	Inmunología	Primario	521	521	Mixta	QOL Appraisal Profile/Rand-36 /General Health Then-Test	No
(Visser et al., 2013)	Países Bajos	2013	Oncología	Primario	202	202	Mixta	SF-36	Si
(Arons et al., 2013)	Países Bajos	2013	Psiquiatría	Primario	220	151	Mixta	VAS	Si
(Boucekine et al., 2013)	Francia	2013	Neurología	Secundarios	580	524	Mixta	MusiQoL y SF-36	Si
(Dabakuyo et al., 2013)	Francia	2013	Oncología	Primario	381	381	Femenino	QLQ-C30, BR23 y EurQOL-EQ-5D	Si
(Fokkema et al., 2013)	Países Bajos	2013	Psiquiatría	Primario	560	250	Mixta	BDI	Si
(P. Gandhi et al., 2013)	Estados Unidos	2013	Cardiología	Secundarios	2317	909	Mixta	SF-36	Si
(P. K. Gandhi et al., 2013)	Estados Unidos	2013	Cardiología	Secundarios	2317	788	Mixta	SF-36	Si
(Lix et al., 2013)	Canadá	2013	Gastrointestinal	Primario	388	388	Mixta	EII (IBDQ) y SF-36	No
(C. E. Schwartz et al., 2013)	Estados Unidos	2013	Ortopedia	Primario	178	169	Mixta	SF-36 /ODI	Si
(Edelaar-Peeters & Stiggelbout, 2013)	Países Bajos	2013	Neurología	Primario	44	44	Mixta	VAS	Si
(Flynn et al., 2013)	Australia	2013	Gerontología	Primario	478	315	Mixta	ICECAP-O	No
(Ahmed et al., 2014)	Canadá	2014	Enfermedades crónicas	Secundarios	776	776	Mixta	SF-36	No

Tabla 11. Artículos incluidos en la revisión sistemática: Enero 2010 – Julio 2020 (3/6)

Autor	País	Año	Especialidad del estudio	Tipo de datos	Muestra (Inicio)	Muestra (Final)	Tipo de población	Instrumento	Response Shift
(Anota et al., 2014)	Francia	2014	Oncología	Secundarios	381	381	Femenino	EORT QLQ-C30/EORT QLQ-BR23	Si
(Barclay-Goddard & Tate, 2014)	Canadá	2014	Cardiología	Secundarios	3983	745	Masculino	SF-36	Si
(Brinksma et al., 2014)	Países Bajos	2014	Oncología	Primario	61	37	Mixta	PedsQL / Cantril ladder	Si
(DeConde et al., 2014)	Estados Unidos	2014	Otolaringología	Primario	514	514	Mixta	SNOT-22	Si
(Hamidou et al., 2014)	Francia	2014	Oncología	Primario	381	381	Femenino	EORTC-QLQ-30 Y BR-23	Si
(J. S. Howard et al., 2014)	Estados Unidos	2014	Ortopedia	Primario	56	48	Mixta	SF-36	No
(Sajobi et al., 2014)	Canadá	2014	Neurología	Primario	80	80	Mixta	QOLIE-31	Si
(C. E. Schwartz et al., 2014)	Estados Unidos	2014	Neurología	Secundarios	859	859	Mixta	NARCOMS/SF-12	Si
(R. M. ten Ham et al., 2014)	Países Bajos	2014	Oncología	Secundarios	1720	1720	Masculino	SF-36	Si
(Elliott et al., 2014)	Estados Unidos	2014	Nefrología	Primario	31	27	Mixta	-	Si
(Von Blanckenburg et al., 2014)	Alemania	2014	Oncología	Primario	480	86	Mixta	WHOQOL-BREF	Si
(Guilleux et al., 2015)	Francia	2015	Diferentes diagnósticos	Secundarios	669	537	Mixta	SF-36	Si
(Liu & Davis, 2015)	Estados Unidos	2015	Otolaringología	Primario	79	32	Mixta	SNOT-20	Si
(Mayo et al., 2015)	Canadá	2015	Medicina Interna	Secundarios	388	388	Mixta	SF-36	No
(Mollerup & Johansen, 2015)	Dinamarca	2015	Dermatología	Primario	306	306	Mixta	VASnow / VASworst	No
(Rutgers et al., 2015a)	Países Bajos	2015	Ortopedia	Primario	74	20	Mixta	VAS	Si

Tabla 11. Artículos incluidos en la revisión sistemática: Enero 2010 – Julio 2020 (4/6)

Autor	País	Año	Especialidad del estudio	Tipo de datos	Muestra (Inicio)	Muestra (Final)	Tipo de población	Instrumento	Response Shift
(Sajobi et al., 2015)	Canadá	2015	Neurología	Secundarios	409	409	Mixta	SF-36	Si
(Traa et al., 2015)	Países Bajos	2015	Oncología	Primario	672	205	Mixta	WHOQOL-Bref	No
(Verdam et al., 2015)	Países Bajos	2015	Oncología	Primario	1157	1157	Mixta	EQ-5D/ EORTC-QLQ-C30 / RSCL	Si
(Phillips et al., 2015)	Singapur	2015	Oncología	Primario	167	167	Mixta	EORTC QLQ-C30	No
(C. Schwartz et al., 2015)	Canadá	2015	Enfermedades crónicas	Primario	172	172	Mixta	SF-36y VAS	No
(Blanchin et al., 2016)	Francia	2016	Enfermedades crónicas	Primario	669	499	Mixta	SF-36	Si
(P. K. Gandhi et al., 2016)	Estados Unidos	2016	Otolaringología	Primario	238	238	Mixta	PAQLQ	No
(Gerlich et al., 2016)	Alemania	2016	Oncología	Primario	522	402	Masculino	EORTC QLQ-C30	Si
(Hollman et al., 2016)	Países Bajos	2016	Ortopedia	Primario	36	36	Mixta	WORC /EuroQol-5D-3L	No
(Lix et al., 2016)	Canadá	2016	Medicina Interna	Primario	606	388	Mixta	SF-36	Si
(Nolte et al., 2016)	Alemania	2016	Psiquiatría	Secundarios	1188	1188	Mixta	ICD-10	Si
(Ousmen et al., 2016b)	Francia	2016	Oncología	Secundarios	381	381	Femenino	EORTC-QLQ-C30/EORTC-QLQ-BR23	Si
(Reissmann et al., 2016)	Alemania	2016	Odontología	Secundarios	554	554	Mixta	OHIP	No
(Sébille et al., 2016)	Francia	2016	Medicina Interna	Primario	390	390	Mixta	PreKit-QOL	No
(Taminiau-Bloem et al., 2016)	Países Bajos	2016	Oncología	Primario	50	24	Mixta	EORTC QLQ-C30	No
(Verdam et al., 2016)	Países Bajos	2016	Oncología	Secundarios	485	437	Mixta	SF-36	Si
(P. Wu, 2016)	Taiwán	2016	Psicología	Primario	320	320	Mixta	BDII	Si
(Arthur et al., 2016)	Reino Unido	2016	Audiología	Primario	16	16	Mixta	GHABP	Si

Tabla 11. Artículos incluidos en la revisión sistemática: Enero 2010 – Julio 2020 (5/6)

Autor	País	Año	Especialidad del estudio	Tipo de datos	Muestra (Inicio)	Muestra (Final)	Tipo de población	Instrumento	Response Shift
(Jakola et al., 2017)	Suecia	2017	Neurología	Primario	210	73	Mixta	EQ-5D-3L	No
(Chen et al., 2017)	Estados Unidos	2017	Salud Mental	Primario	114	114	Mixta	ISI and PSQI	No
(B. Hosseini et al., 2017)	Francia	2017	Medicina Interna	Primario	264	171	Mixta	EQ-5D-3L y SF-6D	No
(Jabrayilov et al., 2017)	Países Bajos	2017	Psiquiatría	Secundarios	540	450	Mixta	OQ-45	No
(Machuca et al., 2017)	Alemania	2017	Odontología	Primario	75	75	Mixta	DHEQ /CRT	Si
(Powell et al., 2017)	Canadá	2017	Salud Mental	Secundarios	2148	2148	Mixta	QOLI-20	No
(Sajobi et al., 2017)	Canadá	2017	Pediatría	Primario	373	373	Mixta	QOLCE-55	Si
(Salmon et al., 2017a)	Francia	2017	Oncología	Secundarios	466	466	Femenino	MFI-20 / QLC-C30	Si
(C. E. Schwartz et al., 2017)	Estados Unidos	2017	Neurología	Secundarios	858	858	Mixta	Rand-12	Si
(Hinz et al., 2017)	Estados Unidos	2017	Ortopedia	Secundarios	90450	90450	Mixta	EQ-5D3	No lo señala
(Spuling et al., 2017)	Alemania	2017	Gerontología	Primario	3854	1764	Mixta	SRH-T1 y SRH-T2	Si
(Tessier et al., 2017)	Francia	2017	Oncología	Secundarios	215	215	Femenino	SWLS y EORTC QLQ-C30	Si
(Verdam et al., 2017a)	Países Bajos	2017	Oncología	Primario	170	170	Mixta	SF-36	Si
(Yang et al., 2017)	Estados Unidos	2017	Neurología	Primario	124	124	Mixta	EQ-5D/UPDRS	No
(B. Hosseini et al., 2017)	Irán	2017	Oncología	Primario	211	211	Mixta	EORTC QLQ-C30	Si
(Aburub et al., 2018)	Canadá	2018	Oncología	Primario	192	97	Mixta	EQ-5D	Si
(Friedrich et al., 2018)	Alemania	2018	Oncología	Primario	350	308	Femenino	EORTC QLQ-C30	Si
(Krägeloh et al., 2018)	Nueva Zelandia	2018	Psicología	Primario	181	181	Mixta	CHIME	Si

Tabla 11. Artículos incluidos en la revisión sistemática: Enero 2010 – Julio 2020 (6/6)

Autor	País	Año	Especialidad del estudio	Tipo de datos	Muestra (Inicio)	Muestra (Final)	Tipo de población	Instrumento	Response Shift
(Murata et al., 2018)	Japón	2018	Oncología	Primario	-	-	Femenino	EORTC QLQ-C30	Si
(Nichols et al., 2018)	Nueva Zelandia	2018	Ortodoncia	Primario	57	57	Mixta	OHIP-14	Si
(Reissmann et al., 2018b)	Alemania	2018	Odontología	Primario	126	126	Mixta	OHIP -49	Si
(Schmidt et al., 2018)	Alemania	2018	Oncología	Secundarios	518	518	Mixta	EORTC QLC-30	No
(C. E. Schwartz et al., 2018)	Estados Unidos	2018	Neurología	Secundarios	2104	2104	Mixta	SF-36	Si
(Wang et al., 2018)	China	2018	Gerontología	Primario	238	238	Mixta	SF-12	Si
(Carlier et al., 2019)	Países Bajos	2019	Psiquiatría	Primario	206	206	Mixta	SQ-48	Si
(Felix et al., 2019)	Alemania	2019	Ortopedia	Primario	203	137	Mixta	EQ-5D-3L	Si
(Friedrich et al., 2018)	Alemania	2019	Cardiología	Primario	479	282	Mixta	EORTC-QLQ-C30	Si
(Powden et al., 2019)	Estados Unidos	2019	Ortopedia	Primario	20	20	Mixta	FAAM-Sport	No
(Preiß et al., 2019b)	Alemania	2019	Oncología	Primario	197	197	Masculino	EORTC QLQ-C30 y PHQ-4	Si
(Rohn et al., 2019)	Italia	2019	Oncología	No especificado	258	258	Mixta	EORTC QLQ-C30	No
(Haagsma et al., 2020)	Países Bajos	2020	Traumatología	Primario	550	550	Mixta	EQ-5D-3L y EQ-VAS	Si
(Machuca et al., 2020)	Reino Unido	2020	Ortodoncia	Primario	145	145	Mixta	OHIP-Edent	Si
(C. E. Schwartz et al., 2020)	Estados Unidos	2020	Medicina General	Secundarios	1481	1481	Mixta	PROMIS-10	No
(R. M. T. ten Ham et al., 2020)	Estados Unidos	2020	Oncología	Primario	3161	1677	Masculino	SF-36 y UCLA (PCI)	Si
(Tew et al., 2020)	Australia	2020	Ortopedia	Secundarios	1892	1553	Mixta	SF-12	No
(Topp et al., 2020)	Estados Unidos	2020	Neurología	Primario	100	93	Mixta	SF-12	No
(Artavia-Mora et al., 2020)	Países Bajos	2020	Virología	Secundarios	3625	3625	Mixta	CD4	No
(Murata et al., 2020b)	Japón	2020	Oncología	Secundarios	368	368	Femenino	EORTC QLQ-C30	Si

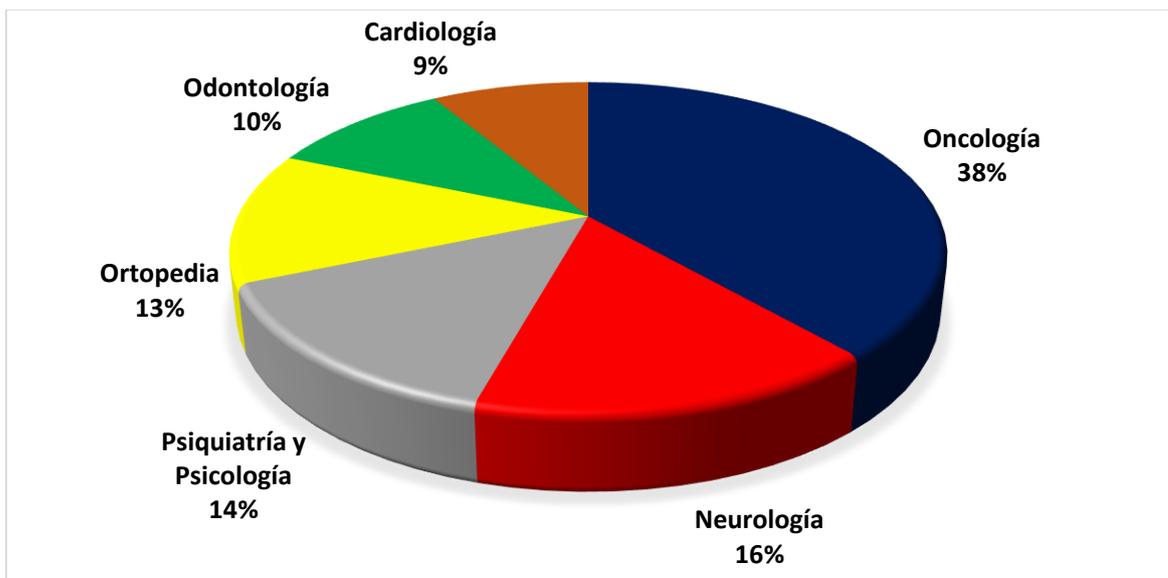
Tabla 12. Características de los artículos incluidos en la revisión (n=107)

Características	Total números de artículos (n=107)	Porcentaje (%)
Revistas		
Quality of Life Research	34	31.8
Health Quality Life Outcomes	10	9.3
European Journal Cancer Care	4	3.7
Journal Clinical Epidemiology	4	3.7
Archives physical medicine rehabilitation	2	1.9
BMC Medical Research Methodology	2	1.9
Social Science Medicine	2	1.9
Otras Revistas	49	45.8
Tamaño de la muestra		
Menos de 200	35	32.7
200 y más	69	64.5
No especificadas	3	2.8
País de la investigación		
Estados Unidos	23	21.5
Países Bajos	17	15.9
Alemania	15	14.0
Canadá	14	13.1
Francia	11	10.3
Japón	3	2.8
Holanda	4	3.7
Reino Unido	3	2.8
Australia	3	2.8
Nueva Zelandia	2	1.9
Singapur	2	1.9
Taiwán	2	1.9
Otros países	8	7.5
Tipo de estudios		
Primarios	74	69.2
Secundarios	31	28.9
No especificado	2	1.9

El análisis de los artículos incluidos permitió examinar el tipo de dato utilizado y se determinó que el 69.2% fueron estudios con datos primarios, 28.9% datos secundarios y 1.9% no fue especificado. Además, se pudo identificar que el 32.7% de las investigaciones utilizaron muestras menores de 200 pacientes, situación que influye en el tipo de método utilizado para la detección del *Response Shift* (Tabla 12).

La Figura 8 muestra las especialidades médicas que impulsaron la mayoría de los estudios fueron Oncología (27), Neurología (11), Ortopedia (9), Psicología/Psiquiatría (10), Odontología (7) y Cardiología (6).

Figura 8. Especialidades evaluadas en los diferentes artículos seleccionados



Diversos cuestionarios fueron utilizados para capturar los datos de pacientes. Algunos estudios se concentraron en aplicar un solo instrumento, mientras que otros incluyeron diferentes mecanismos de recolección de datos. La Tabla 13 describe el número de los instrumentos más utilizados y su codificación usual.

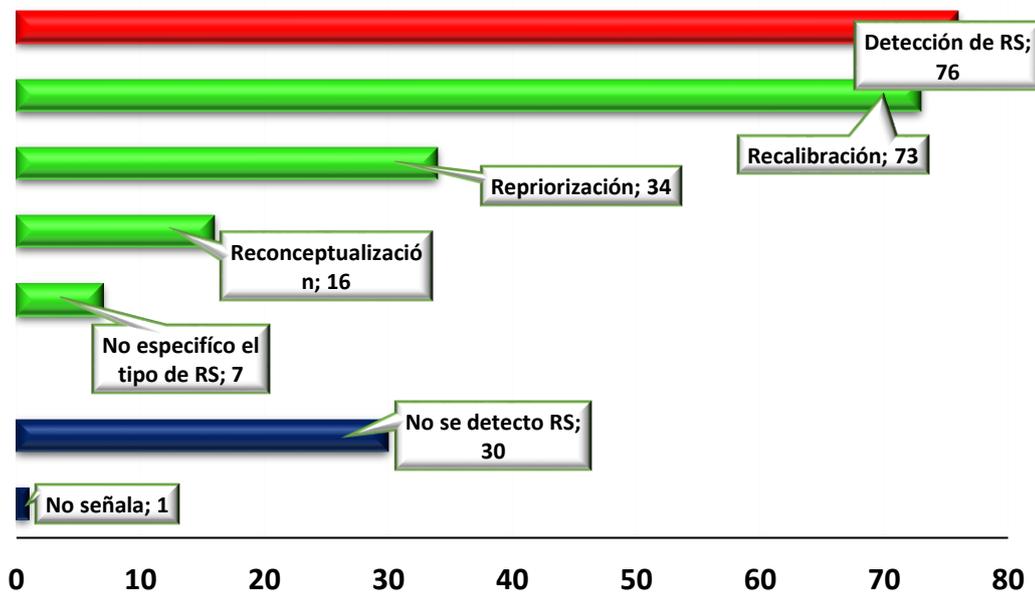
Tabla 13. Instrumentos más utilizados en los estudios evaluados

Instrumentos Usados	Código	Número
European Organization for Research and Treatment of Cancer Quality-of-Life-Questionnaire Core 30	EORTC QLQ-C30	20
36-Item Short Form Survey	SF-36	22
12-Item Short Form Survey	SF-12	7
EuroQol – 5 Dimensions	EQ-5D	10
Oral Health Impact Profile	OHIP	6

2.3.6. Sobre los métodos de detección de *Response Shift*

Del total de 107 artículos, en 76 (71.0%) se detectaron la existencia de *Response Shift*, en 30 (28.0%) no lograron identificarlo, mientras que en un solo artículo no se señala (Figura 9).

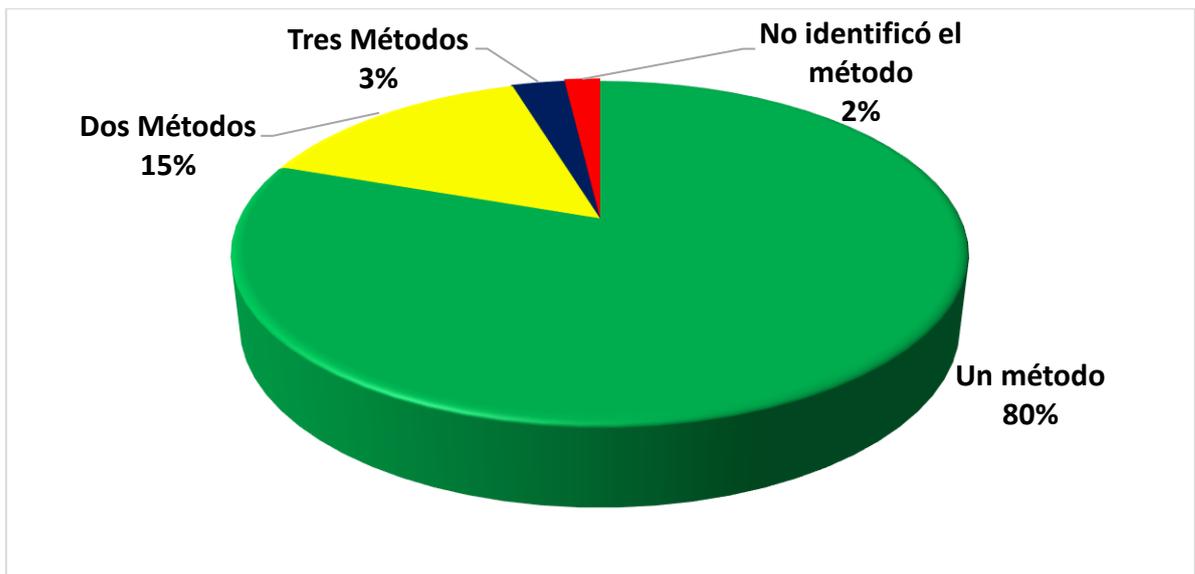
Figura 9. Número de artículos (n=107), detección y tipo de detección de *Response Shift*



Durante esta revisión sistemática, se identificaron los métodos utilizados para la detección de *Response Shift*, los tipos de *Response Shift*, y si los estudios detectaron o no su existencia. En varios estudios, se encontró evidencia de la utilización de más

de un método. En 86 (80.4%) artículos solamente utilizaron un método para la detección de *Response Shift*, mientras que 19 (17.8.0%) utilizaron dos o más métodos, y dos artículos no especificaron el método utilizado, tal como muestra la Figura 10.

Figura 10. Artículos según número de Métodos utilizados para detectar *Response Shift*



La Tabla 14 describe el número de veces que se utilizaron los métodos en las publicaciones. Un total de 41 (31.8%) artículos utilizaron el método *Then-Test*, 35 (27.1%) adoptaron los Modelos de Ecuaciones Estructurales de Oort y 2 (1.5%) artículos señalaron el uso de Modelos de Ecuaciones Estructurales de Schmidt. Otros modelos utilizados fueron la Regresión Lineal Múltiple (7), Regresión de Efectos Mixtos (6), Análisis de Trayectoria Latente (6), Teoría de Respuesta al Ítem (6), Regresión Logística (5), Regresión y Clasificación de Árbol (4) y los Métodos de Importancia Relativa (4). Se utilizaron otros métodos, pero en un número menor de estudios.

Tabla 14. Número de veces que fueron utilizados los métodos para la detección de *Response Shift*

Método	Total (n=129)	Detección de <i>Response Shift</i>		
		Si	No	No identificado
<i>Then-Test</i>	41	32	9	0
Modelo de Ecuaciones Estructurales (Oort)	35	26	9	0
Modelo de Ecuaciones Estructurales (Schmidt)	2	2	0	0
Regresión Lineal Múltiple	7	4	2	1
Método de Importancia Relativa	4	3	1	0
Regresión de Efectos Mixtos	6	5	1	0
Regresión y Clasificación de Árbol	4	4	0	0
Regresión de Bosque Aleatorio	2	2	0	0
Modelo de Regresión Logística	5	1	4	0
Teoría de Respuesta al Ítem	6	3	3	0
Análisis de Trayectoria Latente	6	2	4	0
Otros Métodos	11	7	3	1

Según los métodos de detección utilizados en todos los artículos, que en ocasiones se aplicaron varios métodos en una misma publicación, 91 (70.5%) detectaron *Response Shift*, mientras que 36 (27.9%) indicaron su ausencia. En la revisión, 2 (1.6%) artículos no identificaron la presencia o ausencia de este fenómeno. De los estudios que, si detectaron *Response Shift* mediante diferentes métodos, 73 indicaron la presencia de recalibración, 34 indicaron la existencia de repriorización, 16 destacaron la reconceptualización, y 7 artículos detectaron la existencia de *Response Shift*, pero no especificaron el tipo, tal como se detalla en la Tabla 15.

Estas proporciones se obtienen debido a que 41 artículos detectaron un solo tipo de *Response Shift*, 20 artículos registraron dos tipos de *Response Shift*, solamente 7 estudios lograron identificar los tres tipos simultáneamente. Siete artículos no indicaron que tipo *Response Shift* detectaron.

De los artículos incluidos ($n = 107$), el método tradicional para la detección de *Response Shift* *Then-Test* registró el mayor número de *Response Shift* del tipo recalibración (27), mientras que los Modelos de Ecuaciones Estructurales de Oort registraron similar número de estudios detectando recalibración (24), pero el mayor registro de repriorización (13) y reconceptualización (7) de toda la revisión sistemática

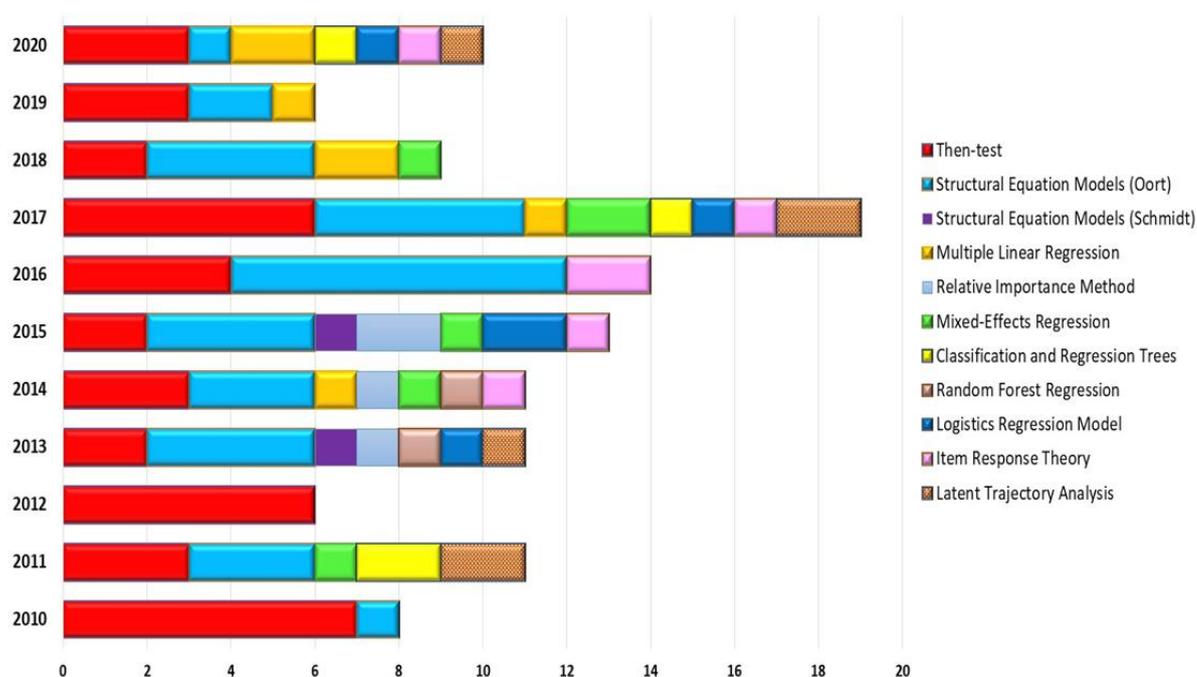
(Tabla 15). Los Modelos de Ecuaciones Estructurales de Schmidt detectaron cambios solamente en los estándares internos en dos (2) artículos. El resto de los métodos también demostraron ser utilizados para la detección de los tres tipos de *Response Shift*, con excepción de los Métodos de Importancia Relativa, Regresión de Efectos Mixtos, Regresión de Bosques Aleatorios y el Análisis de Trayectoria Latente que no identificaron la presencia de reconceptualización.

Tabla 15. Número de artículos según método de detección y tipo de *Response Shift*

Métodos	Tipo de Response Shift			
	Recalibración	Repriorización	Reconcep- tualización	No Indicado
Then-Test	27	6	3	3
Modelos de Ecuaciones Estructurales (Oort)	24	13	7	1
Modelos de Ecuaciones Estructurales (Schmidt)	2	0	0	0
Regresión Lineal Múltiple	2	0	1	1
Método de Importancia Relativa	0	3	0	0
Regresión de Efectos Mixtos	3	2	0	1
Regresión y Clasificación de Árbol	4	2	1	0
Regresión de Bosque Aleatorio	1	1	0	0
Modelo de Regresión Logística	1	1	1	0
Teoría de Respuesta al Ítem	3	3	1	0
Análisis de Trayectoria Latente	2	1	0	0
Otros Métodos	4	2	2	1

A pesar de la efectividad para la detección de *Response Shift*, esta revisión sistemática evidencia el creciente interés por explorar los diferentes métodos estadísticos para el estudio de *Response Shift* (Figura 11). Los métodos más frecuentemente utilizados fueron *Then-Test* y Modelos de Ecuaciones Estructurales de Oort entre los años 2010 y 2020. Sin embargo, durante los últimos años, la figura muestra que otras técnicas alternativas están siendo consideradas para la evaluación de este fenómeno.

Figura 11. Frecuencia de los métodos utilizados para la detección de *Response Shift* por año



2.3.7. Análisis crítico de la revisión sistemática

Luego de la revisión sistemática, el estudio examinaba las investigaciones realizadas en CVRS que buscaban, mediante diferentes métodos, la detección de *Response Shift*. Esta revisión describe cómo se está evaluando en los últimos años la detección de este fenómeno, qué métodos son los más utilizados, y qué tipo de *Response Shift* es el más identificado. Para médicos, profesores e investigadores, esta metodología es fuente útil de evidencias y actualización (Manchikanti et al., 2009), proporcionan elementos para que quienes elaboran políticas clínicas puedan evaluar riesgos, beneficios y afectaciones al cuidado de la salud, así como nuevas iniciativas de investigación (Liberati et al., 2009; Moher et al., 2009).

El mayor número de estudios sobre CVRS y métodos para la detección de *Response Shift* se realizaron en los años 2014 y 2017. Un tercio de las publicaciones incluidas en la revisión están orientadas a los estudios del cáncer (Aburub et al., 2018; Brinksmas et al., 2014; Dabakuyo et al., 2013; Gerlich et al., 2016; Murata et al., 2020b; Neuman et al., 2012; Preiß et al., 2019a; Taminiu-Bloem et al., 2010; Tessier et al., 2017).

El método más utilizado fue *Then-Test* principalmente para datos primarios de los cuales se logró medir recalibración (Arthur et al., 2016; Dempster et al., 2010; Galenkamp et al., 2012; Haagsma et al., 2020; Reissmann et al., 2018b; Rutgers et al., 2015b; Shi et al., 2011). En menor proporción, otros estudios permitieron destacar que el *Then-Test* también puede identificar los otros tipos de *Response Shift*: repriorización y reconceptualización (Höfer et al., 2011; Machuca et al., 2017; Taminiau-Bloem et al., 2010). Similares resultados se mostraron en el estudio de Sajobi et al. (2018) que consistió en una revisión sistemática sobre métodos para detectar *Response Shift*, y donde el 54.5% de las publicaciones utilizaron *Then-Test* para este propósito. De igual forma, en Ilie et al., (2019b), un 60.0% de los estudios de cáncer utilizaron esta metodología, pero señala que su interpretación debe realizarse con precaución ya que es propensa a generar sesgos.

Schwartz (2010) indica que el *Then-Test* posee la ventaja de un fácil manejo y análisis, pero su desventaja yace en los errores aleatorios y/o confusión, y su dificultad de interpretación. Por tanto, este enfoque induce a la utilización de métodos individualizados con mayor rigor estadísticos.

El método de Modelos de Ecuaciones Estructurales de Oort, a diferencia del *Then-Test*, es utilizado en estudios basados en datos primarios y secundarios. Este método multivariante evidenció su capacidad de detectar cambios en todos los tipos de *Response Shift*: recalibración, repriorización y reconceptualización. Diversas investigaciones reiteran los Modelos de Ecuaciones Estructurales de Oort como un método eficaz para detectar los cambios en pacientes, indistintamente de los diferentes tipos de *Response Shift* o los datos utilizados (DeConde et al., 2014; Nakamura et al., 2013; Verdam et al., 2017b; Wang et al., 2018). Esta capacidad de detección puede ser la razón por su mayor interés de uso por parte de investigadores.

El Modelo de Ecuaciones Estructurales de Schmidt fue utilizado sólo en dos estudios detectando únicamente recalibración. En estudios comparativos con el método de Oort Gandhi et al., (2013) sólo se identifica este tipo de *Response Shift*, pero concluyen que ambos enfoques utilizan parámetros diferentes para definir la recalibración: Schmidt mediante cambios en el factor de carga en el tiempo, mientras que Oort por medio de los cambios en intercepto.

Otros métodos utilizados lograron mayor registro de detección de *Response Shift*: Regresión de Efectos Mixtos y Lineal Múltiple (Felix et al., 2019; Lacey et al., 2011; Nichols et al., 2018; C. E. Schwartz et al., 2014), Árbol de Clasificación y Regresión (Li & Schwartz, 2011; Machuca et al., 2020), Teoría de Respuesta al Ítem (Anota et al., 2014; Blanchin et al., 2016) y Método de Importancia Relativa (Sajobi et al., 2014, 2015; C. E. Schwartz et al., 2013).

A pesar de una mayor frecuencia de detección de recalibración, llama la atención que los tres estudios que utilizaron el Método de Importancia Relativa lograron detectar repriorización (Sajobi et al., 2014, 2015; C. E. Schwartz et al., 2013). Los métodos de Regresión de Efectos Múltiples (Lacey et al., 2011; C. E. Schwartz et al., 2014), Árboles de Clasificación y Regresión (Li & Schwartz, 2011; Machuca et al., 2020) y Análisis de Trayectoria Latente (Salmon et al., 2017b) tampoco aportaron información relativa al tipo de *Response Shift* reconceptualización.

Una de las observaciones de este estudio es que el tamaño de la muestra está relacionado con el tipo de método que es utilizado para la detección del *Response Shift*. La utilización del *Then-Test* es más flexible y aplicable a muestra pequeñas o en estudios específicos, mientras que los métodos con un rigor estadístico más fuerte requieren de muestras grandes. De acuerdo con lo planteado por Schwartz et al. (2013), para el uso de métodos multivariantes avanzados que puedan detectar *Response Shift*, el análisis de datos debe incluir un número suficiente de individuos que participen en el estudio, así como del número de parámetros considerados: el modelo, el factor de carga, la relación de los ítems, distribución de los datos y los procedimientos para estimar parámetros.

Ante la necesidad de explorar nuevas metodologías multivariantes para analizar este fenómeno, surge una propuesta alternativa para los estudios de CVRS y la detección de *Response Shift* mediante datos de tres vías utilizando STATIS Dual (Escoufier, 1976b; H. L'Hermier Des Plantes, 1976). La propuesta presentada por Vicente-Galindo (2003) se centra en un procedimiento donde integra los métodos STATIS Dual y la comparación de subespacios de Krzanowski (1979) en los cuales se examina las estructuras factoriales de un conjunto de datos múltiples para identificar la presencia de cambios.

Las revisiones sistemáticas evaluadas en este estudio (Hinds et al., 2018; Ilie et al., 2019a; Nielsen et al., 2017; Powden et al., 2018; Scholten et al., 2017) confirman un continuo uso de los mismos métodos para detectar *Response Shift* y sus diferentes tipos, destacando cada una de sus ventajas y limitaciones. La propuesta anterior es un ejemplo de las oportunidades para continuar examinando otras metodologías estadísticas robustas que permitan abordar esta línea de investigación, y reducir los sesgos en los resultados y las ambigüedades en los análisis.

2.3.8. Consideraciones finales

La revisión sistemática ha demostrado ser una metodología adecuada para la localización y síntesis del nivel de conocimiento sobre el tema de investigación propuesto. Se ha constatado un interés generalizado por el tema de la CVRS y sobre *Response Shift* en diferentes grupos de diagnóstico. Con respecto a los estudios para detección de *Response Shift*, se observa un interés continuo y una consolidación de los métodos estadísticos, siendo los más utilizados *Then-Test*, Modelos de Ecuaciones Estructurales de Oort, Regresión Lineal Múltiple y de Efectos Mixtos, Teoría de Respuesta al Ítem y Análisis de Trayectoria Latente. Los Modelos de Ecuaciones Estructurales de Oort ha resultado ser el más versátil por su capacidad de detectar *Response Shift* y sus diferentes tipos.

Los resultados muestran que no todos los métodos logran detectar *Response Shift* en similares proporciones, la mayoría logró identificar recalibración, y en menor proporción repriorización y en pocos casos reconceptualización. Sin embargo, nuestro estudio pone de manifiesto la necesidad de explorar otros métodos que permitan igual capacidad de detección, con metodologías más robustas y alta precisión. Es recomendable continuar los estudios sobre el fenómeno de *Response Shift*, su detección y tipos desde un abordaje multidisciplinar.

CAPÍTULO III

MÉTODOS ESTADÍSTICOS PARA LA DETECCIÓN DE *RESPONSE SHIFT*

3. MÉTODOS ESTADÍSTICOS PARA LA DETECCIÓN DEL *RESPONSE SHIFT*

3.1. Características de los estudios de calidad de vida relacionado con la salud

Los estudios sobre calidad de vida relacionada con la salud (CVRS) requieren ser sometidos a un proceso de medición cuyas metodologías y técnicas para el manejo de los datos buscan examinar el comportamiento de un individuo o paciente dentro del contexto clínico, frente a un padecimiento y del cual es imprescindible entender las variaciones experimentadas a lo largo de la condición médica. El propósito de esta medición es entender la evolución de la persona estudiada y tomar una decisión pertinente sobre el tratamiento asignado para su mejoramiento.

Dado que es fundamental analizar las condiciones de CVRS y detectar la existencia de cambios, es necesario el establecimiento de una metodología que permita identificar el tipo de estudio sobre CVRS que se desea realizar y las técnicas estadísticas más adecuadas que serán aplicadas con el propósito de examinar los resultados obtenidos y su nivel de confiabilidad entre las premisas planteadas y la realidad observada.

Los estudios de CVRS requieren de la existencia de información obtenida luego de la captación de datos consecutivos sobre una condición clínica que sea investigada. Estos datos son el resultado de un mismo individuo examinado a lo largo del tiempo, en consecuencia, estos estudios son de carácter longitudinal lo cual implica más de dos mediciones a lo largo de un seguimiento. Los datos longitudinales son las mediciones repetidas en los sujetos a lo largo de un periodo de tiempo y que se pueden realizar para exposiciones dependientes del tiempo o resultados recurrentes. (Delgado & Llorca, 2004)

Este tipo de estudios longitudinales pueden ser de dos tipos: prospectivos o retrospectivos. Los estudios prospectivos son aquellos donde el inicio es anterior a los hechos estudiados, por tanto, los datos se recogen a medida que van sucediendo;

mientras que los estudios retrospectivos son aquellos donde el inicio del estudio es posterior a los hechos estudiados y los datos se recogen de archivos o entrevistas sobre hechos sucedidos en el pasado. (Supo, 2012)

Nunes (2009) señala que, a largo del tiempo, los estudios que utilizaban datos consecutivos o medidas repetidas se restringían a estudios experimentales. Pero con el registro progresivo de información histórica sobre fenómenos similares, se desarrollaron otro tipo de estudios ampliamente utilizados por investigadores del tema ya que se incorporan metodologías y técnicas estadísticas más rigurosas. Estos son los estudios longitudinales. Los datos de CVRS son datos longitudinales debido a su recopilación y documentación de carácter cronológica la cual posee características muy particulares. Dada esta naturaleza, su tratamiento estadístico se sustenta en el análisis de datos.

En contextos de ensayos clínicos donde se pretende evaluar el impacto de terapias, generalmente se realiza una evaluación longitudinal de la CVRS durante el período de tratamiento, en los cuales se puede identificar condiciones de cambio. Con miras de lograr los datos para el estudio, se solicita a los pacientes que contesten un cuestionario validado en varios momentos a lo largo del periodo de tratamiento programado y después de finalizado el tratamiento. La aplicación del cuestionario es generalmente determinada con ciclos dentro del tratamiento o con visitas programadas en periodos posteriores. La evaluación de la CVRS termina en el instante previo detallado en el protocolo o cuando la enfermedad obliga a que el paciente sea sometido a otro tratamiento no establecido en el protocolo y que este tratamiento pueda confundir el efecto de la terapia original en la calidad de vida. Por esta razón, los datos de CVRS generan diversos problemas complejos, propios de la obtención del conjunto de datos, sino también de su análisis (Nunes, 2009, pp.113-114).

Schwartz et al. (2015) señala que el manejo de datos longitudinales no está exento de generar estimaciones que incluyan sesgos y pierdan confiabilidad. Estos sesgos se producen principalmente por la falta de datos que se originan por el desgaste del encuestado o por ítems que no son respondidos. El estudio reconoce el impacto de la pérdida de datos en la detección del cambio y que esta afectación es relativamente desconocida hasta la fecha. Para detectar el impacto en los resultados por la pérdida

de datos, los autores utilizaron diferentes metodologías estadísticas para examinar de forma longitudinal la CVRS y, al contrastar los resultados, logran evidenciar la sensibilidad que poseen los estudios de *Response Shift* cuando se presentan datos faltantes. Esta es una importante crítica por parte de los investigadores sobre la utilización de los datos longitudinales en la identificación de *Response Shift* en CVRS y sus posibles sesgos por la falta de datos.

A menudo los efectos de *Response Shift* son bajos y la falla en su detección puede darse por los datos faltantes debido a la deserción o la falta de respuesta de los ítems. Ya que los datos faltantes se encuentran frecuentemente en datos longitudinales de CVRS, se deben considerar estrategias efectivas para el manejo de esta pérdida (Sajobi et al., 2015).

3.2. Métodos tradicionales para el análisis de calidad de vida relacionada con la salud y detectar *Response Shift*

El análisis de la CVRS descansa en la recolección de datos derivada de las revisiones periódicas y repetidas a pacientes que sufren algún padecimiento de carácter médico a lo largo de un intervalo de tiempo y cuyo propósito es identificar los cambios en las respuestas de estos individuos con el fin de evaluar los efectos del tratamiento adoptado.

Las primeras metodologías utilizadas para analizar los datos de CVRS se sustentaron en los estadísticos descriptivos cruzados, en diferentes intervalos de tiempo, con el objetivo de describir la tendencia global de los diferentes grupos de tratamiento o sintetizar datos longitudinales en un estadístico resumen antes de efectuar comparaciones entre grupos de tratamientos. Estos métodos no consideran las variaciones intrínsecas de cada paciente y no captan cambios importantes en la CVRS a lo largo del tiempo, incrementando los errores Tipo I es decir se establece un planteamiento como una afirmación significativa cuando en realidad no lo es.

Otro método basado en el uso de estadísticos reduce considerablemente el número de contrastes. El procedimiento permite controlar la tasa de error Tipo I. Una forma de reducir la dimensionalidad de los datos es reducir un conjunto de n medidas individuales a un valor simple. Otra vía será a través de una media ponderada de las puntuaciones estandarizadas, donde las ponderaciones están basadas en la correlación entre las medidas.

Entre las técnicas estadísticas inicialmente utilizadas se encuentran el Análisis Univariante de Varianza (ANOVA) y el Análisis Multivariante de Varianza (MANOVA) con medidas repetidas. Estos tipos de análisis se basan en el cálculo de la suma de las puntuaciones atribuidas por los pacientes en cada instante en los ítems que miden cada variable latente, asumiendo una distribución normal. Sin embargo, estas técnicas presentan algunas debilidades dado a la correlación producida entre medidas repetidas que afectan los errores estándares y ampliando los errores Tipo I, así como la consideración unidimensional del modelo ANOVA, toda vez que la CVRS es de naturaleza dinámica y multidimensional. Con el propósito de reducir los errores de los modelos anteriores, se han aplicado otras técnicas como el Análisis de Covarianza (ANCOVA), el Análisis Multivariante de la Covarianza (MANCOVA) y la Regresión Múltiple.

3.2.1. *Then- Test*: Técnica para detectar *Response Shift* en CVRS

Así como en otros contextos, la técnica más ampliamente utilizada es el *Then-Test* descrito por [Howard y Dailey \(1979\)](#) donde expone la necesidad de recalibrar la escala de medición de la calidad de vida. El *Then-Test* consiste en solicitar a los pacientes, por medio de instrumentos validados, que luego de someterse a un tratamiento procedan a un juicio retrospectivo de su estado de salud inicial. El principal objetivo de este método es detectar los cambios en los estándares internos y así como la conceptualización y cambios en los valores.

La metodología de aplicación del método *Then-Test* se refiere a medidas realizadas en el momento de aplicar el *Post-Test* pero referidas al momento del *Pre-Test*, es decir, que se le solicita al encuestado que realice una evaluación sobre cómo se sentía al momento del *Pre-Test*. La idea en la que se basa es la siguiente: si se ha producido algún tipo de calibración en la escala de medida entre el *Pre-Test* y el *Post-Test* ambos valores no serían comparables, pero al ser evaluadas las dos experiencias en el momento del *Post-Test*, la recalibración afectará por igual a ambas evaluaciones. Así se pretende asegurar que los estándares de medida para las dos evaluaciones son los mismos y evitar que el cambio Beta sesgue al verdadero cambio Alfa. Si aparecen diferencias entre el *Pre-Test* y el *Then-Test* se asume que se ha producido un cambio Beta. La diferencia entre los valores al realizar el *Then-Test* y el *Post-Test* se acepta como el verdadero cambio Alfa (Galindo-Villardón & Vicente-Galindo, 2004).

Dicho de otra manera, para evaluar los efectos en la calidad de vida de un paciente luego de un tratamiento, se realizan las tres pruebas: *Pre-Test*, *Then-Test* y *Post-Test*. La primera se aplica al inicio del tratamiento, mientras que las otras dos se capturan simultáneamente en un momento posterior del mismo. El *Pre-Test* captura la condición reportada por el paciente antes de iniciar el tratamiento.

En un momento posterior, el *Then-Test* cuestiona sobre esa condición del paciente acerca de antes de iniciar el tratamiento. Como esta prueba realiza preguntas sobre el pasado, las respuestas pueden ser iguales o diferentes entre el *Then-Test* y el *Pre-Test*. Si las respuestas son iguales, se manifiesta una consistencia en los resultados sobre la condición estudiada; pero si son diferentes, se manifiesta un cambio en la calibración de los estándares propios del individuo: esto se refiere a la existencia de un cambio Beta.

Paralelamente al *Then-Test* se aplica el *Post-Test*, el cual captura la percepción del paciente sobre la condición posterior a su tratamiento. Como ambas pruebas se aplicaron simultáneamente, no deben existir cambios en la escala de medición por parte del paciente. En consecuencia, no debe evidenciarse cambios Beta. Bajo esta premisa, cualquier cambio evidenciado entre el *Then-Test* y el *Post-Test* representa el verdadero cambio natural del paciente: esto se refiere a la existencia de un cambio Alfa. La Figura 12 muestra gráficamente las relaciones entre cada una de estas pruebas.

4	(2010)	Kvam et al.	Minimal important differences and response shift in health-related quality of life; a longitudinal study in patients with multiple myeloma
5	(2010)	McPhail & Haines	Response shift, recall bias and their effect on measuring change in health-related quality of life amongst older hospital patients
6	(2010)	Razmjou et al.	The Impact of Response Shift on Perceived Disability Two Years Following Rotator Cuff Surgery
7	(2010)	Taminiau-Bloem et al.	A 'short walk' is longer before radiotherapy than afterwards: a qualitative study questioning the baseline and follow-up design
8	(2011)	Hinz et al.	Response shift in the assessment of anxiety, depression, and perceived health in urologic cancer patients: an individual perspective
9	(2011)	Hon-Yi et al.	Response shift effect on gastrointestinal quality of life index after laparoscopic cholecystectomy
10	(2011)	Höfer et al.	Coronary Intervention Improves Disease Specific Health-Related Quality of Life but Not Individualised Quality of Life: A Potential Response Shift Effect?
11	(2012)	Galenkamp et al.	Estimates of prospective change in self-rated health in older people were biased owing to potential recalibration response shift
12	(2012)	Kimura et al.	Response shift in oral health-related quality of life measurement in patients with partial edentulism
13	(2012)	Neuman et al.	Rectal Cancer Patients' Quality of Life With a Temporary Stoma: Shifting Perspectives
14	(2012)	Reissmann et al.	Impact of response shift on the assessment of treatment effects using the Oral Health Impact Profile
15	(2012)	Xu-Hao et al.	An Exploratory Study of Response Shift in Health-Related Quality of Life and Utility Assessment Among Patients with Osteoarthritis Undergoing Total Knee Replacement Surgery in a Tertiary Hospital in Singapore
16	(2012)	Schwartz et al.	Understanding appraisal processes underlying the Then-Test: a mixed methods investigation
17	(2013)	Dabakuyo et al.	Response shift effects on measuring post-operative quality of life among breast cancer patients: a multicenter cohort study
18	(2013)	Edelaar-Peeters & Stiggelbout	Anticipated adaptation or scale recalibration?
19	(2014)	Brinksman et al.	Exploring the Response Shift Phenomenon in Childhood Patients With Cancer and Its Effect on Health-Related Quality of Life
20	(2014)	Hamidou et al.	Impact of Response Shift on Time to Deterioration in Quality of Life Scores in Breast Cancer Patients
21	(2014)	Howard et al.	Influence of response shift on early patient-reported outcomes following autologous chondrocyte implantation
22	(2015)	Liu & Davis	The significance of response shift in sinus surgery outcomes
23	(2015a)	Rutgers et al.	Osteoarthritis treatment using autologous conditioned serum after placebo

24	(2016)	Hollman et al.	Response shift of the Western Ontario Rotator Cuff index in patients undergoing arthroscopic rotator cuff repair
25	(2016a)	Ousmen et al.	Impact of the occurrence of a response shift on the determination of the minimal important difference in a health-related quality of life score over time
26	(2016)	Taminiau-Bloem et al.	Using a retrospective pretest instead of a conventional pretest is replacing biases: a qualitative study of cognitive processes underlying responses to Then-Test items
27	(2016)	Arthur et al.	An exploratory study identifying a possible response shift phenomena of the Glasgow hearing aid benefit profile
28	(2017)	Jakola et al.	Is there a response shift in generic health-related quality of life 6 months after glioma surgery?
29	(2017)	Hosseini et al.	Nephrectomy Complication Is a Risk Factor of Clinically Meaningful Decrease in Health Utility among Living Kidney Donors
30	(2017)	Machuca et al.	Using classification and regression tree modelling to investigate response shift patterns in dentine hypersensitivity
31	(2017)	Spulling et al.	Response shift in self-rated health after serious health events in old age
32	(2017)	Yang et al.	Response shift: The experience of disease progression in Parkinson disease
33	(2017)	Hosseini et al.	Response shift in quality of life assessment among cancer patients: A study from Iran
34	(2018)	Friedrich et al.	Response shift effects of quality of life assessments in breast cancer survivors
35	(2018a)	Reissmann et al.	Bias in retrospective assessment of perceived dental treatment effects when using the Oral Health Impact Profile
36	(2019)	Powden et al.	Response Shift After a 4-Week Multimodal Intervention for Chronic Ankle Instability
37	(2019b)	PreilB et al.	Response shift effects in the assessment of urologic cancer patients' quality of life
38	(2019)	Friedrich et al.	Response shift effects in patients' assessments of their quality of life after cardiac rehabilitation
39	(2020)	Haagsma et al.	Conventional and retrospective change in health-related quality of life of trauma patients: an explorative observational follow-up study
40	(2020)	Machuca et al.	How peoples' ratings of dental implant treatment change over time?
41	(2020)	ten Ham et al.	Understanding the Major Factors Affecting Response Shift Effects on Health-Related Quality of Life: What the Then-Test Measures in a Longitudinal Prostate Cancer Registry

Sin embargo, otros autores manifiestan sus críticas a esta metodología. [Visser et al. \(2000\)](#) y [Hill y Betz \(2005\)](#) señalan que existe la posibilidad de que los datos del *Then-Test* aparezcan contaminados por problemas de memoria y dificultades que afrontan los pacientes en recordar su situación anterior, particularmente cuando en situaciones

crónicas la memoria puede volverse poco precisa, por lo que la validez de la prueba puede ser cuestionable.

Galindo-Villardón y Vicente-Galindo (2004) señalan que, desde el punto de vista estadístico, la comparación puede existir, y por tanto ser trivial. Al utilizar en esta metodología una distribución *t - Student* para datos apareados, se resuelve el problema de la comparación tanto del *Pre-Test – Post-Test* como del *Then-Test – Post-Test*, considerando que lo más relevante es que los valores que se comparan no estén sesgados.

En resumen, el *Then-Test* se basa en un diseño de estudio *pre* y *post* para probar el cambio de respuesta, donde implica la inclusión de una o más medidas retrospectivas previa a la prueba para detectar *Response Shift* de recalibración. Su principal ventaja es que se puede estimar el *Response Shift* de recalibración, además de que puede usarse para probar múltiples hipótesis de *Response Shift* mientras se controle el error Tipo I.

Por otro lado, la desventaja que presenta el *Then-Test* es que no debe utilizarse para detectar cambios de repriorización o reconceptualización, además de que no debe ser utilizados para probar *Response Shift* en datos secundarios o análisis de datos longitudinales sin ningún diseño de prueba previa retrospectivas (Sajobi et al., 2018).

3.3. Métodos multivariantes para detectar *Response Shift* en calidad de vida relacionada con la salud

Con la incorporación de las metodologías multivariantes, se logra incorporar un amplio abanico de oportunidades con miras a detectar los cambios en la CVRS de un individuo a lo largo del tiempo, así como contrastarlo como parte de un conjunto grupal de pacientes en momentos diferentes de sus tratamientos clínicos. En esta sección, se realizará una revisión de varios autores que desarrollan enfoques metodológicos orientados hacia tratar el contexto de *Response Shift* desde la óptica de las variables y los individuos.

3.3.1. El método de Ahmavaara

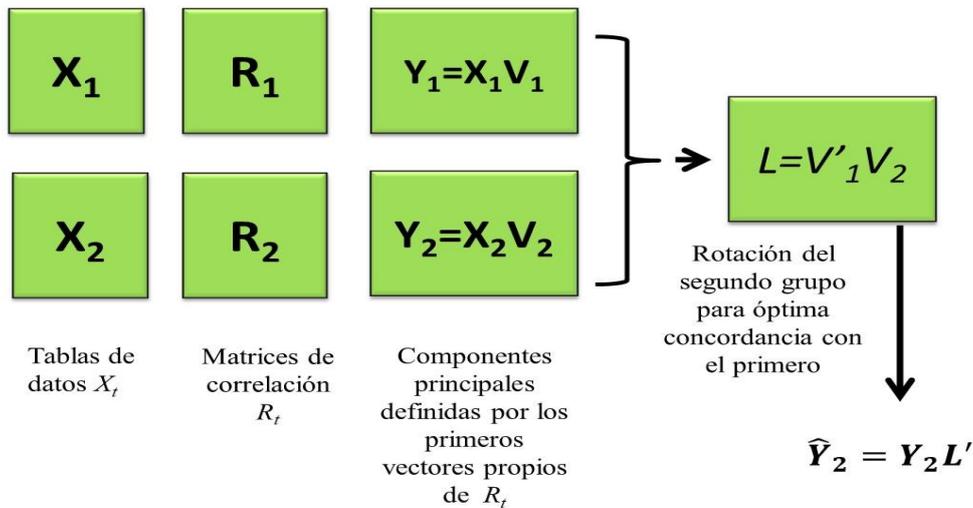
Uno de los primeros métodos utilizados para la comparación de estructuras factoriales fue propuesto por [Ahmavaara \(1954\)](#), considerado como el más sencillo, detallista y más elegante, y puede ser utilizado para las variables de una población estudiada que poseen iguales varianzas. Este es un método muy poco utilizado y raramente referido ([Hamilton, 1967](#)). Sin embargo, una aplicación inicial fue realizada por [Golembiewski et al. \(1976\)](#) con el objetivo de comparar evaluaciones de CVRS.

El objetivo de este método es la de transformar una de las matrices factoriales comparadas logrando que la estructura del primer grupo sea tan similar a la estructura del segundo grupo. Bajo este principio, la aplicación propuesta por Golembiewski supone que las dos matrices utilizadas X_1 y X_2 representarán las dos evaluaciones de calidad de vida de los pacientes comparables al *Pre-Test* y *Post-Test*.

Al comparar ambas evaluaciones se puede detectar la existencia de cambio. Para la aplicación del método, se desarrolla una matriz de transformaciones L , donde las columnas de la matriz están definidas por los factores comunes de X_2 , los cuales surgen en L en el mismo orden de X_2 . Por otra parte, las filas de la matriz L están definidas por los factores comunes de la X_1 , manteniendo el orden. Los elementos de la matriz L así definida, representan los cosenos de los ángulos entre los factores de ambos estudios. Si los resultados del análisis muestran que los elementos son exactamente iguales entre X_1 y X_2 , entonces los respectivos factores del *Pre-Test* y *Post-Test* serán idénticos.

Del producto entre las matrices X_1 y L se obtiene la matriz C , la cual representa el mejor ajuste mínimo cuadrático de X_1 sobre X_2 . Posteriormente se calcula los Coeficiente de Correlación Intraclase entre las matrices C y X_2 . Estos coeficientes indicarán el grado según las dos estructuras las cuales son similares tanto en patrón como en magnitud. El cuadrado de los coeficientes indica el porcentaje de varianza de las dos estructuras, tal como se muestra en el esquema de la Figura 13.

Figura 13. Esquema del método Ahmavaara



Aunque el método de Ahmavaara se presenta como un procedimiento útil para comparar dos soluciones factoriales, para los análisis de CVRS se manifiestan algunas limitaciones:

- Posibilita solamente la comparación de dos estructuras, lo que impide conseguir una visión general del problema cuando el objetivo es comparar estructuras en más de dos tiempos o situaciones,
- Se supone que las matrices en estudio se refieren a los mismos individuos o variables (ítems), lo que en la práctica no siempre sucede,
- La aplicación del procedimiento implica el cálculo de un elevado número de correlaciones intraclase, lo cual produce incrementos significativos en el riesgo de errores Tipo I.

3.3.2. El modelo de Zwinderman

Con el objetivo de cuantificar la magnitud del cambio en CVRS a lo largo del tiempo, [Zwinderman \(1990\)](#) propone un modelo denominado Regresión Logística Latente que plantea la calidad de vida dentro del contexto de salud como una variable latente multidimensional medida por medio de respuestas dicotómicas al mismo conjunto de ítems en diferentes instantes, y que el mismo incluye parámetros relacionados con el

proceso temporal y con los efectos de los tratamientos clínicos, teniendo en cuenta la interacción entre parámetros.

El planteamiento de Zwinderman considera un modelo de medida en el que X_v es la suma de las puntuaciones relativas al paciente v sobre todos los ítems relativos a una determinada variable latente, donde:

$$X_v = \tau_v + \varepsilon_v$$

Donde τ_v es el valor verdadero de una determinada variable latente del paciente v y ε_v un componente de error. Sin embargo, la modelización del cambio (δ_v) representa una función del tiempo y del tratamiento que el paciente v recibe:

$$\delta_v = q_v \mu + \tau$$

Donde τ es el cambio inherente al proceso temporal y μ el cambio debido al tratamiento; mientras que q_v es un parámetro de proporción donde toma el valor 1 si v recibe el tratamiento y 0 en caso contrario.

Zwinderman consideró la falta de datos en el conjunto longitudinal que se reporta en la evolución de los pacientes a lo largo del tratamiento clínico, sin embargo, asume que la ausencia de datos puede ser estimado independientemente de los parámetros subyacentes a la CVRS latente al inicio del tratamiento. Bajo este principio, los pacientes que reciben el mismo tratamiento presentan el mismo patrón de cambio a lo largo del tiempo y que el patrón de cambio no depende de la percepción de la CVRS al inicio del tratamiento. En consecuencia, la técnica propuesta es la suposición *hazard* proporcional del Modelo de Regresión de Cox para el análisis de supervivencia. Dentro de los estratos definidos por cada combinación ítem-paciente se establece que:

$$\text{logit}[p(X_{vit})] - \text{logit}[p(X_{vij})] = \delta_{vt} - \delta_{vi}$$

La restricción de este modelo se basa en la existencia de subgrupos de pacientes que reaccionan de formas diferentes a un tratamiento dado y que la única razón por datos faltantes es la mortalidad de los pacientes, la cual es asociada con puntuación más baja de CVRS. Dado que el modelo de Zwinderman parte del principio de que existe una

relación entre CVRS y mortalidad, y que los resultados pueden estar sesgados, el modelo sugiere combinar el análisis de las medidas de CVRS con el análisis de supervivencia. (Zwinderman, 1990)

La propuesta de Zwinderman representa un significativo aporte a la medición de CVRS ya que permite cuantificar el cambio debido al tratamiento clínico y el cambio en el tiempo. Esto sugiere que se podrá identificar la existencia de un cambio Gamma o Beta.

Algunas limitaciones al modelo propuesto por Zwinderman son:

- Los datos deben ser apareado y dicotómicos en pares de respuestas diferentes en los instantes T_1 y T_2
- No se evalúa al paciente de manera completa sino los datos a nivel de conjunto
- El cambio es igual para todas las variables latentes y sus efectos con otros tratamientos no pueden ser separados; el incremento del número de instantes (T_1, T_2, \dots, T_n) genera un modelo latente logístico más complejo
- La falta de datos, que posee un carácter aleatorio, es superada por la estimación del cambio promedio de los datos observados.

Las debilidades del modelo son fuertemente acentuadas al no considerar los cambios derivados de la autopercepción de los pacientes, pues esta condición subjetiva influye en las evaluaciones de CVRS. Una contribución que permite examinar las relaciones entre las variables estudiadas en el modelo de Zwinderman por medio de un método factorial son las técnicas Biplot.

3.3.3. El modelo Mayo et al.

Una propuesta para identificar a sujetos que experimentan potencialmente el cambio es presentada por Mayo et al. (2008) quienes desarrollan un modelo basado en la comparación longitudinal de la evaluación de pacientes (CVRS observada) y la evaluación esperada (en base a los síntomas y funcionamiento).

El modelo contempla tres pasos:

- El desarrollo de un modelo teórico-predictivo
- La formulación de un modelo residual
- La realización de un análisis de trayectoria grupal de los residuos centrados en los sujetos

El modelo está basado en el análisis de los residuos resultante de los efectos aleatorios que son centrados y utilizados para crear trayectorias. Estos residuos a nivel individual proporcionan información sobre la diferencia del nivel de percepción de la CVRS observada de cada paciente con respecto a la CVRS esperada. Las fluctuaciones en dicha diferencia determinan la presencia de *Response Shift* y la clasificación de las personas en grupos según la dirección e instante.

Los residuos obtenidos a partir del modelo traducen las diferencias entre ambos momentos y son utilizados para identificar a los pacientes que experimentaron cambios. Las fluctuaciones largas en los residuos sugieren la ocurrencia de *Response Shift*. Los residuos referentes a cada paciente son centrados en el respectivo valor medio de forma que remueva el impacto de la distancia entre las puntuaciones registradas.

Para medir el grado de ajuste, se comparan modelos que incluyen diferentes grupos de trayectorias utilizando los Criterios de Información de Akaike (AIC) y Bayesianos (BIC). Se estima para cada modelo la probabilidad del paciente que pertenece a un grupo dado, se evalúa el ajuste y se asigna los pacientes a distintos grupos.

El modelo plantea algunas fortalezas para su utilización como la capacidad de clasificar el cambio de un individuo en lugar del grupo, no es necesario incrementar progresivamente el número de pacientes, separa los grupos en base a los tiempos de medición y la dirección del cambio, y puede utilizarse para el análisis estratificado. Sin embargo, también establece algunas limitaciones como la suficiencia de datos que midan la condición consistente del paciente en los diferentes momentos, no permite

distinguir entre el error aleatorio y el *Response Shift*, y no es capaz de identificar el tipo de cambio ocurrido.

Mayo et al. (2008) presenta una importante contribución en la detección de *Response Shift* por su enfoque a nivel individual del paciente, y el modelo analiza los residuos centrados a lo largo del tiempo.

3.3.4. Los Modelos de Ecuaciones Estructurales

Los modelos de ecuaciones estructurales (SEM, por sus siglas en inglés de *Structural Equation Models*) no representan un conjunto preestablecido de técnicas estadísticas específicas que permiten establecer el cambio y su tipo, más bien corresponde a una familia de procedimientos basados en el análisis de la covarianza que examinan las relaciones causales entre variables cuando éstas son de tipo lineal. Estos modelos se consideran una extensión de varias técnicas multivariantes de regresión múltiple, el análisis factorial y el análisis de senderos (Afifi & Clark, 1990). Además, abarca una familia de modelos entre ellos el análisis de la estructura de la covarianza, análisis de variables latentes, análisis factorial confirmatorio, incluso el análisis LISREL (Jöreskog & Sörbom, 1996), entre otros.

Estos modelos se sustentan en el análisis de relaciones entre dos variables X y Y , ya sean de tipo causal (X causa Y) o no causal, con el objetivo de determinar sus correlaciones observadas.

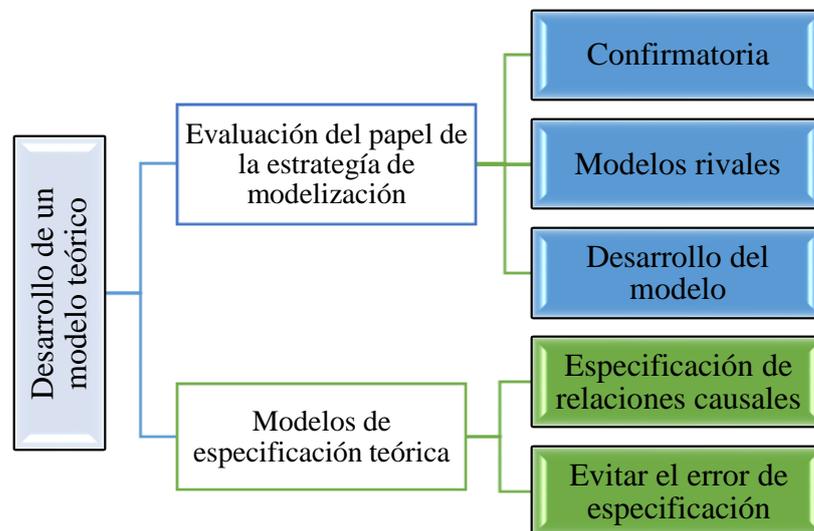
Una relación de causalidad supone que el cambio en una variable produce un cambio en otra. Esto se representa a través del diagrama causal que constituye una representación gráfica de un sistema de ecuaciones donde se especifican tanto las relaciones entre variables dependientes e independientes como las relaciones entre constructos y variables indicadoras. Estos modelos se ajustan a matrices de covarianzas, por lo que las variables sólo suelen ser consideradas centradas con respecto a su media y los parámetros no estandarizados. Las varianzas y covarianzas que se refieren únicamente a las variables causales constituyen por sí mismas como

los parámetros del modelo, siendo las reglas de descomposición para derivar las restantes varianzas y covarianzas.

Los modelos de ecuaciones estructurales permiten proponer el tipo de dirección de las relaciones que se espera encontrar entre las diversas variables contenidas en él (Ruiz et al., 2010). Para el desarrollo de los modelos de ecuaciones estructurales, Hair et al. (1999) describe siete pasos para la modelización:

- (1) El desarrollo de un modelo teórico: Se basa en las relaciones causales en las que el cambio en una variable supone un cambio en otra. El error más crítico en el desarrollo de modelos teóricos es la omisión de una o más variables predictivas claves. Esta omisión trae como consecuencia sesgar la evaluación de la importancia de otra variable (MacCallum, 1986). La Figura 14 presenta la metodología para el desarrollo del modelo teórico.

Figura 14. Metodología del desarrollo de un modelo teórico



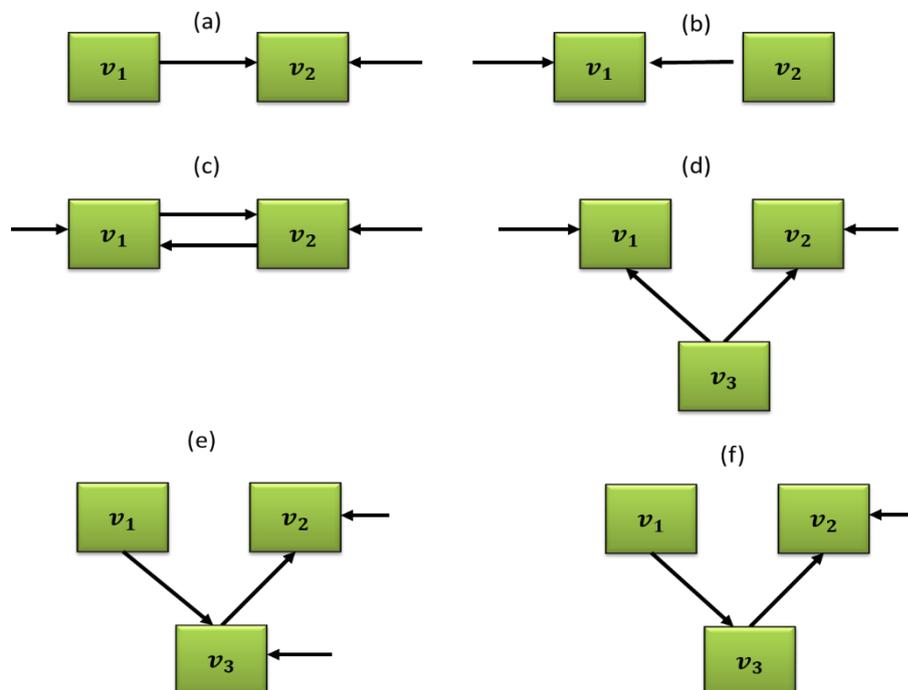
Fuente: MacCallum (1986)

- (2) La construcción de un diagrama de secuencia de relaciones causales: Estos diagramas permiten presentar las relaciones predictivas entre constructo, relaciones asociativas o correlaciones entre constructo e incluso entre

indicadores. Es decir, define constructos endógenos y exógenos y vincula las relaciones de un diagrama de relaciones. Una de las ventajas del diagrama causal es que presenta una figura para cada tipo de relación permitiendo visualizar las relaciones existentes.

Los elementos del diagrama causal representan una serie de símbolos utilizados para su construcción los cuales sintetizan el proceso causal donde un rectángulo representa la variable observada, un círculo representa la variable latente o factor común, una flecha directa representa una relación causal directa de un constructo a otro, una flecha curva entre constructo indica una correlación entre constructos, y dos flechas que conectan dos variables en sentido contrario indica una relación recíproca entre constructos, tal como se muestra en la Figura 15.

Figura 15. Representación de las relaciones causales a través del diagrama de secuencia



- (3) La conversión de un diagrama de secuencias en un conjunto de ecuaciones estructurales y especificación de un modelo de medida: Una vez definido el modelo teórico y representado en un diagrama de secuencias se debe especificar el modelo en términos formales a través de las ecuaciones estructurales que permita vincular los constructos, y especificar el modelo de medida señalando el número de indicadores, teniendo en cuenta la fiabilidad del constructor y la identificación de las correlaciones entre éste y las variables. El objetivo es vincular definiciones operacionales de los constructos con la teoría planteada para llegar a un contraste empírico apropiado. Por ello se establece el modelo estructural y el modelo de medidas.

Las ecuaciones estructurales del modelo lineal de variables latentes vienen definidas a través de la ecuación:

$$\eta = B_{\eta} + \Gamma_{\xi} + \zeta$$

donde,

η es un vector de variables latentes endógenas

B es una matriz de coeficientes para las variables latentes endógenas

Γ es una matriz de coeficientes para las variables latentes exógenas

ξ es un vector aleatorio de variables latentes exógenas

ζ es un vector de errores latentes en la ecuación

El modelo de medida tiene ecuaciones estructurales que representan el vínculo entre las variables latentes y las variables observadas. Las ecuaciones estructurales del modelo de medida están dadas por:

$$x = \Lambda_x \xi + \epsilon$$

$$y = \Lambda_y \eta + \delta$$

donde,

x es un vector de variables observadas indicadoras de ξ

y es un vector de variables observadas indicadoras de η

Λ_x es una matriz de coeficientes que relacionan x con ξ

Λ_y es una matriz de coeficientes que relacionan y con η

ϵ es un vector de errores de medida para x

δ es un vector de errores de medida para y

- (4) Selección del tipo de matriz de entrada y estimación del modelo propuesto: Los modelos de ecuaciones estructurales introducen una matriz de correlaciones o varianzas-covarianzas de todos los indicadores utilizados en el modelo, por lo que el modelo de medida especifica qué indicadores corresponde a cada constructo y las puntuaciones del constructo latente representan las utilizadas en el modelo estructural.

Esta técnica presenta tres supuestos: observaciones independientes, muestra aleatoria de los individuos y la linealidad de todas las relaciones. Los modelos de ecuaciones estructurales son más sensibles a las características distribucionales de los datos especialmente en el cumplimiento de la normalidad multivariante.

La estimación se puede realizar mediante cuatro procesos: estimación directa, *bootstrapping*, simulación y análisis *jackknife*. La estimación directa consiste en establecer un modelo con un procedimiento de estimación elegido, definición de parámetros e intervalo de confianza de cada parámetro estimado en base al error muestral.

El *bootstrapping*, que es una estimación de parámetro e intervalos de confianza basado en estimaciones múltiples (Bollen & Stine, 1992), se realiza en cuatro pasos: primero, se diseña una muestra original que actúe como población para los efectos muestrales, luego se vuelve a muestrear la muestra original para generar un gran número de nuevas muestras, seguido de estimar el modelo de

cada muestra nueva guardando los parámetros estimados y, por último, las estimaciones de los parámetros finales se calculan como la media de las estimaciones de los parámetros de todas las muestras.

Otro método de estimación es la simulación basada en muestras múltiples y modelos estimados. Estos procesos difieren del *bootstrapping* ya que, durante la generación de nuevas muestras, el programa de simulación puede cambiar ciertas características de la muestra para cumplir con las que se propone el investigador.

El análisis *jackknife* que consiste en crear N muestras repetidas nuevas, donde N es el tamaño de muestra original, donde cada vez que se crea una nueva muestra se omite una observación diferente (N-1) a partir de la muestra original.

- (5) Valoración de la identificación del modelo estructural: En esta etapa es importante considerar el principio de seleccionar una ecuación bien diferenciada para estimar cada coeficiente disminuyendo el número de incógnitas en la ecuación, ya que a medida que el modelo estructural se hace más complejo es más difícil lograr identificarlo. Para la identificación del modelo estructural se debe considerar establecer los grados de libertad, además de diagnosticar y solucionar los problemas de identificación.

Los grados de libertad representan la diferencia entre el número de correlaciones o covarianzas y el número efectivo de coeficientes en el modelo propuesto. Un grado de libertad es un elemento no restringido de la matriz de datos y se calcula a través de la siguiente ecuación:

$$gl = \frac{1}{2}[(p + q)(p + q + 1)] - t$$

donde,

p = número de indicadores endógenos

q = número de indicadores exógenos

t = número de coeficientes estimados en el modelo propuesto

En relación con diagnosticar los problemas de identificación, se pueden llevar a cabo contraste cuando la ecuación se identifica para determinar si los resultados son inestables debido al nivel de identificación. De manera inicial, el modelo puede ser reestimado varias veces, cada uno con un valor de partida diferente. Se puede especificar un valor inicial para cualquier parámetro estimado, es decir un punto de partida del proceso de estimación. Si los resultados no convergen en el mismo punto para diversos valores de partida, la identificación debería examinarse más a fondo.

Como prueba adicional, se evalúa el efecto de la identificación como un único coeficiente que consiste en estimar el modelo y obtener la estimación del coeficiente. Luego se fija el coeficiente a su valor estimado y se reestima la ecuación. Si el ajuste conjunto del modelo varía significativamente, esto indica problema de identificación.

Otro enfoque es observar los posibles síntomas de un problema de identificación. Estos pueden ser los errores estándar muy elevados para uno o más coeficientes, la incapacidad de programar para invertir la matriz de información, estimaciones muy poco razonables o estimaciones imposibles, y las elevadas correlaciones entre los coeficientes estimados.

En relación con la solución de los problemas de identificación, se establece identificar más restricciones para el modelo como la de eliminar algunos coeficientes estimados. Para ello, [Wardrop, et al. \(1993\)](#) recomienda construir un modelo teórico con el mínimo de coeficientes estimados que puedan ser justificados, fijar las varianzas de error de medida de los constructos si es posible, fijar algunos coeficientes estructurales conocidos con fiabilidad y eliminar las variables problemáticas.

- (6) Evaluación de los criterios de calidad de ajuste: Al establecer el modelo de estructura es indispensable evaluar la calidad de ajuste en diferentes niveles, tales como la evaluación del modelo conjunto, el modelo de medida y el modelo estructural. La calidad de ajuste mide la correspondencia entre la

matriz de entrada observada con la que se predice mediante el modelo propuesto.

Las medidas de calidad de ajuste son de tres tipos: medidas absolutas del ajuste que evalúan el ajuste global del modelo, medidas del ajuste incremental que cuales comparan el modelo propuesto con otro modelo especificado por el investigador, y las medidas de ajuste de parsimonia que ofrecen una comparación entre modelos con diferentes números de coeficientes estimados, siendo el propósito determinar la cantidad de ajuste conseguido por cada coeficiente estimado.

La medida del ajuste absoluto está fundamentada en la bondad de ajuste para los modelos de ecuaciones estructurales (Kline, 2005). Esta medida permite contrastar la hipótesis nula cuando el modelo es adecuado, en este sentido cuando los valores chi-cuadrado (χ^2) son altos en comparación con los grados de libertad del modelo, esto sugiere la existencia de un pobre ajuste. El modelo tiene un ajuste aceptable si los valores de chi-cuadrado dividido por los grados de libertad son de 2 a 3, y con límites hasta 5. (Hair et al., 1999)

Una de las desventajas de la medida chi-cuadrado es su sensibilidad a las diferencias del tamaño muestral, especialmente en los casos en que el tamaño muestral excede los 200 registros. A medida que este tamaño de la muestra aumenta, esta medida tiene una gran tendencia a indicar diferencias significativas para modelos equivalentes.

El parámetro de no centralidad y no centralidad a escala (NCP) es útil cuando el tamaño de muestra es pequeño. Definido como el resultado de la búsqueda de los estadísticos de una medida alternativa al ratio-estadístico de verosimilitud chi-cuadrado que se vean menos afectados por el tamaño de la muestra. Hair et al. (1999) lo calcula como sigue:

$$NCP = \chi^2 - \text{Grado de libertad}$$

Para estandarizar el NCP y obtener el parámetro no centralizado a escala se procede a dividirlo entre el tamaño de la muestra, es decir

$$SNCP = \frac{\chi^2 - \text{Grado de libertad}}{\text{Tamaño de la muestra}}$$

El error de aproximación cuadrático medio (RMSEA) representa el ajuste anticipado con el valor total de la población y no con el valor de la muestra extraída de estimación (Lévy, 2003). El RMSEA se puede interpretar como el error de aproximación medio por grado de libertad. Cuando los valores se mantienen en torno a 0.05, se consideran aceptables.

El RMSEA se define a partir de las comparaciones entre la matriz de varianza-covarianzas población y la matriz ajustada, y se establece como sigue:

$$RMSEA = \sqrt{\frac{NCP}{N \times \text{grados de libertad}}}$$

Para las medidas del ajuste incremental, se aplican los índices de Tucker-Lewis o de Bentler-Bonett. El primero es un índice de ajuste no normado (NNFI) planteado por Tucker y Lewis (1973). Este considera la parsimonia del modelo indicando que, si se le agregan más parámetros al modelo, el índice sólo aumenta si el estadístico χ^2 disminuye en mayor medida que los grados de libertad. Si el modelo es correcto, la esperanza del índice será aproximadamente igual a la unidad para cualquier tamaño de muestra.

El índice Tucker-Lewis se expresa:

$$TLI = \frac{(\chi_{nulo}^2 / gl_{nulo}) - (\chi_{propuesto}^2 / gl_{propuesto})}{(\chi_{nulo}^2 / gl_{nulo}) - 1}$$

donde,

χ_{nulo}^2 es el estadístico chi-cuadrado del modelo particular en estudio

$\chi_{propuesto}^2$ el chi-cuadrado del modelo

Un valor recomendado para el índice TLI es un nivel 0.90 o superior. Esta medida también se puede utilizar para comparar entre modelos alternativos mediante la sustitución del modelo alternativo por el modelo nulo (Hair et al., 1999).

Por su lado, el índice de ajuste normal (NFI), propuesto por Bentler y Bonett (1980), es una medida que va de 0 (ningún ajuste) a 1.0 (ajuste perfecto), y evalúa la disminución del estadístico chi-cuadrado de un modelo particular respecto al modelo base escogido, calculado de la siguiente forma:

$$NFI = \frac{\chi_{nulo}^2 - \chi_{propuesto}^2}{\chi_{nulo}^2}$$

donde,

χ_{nulo}^2 es el estadístico chi-cuadrado del modelo particular en estudio

$\chi_{propuesto}^2$ es el chi-cuadrado del modelo

No existe un valor absoluto que indique un ajuste aceptable, pero un valor recomendado normalmente es 0.90 o superior.

Las medidas de ajuste de parsimonia, las cuales utilizan el criterio de información de Akaike (AIC), propuesto por Akaike (1987), son una medida de información teórica, útil para comparar modelos sin necesidad de dividir la muestra, y no evalúa modelos aislados. El AIC se calcula:

$$AIC = \chi^2 + 2p$$

donde,

χ^2 es el valor del estadístico chi-cuadrado del modelo

2 el número de penalidad relativas

p representa el número de parámetros estimados.

Los valores AIC cercanos a cero indican un mejor ajuste y una mayor parsimonia. Con ello se obtiene, no sólo un buen ajuste de las covarianzas y correlaciones previstas y observadas, sino que también un modelo no propenso al sobreajuste.

- (7) Interpretación y modificación del modelo: Al realizar la evaluación del modelo se procede a revisar los resultados y sus correspondencias con la teoría propuesta. Estas evaluaciones se deben realizar considerando los resultados estandarizados, los índices de modificación y los cambios potenciales del modelo.

La especificación del modelo ayuda al investigador a saber si el primer modelo obtenido es el mejor, por lo que es necesario buscar métodos para mejorar el ajuste añadiendo o eliminar los parámetros estimados del modelo original. La interpretación permite establecer el modelo adecuado, además de aceptar o rechazar las hipótesis planteadas.

Jöreskog (1993) sintetiza las aplicaciones de los modelos de ecuaciones estructurales en tres objetivos diferentes: confirmar un modelo, elegir modelos alternativos y generar un modelo. La confirmación de un modelo se refiere a la etapa cuando el investigador posee un modelo que se acepta o rechaza en base a la correspondencia con las relaciones evidenciadas en un determinado conjunto de datos y tiene un objetivo estrictamente confirmatorio. Elegir modelos alternativos corresponde cuando el investigador dispone de varios modelos posibles y pretende seleccionar uno, y la generación de modelos se refiere a la etapa donde se establece un modelo inicial que es sustentable teórica como estadísticamente.

Para el estudio de la detección y categorización del *Response Shift*, los modelos de ecuaciones estructurales aprovechan el análisis de las estructuras de covarianza usando el Análisis Factorial Confirmatorio (AFC) para identificar reconceptualización y recalibración a través del análisis de cambios en la estructura y patrón factorial a lo largo del tiempo.

3.3.4.1. Los modelos de Curvas de Crecimiento Latente

Uno de los modelos más utilizados para el análisis de datos longitudinales son las Curvas de Crecimiento Latente como parte de los modelos de ecuaciones estructurales. El análisis de Curvas de Crecimiento Latente permite examinar medidas repetidas como mecanismos para identificar la ocurrencia de cambios con datos de CVRS, a su vez de investigar el comportamiento implícito en una variable respuesta a lo largo del tiempo (McArdle & Epstein, 1998; Meredith & Tisak, 1990; Muthén, 1991; Willett & Sayer, 1994). El objetivo se centra en estimar los factores de crecimiento que no son observados y que generan las medidas repetidas observadas. Las curvas enfocan las trayectorias de crecimiento de cada individuo, lo que permite modelar la tendencia del cambio individual o de grupos.

Los modelos de Curvas de Crecimiento Latente se definen en cuatro evaluaciones de una variable observable Y sobre un conjunto de individuos como una función de trayectoria común lineal (Curran et al., 2004; Curran & Hussong, 2001, 2003)

$$y_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{it}$$

donde,

y_{it} representa la medida obtenida del individuo i en el instante t

α_i Es la intersección de la trayectoria subyacente para el individuo i

$\lambda_t = 0,1,2,3$, para las $T=4$ evaluaciones

β_i es la pendiente lineal de la trayectoria subyacente para el individuo i

λ_t Coeficiente del tiempo

ε_{it} Error aleatorio asociado al individuo i en el instante t

El intercepto representa el valor inicial asociado a cada individuo, que es el mismo a lo largo de los diferentes tiempos, la pendiente se relaciona con la tasa de crecimiento

de cada individuo. Ambos son modelizados como variables latentes necesarias para describir el crecimiento.

El intercepto y pendiente son los componentes de la función de crecimiento y se pueden plantear de la siguiente manera:

$$\alpha_i = \mu_\alpha + \varepsilon_{\alpha i}$$

$$\beta_i = \mu_\beta + \varepsilon_{\beta i}$$

Siendo las μ_α y μ_β la intersección media y las agrupaciones sobre todos los individuos y $\varepsilon_{\alpha i}$ y $\varepsilon_{\beta i}$ las desviaciones de cada individuo de la media del grupo.

El primer paso del análisis consiste en ajustar el modelo de curva latente presentado al conjunto de datos disponibles. Las medidas repetidas son utilizadas como indicadores múltiples de los dos factores latentes, donde el primer factor representa el intercepto de la trayectoria y el segundo la pendiente. El modelo se define restringiendo las cuatro cargas factoriales del intercepto iguales a 1 y las cuatro cargas factoriales relativas a la pendiente igual a $\lambda_t = t - 1$.

Los efectos fijos son las medias de cada factor latente (μ_α y μ_β) que son calculadas a partir de todos los individuos de la muestra. Los efectos aleatorios son representados por cuatro parámetros: la varianza de cada factor latente que refleja el grado de variabilidad individual en los intercepto y pendientes sobre todos los individuos de la muestra (ψ_α y ψ_β), la covarianza entre los dos factores latentes que se interpreta como el grado de asociación entre la variación individual de los intercepto y la variación de las pendientes ($\psi_{\alpha\beta}$), y la varianza residual para cada medida repetida que traduce la varianza no explicada (σ_ε^2). En conjunto, estos parámetros captan la trayectoria de la muestra en estudio, el grado de variabilidad sobre las trayectorias individuales en torno de los valores medios y la proporción de variabilidad no explicada por la trayectoria.

Este planteamiento establece un modelo de carácter lineal. Sin embargo, cuando las medidas repetidas obtenidas pueden relacionarse con el tiempo de forma no lineal de manera que el cambio en y no es constante entre evaluaciones. Esta evaluación puede

asumir muchas formas distintas como cuadrática, exponencial o funciones que son estimadas a partir de las características de la muestra en estudio.

El modelo cuadrático para las trayectorias individuales incluye un tercer factor latente de forma que capta la curvatura presente en las trayectorias individuales, el cual se plantea de la siguiente manera:

$$y_{it} = \alpha_i + \lambda_t \beta_{Li} + \lambda_t^2 \beta_{Qi} + \varepsilon_{it}$$

donde

y_{it} representa la medida obtenida del individuo i en el instante t

α_i es la intersección de la trayectoria subyacente para el individuo i

β_{Li} es la pendiente lineal de la trayectoria subyacente para el individuo i

β_{Qi} es la pendiente cuadrática de la trayectoria para el individuo i

Las cargas factoriales reflejan las relaciones entre el factor cuadrático latente y las medidas repetidas que son los cuadrados de los factores lineales.

El modelo lineal implica un cambio constante en y entre evaluaciones consecutivas, mientras que el modelo cuadrático implica un cambio diferencial. Por otra parte, los modelos exponenciales destacan los incrementos o disminuciones de manera proporcional a los incrementos o disminuciones para un evento previo, y se plantea de la siguiente forma:

$$y_{it} = \alpha_i + (1 - e^{-\gamma \lambda_t}) \beta_i + \varepsilon_{it}$$

donde,

y_{it} representa la medida obtenida del individuo i en el instante t

α_i es la intersección de la trayectoria

β_i es la magnitud del cambio implícito de una observación relativa al nivel inicial

γ es la tasa de crecimiento en y a lo largo del tiempo

Cuando los modelos anteriores no proporcionan un buen ajuste, un modelo alternativo es estimar la forma directamente a partir de los datos. Esto implica la estimación de un factor de intercepto y de un factor de cambio latente simple donde el conjunto de las cargas factoriales del factor latente pendiente son libremente estimados a partir de los datos en vez de fijarlos. En este caso, este factor latente capta la no linealidad del cambio a través del tiempo. El estadístico chi-cuadrado (χ^2) puede proporcionar una estimación del grado de mejoramiento al ajustar un modelo relativamente a otro.

Si se encuentra variabilidad individual significativa en los factores intercepto y pendiente, se procede a introducir variables explicativas de forma a predecir la variabilidad observada. Estableciendo dos variables explicativas z_1 y z_2 , el modelo se formula:

$$\alpha_i = \mu_\alpha + \gamma_1 Z_{1i} + \gamma_2 Z_{2i} + \zeta_{\alpha i}$$

$$\beta_i = \mu_\beta + \gamma_3 Z_{1i} + \gamma_4 Z_{2i} + \zeta_{\beta i}$$

Este es un modelo de crecimiento condicional debido a que las diferencias individuales observadas en el punto inicial y el cambio a través del tiempo están siendo modelizados como función de la variable z . Por otra parte, los parámetros de regresión γ asociados a las variables explicativas son estimados de forma común a todos los individuos.

A pesar de que los modelos de Curvas de Crecimiento Latente pueden sugerir la presencia de *Response Shift*, éstos no llegan a estimar directamente los componentes de recalibración, repriorización y reconceptualización. Por lo que se convierten en un primer paso para evaluar la trayectoria de un conjunto de datos (pacientes) a lo largo del tiempo. Se han identificado un conjunto de investigaciones recientes que utilizan los modelos de Curvas de Crecimiento Latente para la identificación del *Response Shift* en datos de CVRS, los cuales se detallan en la Tabla 17.

Tabla 17. Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift utilizando los modelos de Curvas de Crecimiento Latente (2010-2020)

No.	Año de Publicación	Autor Principal	Título de la investigación
1	(2011)	Amed et al.	Using latent trajectory analysis of residuals to detect response shift in general health among patients with multiple sclerosis article
2	(2011)	Schwartz et al.	Response shift in patients with multiple sclerosis: an application of three statistical techniques
3	(2013)	Flynn et al.	Quantifying response shift or adaptation effects in quality of life by synthesizing best-worst scaling and discrete choice data
4	(2017)	Powell et al.	Changes in quality of life from a homelessness intervention: true change, response shift, or random variation
5	(2017b)	Salmon et al.	Identifying patterns of adaptation in breast cancer patients with cancer-related fatigue using response shift analyses at subgroup level
6	(2020)	Tew et al.	Patient-reported outcome measures (PROMs): can they be used to guide patient-centered care and optimize outcomes in total knee replacement?

Sin embargo, una ampliación de este modelo es la propuesta por [Zee \(1998\)](#) que presenta a la Curva de Crecimiento Latente bajo la ocurrencia de cambios en el estado de los pacientes para modelar los datos de CVRS.

3.3.4.2. El modelo Zee

El modelo de [Zee \(1998\)](#) sugiere la extensión del modelo de Curvas de Crecimiento Latente presentado para controlar el efecto de los estados de salud experimentados por los pacientes. Para ello se introduce una variable que cambia en el tiempo y que representa el *status* o estado de salud. Para el autor, el *status* representa el estado de salud en el instante que se evaluó un paciente. Estos estados de salud pueden ser clasificados como: en tratamiento, fuera de tratamiento, antes de la administración del tratamiento y después de la administración del tratamiento.

La incorporación de los estados de salud se basa en la suposición de que la CVRS de los pacientes es afectada por una condición específica, en especial cuando el objetivo es identificar los impactos a corto y largo plazo de un tratamiento administrado en oposición a un tratamiento estándar.

Zee plantea su modelo estableciendo la ocurrencia de datos faltantes y de medidas obtenidas a partir de visitas irregulares, el cual se presenta como:

$$y_{ri} \sim N \left(\alpha_r \chi_{ri} + \sum_{s=1}^M \delta_{ris} (\beta_{r0s} + \beta_{r1s} t_s + \beta_{r2s} t_s^2 + \dots + \beta_{rps} t_s^p), \Sigma_r \right)$$

donde,

$$\delta_{risj} = \begin{cases} 0 & \text{si la medida no ha sido obtenida en el estado } s \\ 1 & \text{si la medida ha sido obtenida en el estado } s \end{cases}$$

t_{sj} es un vector q_{ris} dimensional que contiene los instantes en que han sido medidos los datos de CVRS. El modelo de curva de crecimiento modificado para un vector de observaciones y_{ri} , tiene distribución normal multivariante con media

$$y_{rij} = f(\varpi_r; x_{ri}, \delta_{risj}, t_{sj}) = \alpha_r x_{ri} + \delta_{ris} (\beta_{r0s} + \beta_{r1s} t_{sj} + \beta_{r2s} t_{sj}^2 + \dots + \beta_{rps} t_{sj}^p)$$

donde ϖ_r es el vector de parámetros $(\alpha_r, \beta_{r0s}, \beta_{r1s}, \beta_{r2s}, \dots, \beta_{rps})^T$, y matriz de varianza-covarianza intraclase Σ_r con parámetros δ_{rs}^2 y ρ_{rs} , $s = 1, \dots, M$. La estructura de correlación intraclase puede ser definida dentro de cada grupo de tratamiento y dentro de cada estado de salud con el objetivo de volver este modelo tan realista cuanto posible en la práctica.

Las curvas de crecimiento globales asociadas a los tratamientos pueden ser comparadas dado el valor medio del estado de CVRS. En un ensayo clínico aleatorio, el objetivo es comparar la CVRS entre grupos de tratamiento mediante un contraste apropiado dentro del modelo para demostrar el efecto del tratamiento. El contraste entre grupos de tratamiento se puede realizar una vez se obtengan las estimaciones de los parámetros y las respectivas varianzas asintóticas.

La propuesta de Zee de los estados de salud experimentados por los pacientes con el objetivo de describir un patrón global con datos faltantes y medidas obtenidas a partir de visitas irregulares, permite justificar la existencia de distintas fases del tratamiento y que la CVRS puede variar, por tanto, se puede utilizar una variable dependiente del tiempo como partición de las medidas de CVRS repetidas. Si se comparan tratamientos sin tener en cuenta los estados de salud, se obtendrán resultados sesgados debido a la eventual diferencia de eficacia de los mismos tratamientos. Por tanto, el control de los estados de salud experimentados por los pacientes permitirá reducir eventuales sesgos en la comparación de tratamientos.

El modelo de [Zee \(1998\)](#) permite superar algunos problemas propios del análisis de datos de CVRS como: la idea general sobre el patrón evolutivo de la CVRS, la posibilidad de contrastar la diferencia en el patrón global de CVRS entre tratamientos, la posibilidad de las estimaciones asociadas a la varianza y correlación entre los estados de salud y los tratamientos, y la aleatoriedad de los datos faltantes.

Sin embargo, la propuesta también posee algunas limitantes: no se posibilita la cuantificación del cambio asociado a las variables latentes y sólo la evolución del patrón global de CVRS, supone que las medidas de CVRS poseen distribución normal multivariante, supone que la matriz de varianzas-covarianzas tiene estructura simétrica, ignora los mecanismos de adaptación por parte del paciente, y no considera la ocurrencia de fenómenos de recalibración, repriorización o reconceptualización en los pacientes.

3.3.4.3. El modelo Schmitt

En el marco del desarrollo de los modelos de ecuaciones estructurales, [Schmitt \(1982\)](#) presentó un método para valorar diferentes tipos cambios y la falta de consistencia de los resultados a lo largo del tiempo, lo cual implica examinar mediciones tomadas antes y después de una intervención para determinar el alcance de la diferencia en el patrón de factores de cargas, la escala métrica o unidades de medidas. Este planteamiento se basa en el análisis factorial confirmatorio.

El análisis de los datos se realiza en cuatro etapas: prueba de homogeneidad de matrices de varianza-covarianza, prueba de igualdad del patrón de factor de las mediciones, prueba de igualdad de unidades de escala y varianzas de factor, y prueba de igualdad de unicidades.

La prueba de homogeneidad de matrices de varianza-covarianza se realiza a través de la prueba Box. La misma consiste en obtener una diferencia significativa entre matrices antes y después de una intervención que sugiere que: el patrón de factor es diferente o que ha ocurrido cambio Gamma, que las unidades de escala son diferentes o que ha ocurrido un cambio Beta, o que las unidades son diferentes o que tiene lugar una combinación de lo mencionado.

Si la prueba Box indica que las matrices de varianza-covarianza son similares, es posible estimar el único patrón de factor.

La prueba de igualdad del patrón de factor de las mediciones (cambio Gamma) permite determinar hasta qué punto las medidas tomadas antes y después de la intervención miden el mismo concepto. Esto señala que, al medir los mismos factores, si el número de factores y el patrón de carga sigue siendo el mismo, se podrá concluir que no existe evidencia de cambio Gamma.

El modelo de medida aplicada es el modelo formal del análisis de factor común dado por:

$$x = \Lambda f + u$$

donde,

x es el vector de respuesta observadas

Λ es el conjunto de factores de carga

f es la matriz de factores comunes

u son los factores únicos

La matriz esperada de varianza-covarianza de x se plantea en $\Sigma = \Lambda\Phi\Lambda' + \Theta$, siendo Φ la matriz de varianza-covarianza de los factores comunes y Θ es la matriz diagonal de varianza de los factores únicos.

La prueba de igualdad de unidades de escalas y varianzas de factor (cambio Beta) surge cuando el modelo tiene el mismo número de factores antes y después de la intervención y se acepta la prueba de igualdad de factores de carga. Las cargas para las mediciones antes y después serán iguales y sólo se necesita una matriz Λ que representa los factores de cargas de ambos periodos.

Se manifiesta un cambio Beta cuando se presenta una disminución significativa en el ajuste entre las matrices observadas y las obtenidas con el modelo.

La prueba de igualdad de unicidades corresponde a una prueba que involucra una restricción de igualdad en los elementos similares de las matrices Θ antes y después de la intervención, cuando las varianzas de factores comunes son iguales. Una diferencia chi-cuadrado significativa o un decremento significativo de las varianzas generados por el modelo podría ser una señal que existe una diferencia con respecto a la fiabilidad de las mediciones antes y después de la intervención.

La aplicación del modelo Schmitt permite proporcionar una prueba general de la similaridad de las matrices de varianza-covarianza antes y después de la intervención, así como varios tipos de cambio y el grado de varianza asociados, en adición permite valorar la posibilidad del cambio a través de una medición antes de la intervención y una medición después de la intervención.

3.3.4.4. El planteamiento de Taris

De acuerdo con los planteamientos de [Schmitt \(1982\)](#) y [Taris \(2000\)](#) propone un procedimiento utilizando el análisis factorial confirmatorio que simplifica en tres pasos la detección de la invarianza de las estructuras de factores a través del tiempo asumiendo que los ítems de interés fueron medidos en dos puntos distintos del tiempo.

Este procedimiento consiste en probar una hipótesis de invarianza estructural dentro del marco del análisis factorial confirmatorio. Este análisis es el más apropiado para probar las hipótesis ya que permite realizar especificaciones precisas de las relaciones entre las variables observadas y las dimensiones subyacentes. Adicionalmente, proporciona pruebas para el ajuste de un modelo factorial específico a los datos. En este sentido se puede valorar el modelo con el mejor ajuste a los datos en los diferentes momentos.

Los pasos propuestos por [Taris \(2000\)](#) incluyen la igualdad de las matrices de varianza-covarianza a través del tiempo, cambio Gamma y cambio Beta.

El procedimiento inicia con la prueba de igualdad de las matrices de varianza-covarianza entre las variables en estudio a través del tiempo. Las matrices de datos obtenidos en los diferentes periodos se analizan como si hubiesen sido obtenidas de muestras estadísticamente independientes y, al plantear el modelo, las matrices de varianza-covarianza se restringen para ser iguales entre sí.

En el segundo paso, cambio Gamma, si se evidencia que existe diferencias en las estructuras factoriales de los diferentes periodos de estudio, entonces se verifica si la misma estructura de factor y patrones de factores de carga se aplica en cada punto del tiempo. Por otro lado, si el número de factores en la estructura simple es el mismo a través del tiempo, es necesario comprobar que el patrón de factores de carga tampoco cambia. En consecuencia, no se detecta la ocurrencia de cambio Gamma.

El procedimiento inicia estableciendo un modelo sencillo con sólo un factor y todos los ítems cargados sobre este único factor, se realiza el análisis y la verificación de la prueba chi-cuadrado y los índices de ajustes correspondientes. Si la primera verificación produce un rechazo del modelo, entonces se plantea un modelo con dos factores hasta conseguir el modelo más simple que otorga el mejor ajuste.

En el tercer paso, la detección del cambio Beta se presenta si las cargas de los ítems en las dimensiones subyacentes se restringen de manera que sean iguales a través del tiempo. El modelo debe ajustar a los datos, así como la comparación de los índices de ajustes obtenidos para el modelo restringido con aquellos obtenidos para el modelo sin

restricción. Esto debe indicar que el ajuste del modelo restringido no es significativamente peor, de tal manera que es posible asumir que los factores de carga sean iguales (Rios, 2006).

Si la estructura básica del modelo es la misma en diferentes ocasiones, alguna parte del este se puede restringir para ser invariante a través del tiempo. La comparación de un modelo particular restringido con el ajuste de un modelo similar sin restricción puede revelar si las restricciones impuestas son específicamente favorables o no. Si el modelo con las restricciones de invarianza estructural a través del tiempo es estadísticamente válido, entonces no ocurre cambio Gamma o cambio Beta. En caso contrario, puede haber ocurrido uno de estos cambios o ambos.

3.3.4.5. El modelo Oort

Oort (2005) propone una metodología basada en los modelos de ecuaciones estructurales para detectar los componentes de *Response Shift* (recalibración, repriorización y reconceptualización) y medir el verdadero cambio aplicable a uno o más constructos relacionados a CVRS a través de múltiples ítems. El procedimiento propuesto se basa en el análisis de las medias y covarianzas sólo para detectar *Response Shift*, y el verdadero cambio si el fenómeno es experimentado por una parte substancial de los individuos utilizando datos de dos tiempos de medida. Esta diferenciación posibilita la detección de varios tipos de *Response Shift* y la medición del verdadero cambio.

Este procedimiento ha sido formulado para detectar *Response Shift* y cambios en un grupo de individuos y no a nivel individual, ya que los modelos de ecuaciones estructurales se centran en las medias y covarianzas, que son estadísticos de grupo. Bajo la premisa que, si muchos sujetos cambian según el mismo patrón evidenciando el mismo nivel de cambio individual, ese cambio será evidente en los estadísticos de grupo.

El modelo propuesto estará dado por:

$$x = \tau + \Lambda\xi + \nabla\varepsilon$$

La estructura de medias y covarianza:

$$E(x) = \mu = \tau + \Lambda k$$

$$Cov(x, x') = \Sigma = \Lambda\Phi\Lambda' + \Delta\Theta\Delta'$$

donde,

x : vector de observaciones de J variables en K ocasiones para un individuo

τ : vector de intercepto para J variables en K ocasiones

Λ : matriz de factores de carga para los K periodos de estudio

ξ : vector de valores de las R variables latentes en los periodos de estudio

Δ : matriz de carga de los factores residuales para las K ocasiones

ε : vector de factores residuales para las J variables en K ocasiones

k : vector de medias de las variables latentes en los K periodos de estudio

Φ : matriz de varianza-covarianza de las R variables latentes para todos los K

Θ : matriz de varianza-covarianza de los factores únicos o residuales en los K periodos de estudio

De acuerdo con lo planteado por [Oort \(2005\)](#), no siempre es posible detectar todos los tipos cambios por las relaciones que pueden existir entre sí, y que para detectar uno es necesario suponer la no existencia del otro.

Durante la modelación de ecuaciones estructurales relacionadas con la identificación del *Response Shift*, se puede puntualizar lo siguiente:

- Los factores comunes o variables latentes representan las dimensiones del fenómeno en estudio y a la vez son definidos por las variables observadas. Por lo tanto, existe un patrón de carga que relaciona ciertas variables observadas con cada una de las variables latente. Si el número de factores difiere de una ocasión a otra y los patrones de carga de las variables latentes son tan diferente que no es posible identificar algún factor que mantengan su significado a través de las ocasiones, entonces la reconceptualización ha sido tan fuerte que no es posible continuar con la prueba de los otros tipos de cambios.
- Si el valor del factor de carga de una variable observada particular cambia significativamente de una ocasión a otra se puede señalar la ocurrencia de una repriorización.
- Considerando que la recalibración es el cambio de los estándares de medición internos de los individuos, es posible considerar dos tipos. Si el cambio afecta a todas las opciones de respuesta en la misma dirección y en la misma extensión, entonces la recalibración es uniforme y sólo afecta a las medias de las variables observadas. Mientras, si las escalas de medición pueden estirarse o encogerse completa o parcialmente, la recalibración será no uniforme y aparece tanto en la estructura de medias como en la estructura de covarianza de las variables observadas.

Los posibles cambios pueden identificarse a través de los parámetros del modelo lineal de variable latente, es decir matrices que caracterizan las estructuras de medias y covarianzas del modelo. Cada tipo de cambio en este modelo es asociado con una hipótesis de invarianza.

La matriz Λ de efectos de carga puede presentar cambios en los valores a través del tiempo y cambios en los patrones de los valores iguales a cero y diferente a cero; variables observadas iguales a cero para un determinado factor no pertenecen al dominio que define este factor.

Estos cambios en la matriz Λ indican que pueden existir cambio en el peso que cada variable observada tiene sobre los factores, cambio en el número de factores en cada

ocasión y cambio en los patrones de factores de carga (el conjunto de variable que indica un determinado factor no es el mismo en todas las ocasiones). Cambio en el peso hace referencia a repriorización, cambio en el número de factores y cambio en el patrón se refiere a reconceptualización. Para probar la existencia de este tipo de cambio es necesario la invarianza de las cargas $\Lambda_1 = \Lambda_2 \dots = \Lambda_k$ y la invarianza de los patrones de cargas $\text{Pat}(\Lambda_1) = \text{Pat}(\Lambda_2) \dots = \text{Pat}(\Lambda_k)$, a través del tiempo. Si esta hipótesis nula de invarianza se rechaza, entonces ha ocurrido repriorización y reconceptualización, respectivamente.

Los factores específicos representan aquellas partes de las variables observadas que no dependen de las variables latentes ni los errores de medidas, es decir, dependen únicamente de la respuesta de brindan los individuos y los propios cuestionarios. Entonces la presencia de recalibración implica falta de invarianza de los patrones de los factores específicos y en los intercepto τ que son característicos de los instrumentos de medición.

La recalibración uniforme se evidencia cuando existe cambio en la medición de los factores específicos y en los intercepto; pero debido a los requerimientos de identificación, no es posible probar la hipótesis de invarianza de medidas de ambos parámetros, sino que se debe asumir invarianza de uno de ellos como verdadera para probar la otra.

En los modelos de factor común usuales, tal como el modelo lineal de variables latentes del análisis factorial confirmatorio, no se hace distinción entre factores específicos y factores de error, solamente considera los factores residuales que engloban ambos factores mencionados. La hipótesis de invarianza para detectar recalibración uniforme está dada en los intercepto que deben ser iguales a través del tiempo $\tau_1 = \tau_2 \dots = \tau_k$.

La recalibración no uniforme se evidencia en los cambios de la varianza de los factores específicos, pero por los requerimientos de identificación se debe imponer la invarianza de las cargas de los factores específicos. Con el modelo lineal de modelos latentes se prueba la existencia de recalibración no uniforme cuando existen cambios

en las varianzas de los factores residuales a través de los momentos, por tanto, la hipótesis a probar es $Diag(\Theta_{11}) = Diag(\Theta_{22}) \dots = Diag(\Theta_{kk})$.

El cambio Alfa en términos de los parámetros del modelo lineal de variables latentes es el cambio en las medidas de las variables latentes k . De esta manera, el cambio verdadero puede ser investigado probando la hipótesis nula de invarianza de k a través del tiempo $k_1 = k_2 \dots = k_k$.

Mediante la hipótesis de invarianza de medida también puede probarse otros tipos de cambios, tales como el cambio en las varianzas de los factores comunes, el cual significa que los individuos son más homogéneos respecto a los factores comunes. De igual forma, el cambio en k (cambio Alfa) es denominado cambio verdadero en las medias, el cambio que se observa al rechazar la hipótesis nula $H_0: Diag(\Theta_{11}) = Diag(\Theta_{22}) \dots = Diag(\Theta_{kk})$ se denomina cambio verdadero en las varianzas.

Un cambio en las correlaciones entre las variables latentes significa que las interrelaciones entre los factores comunes cambian. Este cambio puede ser interpretado como una repriorización o reconceptualización en la medida de uno o más factores de orden superior. La matriz de correlaciones de factor común puede denominarse Φ^* , y la hipótesis de invarianza: $\Phi_{11}^* = \Phi_{22}^* \dots = \Phi_{kk}^*$ (Rios, 2006).

En términos generales, los modelos de ecuaciones estructurales son un conjunto de técnicas estadística no predeterminadas que utilizan el análisis de covarianzas para detectar cambios entre los grupos de individuos o dimensiones de datos. Sin embargo, la limitación de las técnicas tradicionales donde no se identifica el cambio individual a lo largo de cada grupo o dimensión puede ser subsanada con otras técnicas centradas en el análisis factorial para detectar *Response Shift* a nivel individual y categorizar el tipo de cambio. La Tabla 18 describe publicaciones recientes que han utilizado los modelos de ecuaciones estructurales de Oort (2005) para la identificación de *Response Shift*.

Tabla 18. *Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift, utilizando el Modelo Oort (2010-2020)*

No.	Año de Publicación	Autor Principal	Título de la investigación
1	(2010)	King-Kallimanis & Oort	Using structural equation modelling to detect measurement bias and response shift in longitudinal data
2	(2011)	King-Kallimanis et al.	Using structural equation modeling to detect response shift in performance and health-related quality of life scores of multiple sclerosis patients
3	(2011)	Schwartz et al.	Response shift in patients with multiple sclerosis: an application of three statistical techniques
4	(2011)	Barclay-Goddard et al.	Health-Related Quality of Life After Stroke: Does Response Shift Occur in Self-Perceived Physical Function?
5	(2013)	Visser et al.	The role of recalibration response shift in explaining bodily pain in cancer patients undergoing invasive surgery: an empirical investigation of the Sprangers and Schwartz model
6	(2013)	Fokkema et al.	Response Shifts in Mental Health Interventions: An Illustration of Longitudinal Measurement Invariance
7	(2013)	Gandhi et al.	Assessment of response shift using two structural equation modeling techniques
8	(2013)	Gandhi et al.	Influence of explanatory and confounding variables on health-related quality of life after controlling for measurement bias and response shift in measurement
9	(2014)	Ahmed et al.	Minimal evidence of response shift in the absence of a catalyst
10	(2014)	Barclay & Tate	Response shift recalibration and reprioritization in health-related quality of life was identified prospectively in older men with and without stroke
11	(2014)	DeConde et al.	Response Shift in Quality of Life After Endoscopic Sinus Surgery for Chronic Rhinosinusitis
12	(2015)	Mayo et al.	How are you? Do people with inflammatory bowel disease experience response shift on this question?
13	(2015)	Guilleux et al.	RespOnse Shift ALgorithm in Item response theory (ROSALI) for response shift detection with missing data in longitudinal patient-reported outcome studies
14	(2015)	Traa et al.	Evaluating quality of life and response shift from a couple-based perspective: a study among patients with colorectal cancer and their partners
15	(2015)	Verdam et al.	Taking into account the impact of attrition on the assessment of response shift and true change: a multigroup structural equation modeling approach
16	(2016)	Gandhi et al.	An item-level response shift study on the change of health state with the rating of asthma-specific quality of life: a report from the PROMIS Pediatric Asthma Study
17	(2016)	Gerlich et al.	Prostate cancer patients' quality of life assessments across the primary treatment trajectory: 'True' change or response shift?
18	(2016)	Lix et al.	Response shift and disease activity in inflammatory bowel disease
19	(2016)	Nolte et al.	On the validity of measuring change over time in routine clinical assessment: a close examination of item-level response shifts in psychosomatic inpatients
20	(2016)	Reissmann et al.	Longitudinal measurement invariance in prospective oral health-related quality of life assessment

21	(2016)	Sébille et al.	Prospective, multicenter, controlled study of quality of life, psychological adjustment process and medical outcomes of patients receiving a preemptive kidney transplant compared to a similar population of recipients after a dialysis period of less than three years – The PreKit-QoL study protocol
22	(2016)	Verdam et al.	Using structural equation modeling to detect response shifts and true change in discrete variables: an application to the items of the SF-36
23	(2016)	Wu	Response Shifts in Depression Intervention for Early Adolescents
24	(2017)	Chen et al.	Are the Insomnia Severity Index and Pittsburgh Sleep Quality Index valid outcome measures for Cognitive Behavioral Therapy for Insomnia? Inquiry from the perspective of response shifts and longitudinal measurement invariance in their Chinese versions
25	(2017)	Jabrayilov et al.	Longitudinal measurement invariance of the Dutch Outcome Questionnaire-45 in a clinical sample
26	(2017)	Sajobi et al.	Response shift in parents' assessment of health-related quality of life of children with new-onset epilepsy
27	(2017a)	Salmon et al.	Identifying patterns of adaptation in breast cancer patients with cancer-related fatigue using response shift analyses at subgroup level
28	(2017a)	Verdam et al.	Structural equation modeling–based effect-size indices were used to evaluate and interpret the impact of response shift effects
29	(2018)	Friedrich et al.	Response shift effects of quality of life assessments in breast cancer survivors
30	(2018)	Krägeloh et al.	Response Shift After a Mindfulness-Based Intervention: Measurement Invariance Testing of the Comprehensive Inventory of Mindfulness Experiences
31	(2018)	Schwartz et al.	Quality of Life and Adaptation in People With Spinal Cord Injury: Response Shift Effects From 1 to 5 Years Postinjury
32	(2018)	Wang et al.	Using structural equation modeling to detect response shift in quality of life in patients with Alzheimer's disease
33	(2019)	Carlier et al.	Testing for response shift in treatment evaluation of change in self-reported psychopathology amongst secondary psychiatric care outpatients
34	(2019)	Friedrich et al.	Response shift effects in patients' assessments of their quality of life after cardiac rehabilitation
35	(2020a)	Murata et al.	Response Shift–Adjusted Treatment Effect on Health-Related Quality of Life in a Randomized Controlled Trial of Taxane Versus S-1 for Metastatic Breast Cancer: Structural Equation Modeling

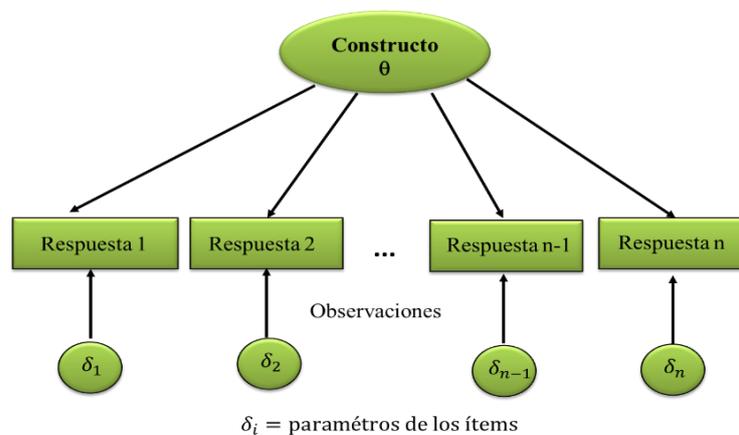
3.4. Nuevos métodos multivariantes propuestos para la detección del *Response Shift* en calidad de vida relacionada con la salud

3.4.1. El modelo de la Teoría de Respuesta al Ítem

La Teoría de Respuesta al Ítem (TRI) es otra forma de relacionar las medidas observadas y un constructo basada en la relación de las puntuaciones de cada uno de los ítems a través de los patrones de las respuestas, lo que permite obtener una estimación del valor del sujeto en el constructo (Martínez-Arias, 1995).

De acuerdo con lo planteado por Atorresi et al.(2009), siguiendo las líneas de Rasch (1960) y Birnbaum (1968), la TRI establece una relación entre el comportamiento de un sujeto frente a un ítem y el rasgo responsable de esta conducta, lo que se denomina rasgo latente. El mismo se basa en la probabilidad de obtener una determinada respuesta a un ítem para cada nivel de rasgo medido por este, donde su principal objetivo es la construcción de un instrumento de medición con propiedades invariantes entre la población. La Figura 16 muestra una representación conceptual de la formulación de este modelo.

Figura 16. Representación del modelo de la Teoría de Respuesta al Ítem

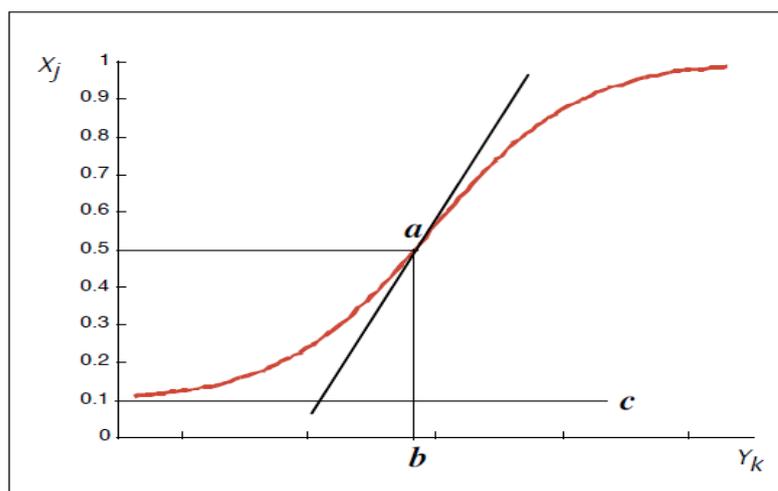


3.4.1.1. Características generales y supuestos de los modelos de la Teoría de Respuesta al Ítem

Las características más importantes de los modelos de la TRI son:

- La TRI parte de la premisa de que el comportamiento de un sujeto ante un ítem puede explicarse en términos de una o varias características denominadas rasgos o aptitudes latentes, las cuales no pueden observarse directamente.
- La relación entre el rendimiento o la conducta de un individuo en un ítem y el conjunto de rasgos responsables de dicho rendimiento se describe mediante una función monótona creciente denominada Curva Característica del Ítem (Figura 9). Mediante esta curva se representa una relación funcional entre la proporción de respuestas correctas a un ítem y el nivel del atributo, y representa la probabilidad de contestar una categoría de respuesta de un ítem i a un determinado nivel, siendo los rangos de esta probabilidad entre 0 a 1, es decir, $0 < P_i(\theta) < 1$.
- Los rasgos que mide un ítem determinarán la especificación de la dimensionalidad. Los determinantes de las respuestas son un conjunto de k rasgos que estará representado por un vector: $\theta' = [\theta_1, \theta_2, \dots, \theta_k]$
- Todo sujeto y todo ítem puede situarse en un punto en el espacio del rasgo o atributo, en este sentido, cada individuo se caracteriza por tener una posición θ_s en el espacio que representa el valor del rasgo.
- Las unidades básicas de análisis serán los ítems de la prueba y los modelos matemáticos de la TRI se establecerán por el nivel de ítem y no de la prueba completa.

Figura 17. Representación de los parámetros de la Curva Característica del Ítem



Los supuestos de los modelos de la TRI son:

- Se asume la existencia de una variable no observada, lo que explica la respuesta de las personas a una prueba o ítem.
- Debe existir la unidimensionalidad del ítem, esto refiere a que el ítem mide un solo rasgo.
- Existe una independencia local, esto supone que manteniendo constante la aptitud (θ) que explica el rendimiento en la prueba, las respuestas de los individuos a un par de ítems son estadísticamente independientes. (Attorresi et al., 2009)

3.4.1.2. Modelo logístico de dos parámetros

Bajo el supuesto de modelos unidimensionales y respuestas politómicas, se proponen el Modelo de Respuesta Nominal (Bock, 1997), el Modelo de Crédito Parcial (Masters, 1982; Wright, 1997), el Modelo de Crédito Parcial Generalizado (Muraki, 1992) y el Modelo de Respuesta Graduada, éste último es una extensión del modelo logístico de dos parámetros y basado en las diferencias entre las funciones de respuestas categóricas (Samejima, 1969).

Para un modelo logísticos de dos parámetros, la ecuación de proporcionalidad es la probabilidad de respuesta correcta en un ítem dicotómico y está dada de la forma:

$$P_i(\theta) = \frac{1}{1 + e^{-a_i(\theta - b_i)}}$$

Donde $P_i(\theta)$ es la probabilidad de responder correctamente a un ítem i , dado un nivel de aptitud θ_a y a_i es el parámetro de discriminación del ítem y b_i es el parámetro de dificultad del ítem.

3.4.1.3. Aplicación de los modelos de la Teoría de Respuesta al Ítem en la detección de *Response Shift*

El uso de los modelos de la TRI para examinar *Response Shift* involucra probar los cambios existentes en las relaciones entre los ítems discretos y los factores de rasgos latente en términos de una función de enlace logística (Sajobi et al., 2018). Los dos principales enfoques de la TRI para la detección del *Response Shift* son los siguientes:

- El modelo logístico lineal con supuestos relajados (Anota et al., 2014)
- El modelo ROSALI (Guilleux et al., 2015)

Anota et al. (2014) señala que la TRI puede ser considerada para explorar los efectos de Response Shift y, contrario a los modelos de ecuaciones estructurales, éstos no están basados en las ponderaciones observadas sino en las respuestas directas de los ítems aplicados a un individuo mediante un cuestionario. En consecuencia, las respuestas de los ítems en la TRI juegan un significativo y la relación entre la respuesta a los ítems y sus rasgos latentes no manifiestan comportamientos lineales.

Para ello, se propone el modelo logístico lineal con supuestos relajados (*Lineal Logistic with Relaxed Assumptions*, por sus siglas en inglés LLRA) el cual permite detectar el *Response Shift* de recalibración en una evaluación de datos retrospectivos a lo largo de dos o más mediciones (Anota et al., 2014).

Por su lado, [Guilleux et al. \(2015\)](#) propone el algoritmo ROSALI (siglas en inglés para *RespOnse Shift ALgorithm in Item response theory*) que utiliza los modelos TRI de tipo crédito parcial longitudinal y crédito parcial graduado como modelos de medición para evaluar el *Response Shift* de recalibración uniforme y no uniforme, así como de repriorización. ROSALI se basa en la unidimensionalidad de los modelos TRI, por lo que no se puede utilizar para detectar reconceptualización lo cual requiere un contexto multidimensional.

Los diferentes pasos para el algoritmo son los siguientes:

- Estimar las dificultades del ítem en la primera ocasión de medición
- Establecer de un modelo de medición
- Ajustar un modelo sin *Response Shift* y la evaluación general del *Response Shift*
- Detección del *Response Shift*
- Estimación del cambio real ([Guilleux et al., 2015](#))

Entre las ventajas de la utilización de los modelos TRI para la detección de *Response Shift*, se destacan:

- Permite su detección a nivel del ítem.
- No se basa en supuesto de normalidad debido a que usa modelo de medición adaptado a elementos categóricos.
- Puede utilizarse para estimar el tamaño del efecto en *Response Shift*.
- La LLRA no requiere unidimensionalidad de ninguno de los elementos, ni multidimensionalidad para estimar la calibración del *Response Shift* en los datos.
- Puede utilizar con múltiples métodos de prueba para controlar el error Tipo I asociado con pruebas múltiples de hipótesis de *Response Shift*.

Tabla 19. Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del *Response Shift*, utilizando los Modelos de la Teoría de Respuesta al Ítem (2010-2020)

No	Año de Publicación	Autor Principal	Título de la investigación
1	(2014)	Anota et al.	Item response theory and factor analysis as a mean to characterize occurrence of response shift in a longitudinal quality of life study in breast cancer patients
2	(2015)	Guilleux et al.	RespOnse Shift ALgorithm in Item response theory (ROSALI) for response shift detection with missing data in longitudinal patient-reported outcome studies
3	(2016)	Blanchin et al.	The Guttman errors as a tool for response shift detection at subgroup and item levels
4	(2016)	Sébille et al.	Prospective, multicenter, controlled study of quality of life, psychological adjustment process and medical outcomes of patients receiving a preemptive kidney transplant compared to a similar population of recipients after a dialysis period of less than three years – The PreKit-QoL study protocol
5	(2017)	Jabrayilov et al.	Longitudinal measurement invariance of the Dutch Outcome Questionnaire-45 in a clinical sample
6	(2020)	Schwartz et al.	Does response shift impact interpretation of change even among scales developed using item response theory?

De igual manera, estos modelos presentan algunas desventajas:

- Requiere de muestras de grandes tamaños como el número de elementos para ayudar a precisar la estimación de los parámetros a nivel de ítem.
- Se requerirá un enfoque de TRI multidimensional para detectar reconceptualización.
- El modelo LLRA se desarrolló utilizando datos de pruebas retrospectivas, pero no han sido probado con datos longitudinales secundarios.

La Tabla 19 presenta un conjunto de investigaciones recientes que aplican los modelos de la TRI para la detección del *Response Shift*.

3.4.2. Modelos de Efectos Mixtos

Los modelos de efectos mixtos se fundamentan en el estudio de la relación lineal entre dos variables donde se ha realizado muestreo por agrupación (*clúster sampling*) (Goldstein, 1986).

La aplicación de este método se da cuando en una investigación existen, además de los sujetos, otras unidades de análisis permitiendo la existencia de agrupaciones naturales que queden reflejada en la estructura de los datos (Oliver et al., 2000). Este análisis modela explícitamente la relación jerárquica de los datos eliminando los sesgos y proporcionando estimaciones sobre la variabilidad y replicabilidad de los coeficientes de regresión en los distintos contextos de investigación y sobre una influencia de estos en el comportamiento del individuo.

A pesar de que el campo de aplicación de este modelo ha sido fundamentalmente en las ciencias sociales y en la educación, se ha observado de manera progresiva la incorporación del modelo en el área de salud. Un ejemplo de ello es un estudio desarrollado por Phillips (2015) donde examinan el desempeño de un cuestionario sobre CVRS y las puntuaciones obtenidas para detectar el cambio en calidad de vida.

Se ajustó un modelo de efectos mixtos para una escala global de calidad de vida y cada una de las escalas funcionales que permitieran estimar las tendencias a lo largo del tiempo. Se esperaba que la escala global de calidad de vida y las escalas específicas mostraran una tendencia a disminuir con el tiempo en esta población de estudio (Phillips et al., 2015).

El modelo se plantea de la siguiente manera:

$$y_{ij} = (b_0 + \beta_{0i}) + (b_1 + \beta_{1i})periodo_{ij} + e_{ij}$$

donde,

b_0 y b_1 indican el nivel medio de los grupos de calidad en la línea base y ponderan el cambio de la media de la calidad de vida por periodo

β_{0i} y β_{1i} indican la desviación del paciente i desde el intercepto y pendiente, respectivamente, comparado con el promedio del grupo.

En el contexto de datos en CVRS, esta metodología modela los datos longitudinales en las ponderaciones globales en función a las ponderaciones de dominio. De esta forma, las pruebas significativas de las interacciones de los efectos se utilizan como evidencia de *Response Shift* (Sajobi et al., 2018).

La principal ventaja de aplicar los modelos de efectos mixtos está que la variación específica del sujeto se puede modelar además que puede adaptarse a datos longitudinales incompletos. Sin embargo, las limitaciones que tiene el método es que sólo se puede aplicar a datos obtenidos de un instrumento que mida la CVRS tanto dominio global como específico, además de que se basa en el supuesto de la distribución normal multivariante. La Tabla 20 presenta investigaciones que han aplicado el modelo de efectos mixtos para la detección de *Response Shift*.

Tabla 20. Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del *Response Shift*, utilizando el Modelo de Efectos Mixtos (2010-2020)

No.	Año de Publicación	Autor Principal	Título de la investigación
1	(2011)	Lacey et al.	Compared to what? A joint evaluation method for assessing quality of life
2	(2014)	Schwartz et al.	Fluctuations in appraisal over time in the context of stable versus non-stable health
3	(2015)	Phillips et al.	Summary scores captured changes in subjects' QoL as measured by the multiple scales of the EORTC QLQ-C30
4	(2017)	Tessier et al.	Does the relationship between health-related quality of life and subjective well-being change over time? An exploratory study among breast cancer patients
5	(2017b)	Salmon et al.	Identifying patterns of adaptation in breast cancer patients with cancer-related fatigue using response shift analyses at subgroup level
6	(2018)	Nichols et al.	Long-term changes in oral health-related quality of life of standard, cleft, and surgery patients after orthodontic treatment: A longitudinal study

3.4.3. Medidas de Importancia Relativa

El término de Importancia Relativa lo define [Kruskal y Majors \(1989\)](#) como el grado en que contribuye cada variable a la predicción de un criterio, es decir a la significancia estadística utilizada como medida de importancia. Por su parte, [Johnson y Lebreton \(2004\)](#) la define como la contribución proporcional que cada predictor que hace al R^2 , considerando tanto su efecto directo como su efecto cuando se combinan con las otras variables en la ecuación de regresión. Diferentes autores consideran estas medidas como estimaciones de ponderaciones o rasgos variables para grupos de discriminación o predicción ([Huberty & Wisenbaker, 1992](#); [Sajobi et al., 2011](#); [Thomas et al., 2008](#)). Mientras [Lix et al. \(2013\)](#), utiliza las medidas de importancia relativa para describir la capacidad que tienen las medidas estadísticas para discriminar entre grupos o predecir la pertenencia a un grupo.

La utilización de estas medidas desarrollada para datos que han sido recolectados en dos ocasiones es considerada para evaluar los dominios como resultados, los coeficientes de la función discriminante estandarizados ([Huberty & Wisenbaker, 1992](#)) y el coeficiente de razón discriminante ([Thomas, 1992](#)), y para los dominios como variables explicativas utiliza el coeficiente de regresión logística estandarizado ([Menard, 2004](#)) y el Índice de Pratt de regresión logística ([Thomas et al., 2008](#)).

Para el análisis de los datos de CVRS, [Lix et al. \(2013\)](#) propone dos métodos basados en las medidas de Importancia Relativa para detectar *Response Shift* de repriorización en múltiples dominios de CVRS. El primer método se basa en una prueba de cambio en las ponderaciones, y el segundo en una prueba de cambio en los rangos del dominio.

Entre los criterios identificados en la utilización de estas medidas, se recomienda no utilizar una única medida de manera uniforme, debido a que se obtienen diferentes atributos de importancia tales como:

- la sensibilidad a las diferentes características de los datos incluidas las correlaciones entre los dominios.
- las desviaciones de una distribución normal y a los efectos que ocurren cuando una variable contribuye directamente a otra.
- al tamaño de la muestra.

Lix et al. (2013) establece un conjunto de pasos para implementar métodos que permiten detectar *Response Shift* de repriorización. El análisis inicia preparando la base de datos, esto involucra investigar los datos para los dominios altamente correlacionados, los cuales pueden ser excluidos y observaciones faltantes pueden ser imputadas. Los métodos pueden resultar en un caso de eliminación de observaciones faltantes a menos que se adopte un método de imputación. Los métodos aplicados por el autor están basados en la prueba de cambio en las ponderaciones y en la prueba de cambio en los rasgos del dominio.

3.4.3.1. Métodos 1: Método basado en la prueba de cambio en las ponderaciones

Los pesos obtenidos a partir de los datos centrados en cada ocasión de medición son denotados por W_{k1} y W_{k2} para k -ésimo dominio de CVRS ($k= 1, 2, \dots, m$) en las ocasiones 1 y 2 respectivamente. Se plantea la hipótesis de que no existen cambios en los pesos k -ésimo dominio frente a la hipótesis alternativa de que existe cambio en los pesos k -ésimo dominio; es decir:

$$H_{Owk}: d_k = 0; \quad \text{donde } d_k = W_{2k} - W_{1k}$$

$$H_{Awk}: d_k \neq 0$$

La distribución muestral de d_k se aproxima usando el método de *bootstrap*, donde los Y_{ij} denota el vector $2m \times 1$ de las puntuaciones de dominio para el i -ésimo participante del estudio ($i=1, \dots, n$) en el j -ésimo grupo ($j=1,2$) se muestrean aleatoriamente con reemplazo en cada grupo para generar B muestras cada una con tamaño N donde $N = n_1 + n_2$

El valor p utilizado para probar H_{Owk} es aproximado por la proporción de los valores de la distribución muestral empírica dada por:

$$p_{dk} = \frac{1}{B} \left[\sum_{b=1}^B I(d_{bk}^* > d_k) + \sum_{b=1}^B I(d_{bk}^* < -d_k) \right]$$

Donde I es la función indicadora que toma valor de 1 si la desigualdad es verdadera y 0 si la desigualdad es falsa y d_{bk}^* es la estadística de diferencia para la b -ésima muestra ($b=1,\dots,B$). El rechazo de la hipótesis nula indica presencia de *Response Shift* de repriorización en el k dominio.

Para controlar la tasa de error Tipo I, se adopta el procedimiento de [Dunn \(1961\)](#) y [Hochberg \(1988\)](#) que plantean la hipótesis nula específica de cada dominio y prueban la significancia con $\alpha^* = \alpha/m$.

3.4.3.2. Método 2: Prueba de cambio en los rangos del dominio

Esté método se basa en los rangos de importancia relativa utilizando la siguiente ecuación para probar la diferencia entre los rangos:

$$q_k = \frac{r_{k1} - r_{k2}}{m - 1}$$

Siendo r_{k1} y r_{k2} los rangos de k -ésimo dominio en los momentos 1 y 2 respectivamente. Los momentos están clasificados por separado con un rango de 1 que indican la clasificación más alta, es decir, el más importante dominio para la discriminación de grupo.

Por otro lado, los rangos se asignan en base a los valores absolutos de los pesos de los coeficientes de la función discriminante estandarizados, los coeficientes de regresión logística estandarizados, los pesos del coeficiente de razón discriminante o el índice de Pratt para regresión logística.

Se plantea la siguiente hipótesis de que no existe cambio en los rangos de k -ésimo dominio:

$$H_{ORk}: q_k = 0$$

$$H_{ARk}: q_k \neq 0$$

La importancia relativa de los rangos r_{k1}^* y r_{k2}^* se calcula para cada muestra junto con la puntuación de diferencia q_k .

El valor p utilizado para probar H_{ORk} es aproximado por la proporción de los valores de la distribución muestral empírica dada por:

$$p_{qk} = \frac{1}{B} \left[\sum_{b=1}^B I(q_{bk}^* > q_k) + \sum_{b=1}^B I(q_{bk}^* < -q_k) \right]$$

donde q_{bk}^* es la diferencia estadística de la muestra.

Entonces, se rechaza la H_{ORk} de p_{qk} si es menor o igual que el nivel de significancia seleccionado. Por lo que el rechazo de H_{ORk} indica presencia de *Response Shift* de repriorización en el k de los dominios de la CVRS. La Tabla 21 presenta investigaciones orientadas a la detección de *Response Shift* utilizando las medidas de Importancia Relativa.

Tabla 21. *Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift, utilizando las Medidas de Importancia Relativa (2010-2020)*

No.	Año de Publicación	Autor Principal	Título de la investigación
1	(2013)	Schwartz et al.	Changing values, changing outcomes: the influence of reprioritization response shift on outcome assessment after spine surgery
2	(2014)	Sajobi et al.	Changes in quality of life after epilepsy surgery: The role of reprioritization response shift
3	(2015)	Sajobi et al.	Identifying reprioritization response shift in a stroke caregiver population: a comparison of missing data methods
4	(2015)	Mayo et al.	How are you? Do people with inflammatory bowel disease experience response shift on this question?

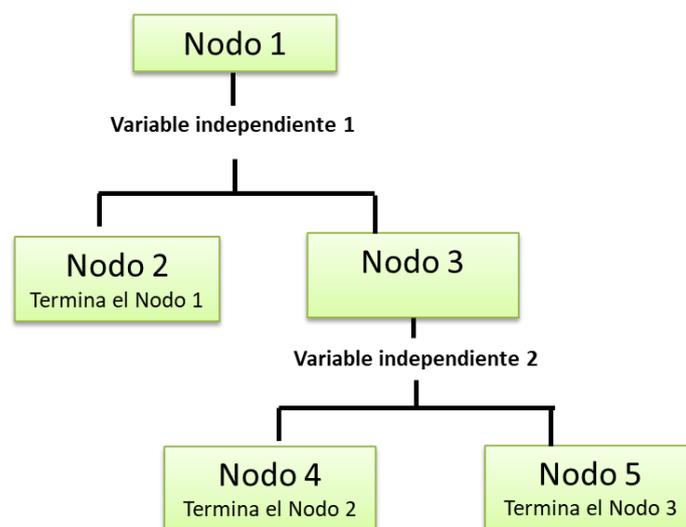
3.4.4. Método de Clasificación y Regresión de Árbol

El método de Clasificación y Regresión de Árbol (CRT por las siglas en inglés de *Classification and Regression Trees*) desarrollado por Breiman et al. (1984) se encuentra identificado en la literatura como CART, CRT, C&RT, RPART, RTA. Este

método consiste en una división binaria que particiona de forma recursiva el conjunto de datos en subgrupos disjuntos llamados hojas. Este método de regresión no paramétrico basado en la partición recursiva de la data en subgrupos homogéneos (Sajobi et al., 2018), y a su vez, tiene la capacidad de segmentar poblaciones de manera eficiente en subgrupos significativos (Lemon et al., 2003).

El CRT se basa en un algoritmo que divide iterativamente el conjunto de datos en dos submuestras de acuerdo con una regla binaria: $PCS < 50$. Esta regla de división se basa en que la elección de las variables explicativas se encuentra dentro de un umbral para esta variable, y la elección se realiza de manera que minimice la heterogeneidad de las submuestras. Los árboles de regresión se construyen utilizando el criterio de "desviación". Las dos submuestras obtenidas son particionadas de la misma manera hasta que existan pocas observaciones (generalmente cinco) en las muestras obtenidas. Este procedimiento produce un árbol que puede tener muchos nodos terminales. El valor medio de la variable de salida se asigna a cada hoja calculada sobre las observaciones dentro de la región correspondiente. Por lo que un árbol de clasificación divide el espacio X y proporciona un valor estimado. Para ello, $Pr(Y = s | X \in A_k)$ en cada región. Su representación conceptual se muestra en la Figura 18.

Figura 18. Modelo de Clasificación y Regresión de Árbol



Una de las ventajas del CRT es que requiere de una muestra pequeña con un mínimo de 10 eventos por variables para obtener una predicción razonable modelando un rendimiento estable (van der Ploeg et al., 2014)

El método de CRT crea un árbol de decisiones utilizando la selección de variables paso a paso para identificar subgrupos exclusivos de una población. De acuerdo con lo planteado por Breiman et al. (1984) y Lemon et al. (2003) la metodología para su elaboración es la siguiente:

- El árbol actúa como una representación con nodos terminales (hojas) que representa una celda de la partición, cada una con un simple modelo que se aplica sólo a esa celda.
- Cada nodo está dividido a través de la mejor variable, maximizando la pureza de los nodos resultantes. Un nodo se considera "puro" cuando todos los casos tienen el mismo valor para la variable dependiente.
- Si falta la variable principal de división para una observación individual, se considerará como variable principal la que tiene el mejor patrón similar relativo a la variable resultado.

Para evaluar la confiabilidad del árbol, Blockeel y Struyf (2003) sugieren realizar una validación cruzada de 10 veces, donde el conjunto de datos se divide en 10 partes seleccionadas al azar y aproximadamente iguales con cada parte que contiene una distribución similar de datos. Las primeras nueve partes de los datos (90%) se utilizan para construir el árbol más grande posible, y el 10% restante se utiliza para obtener estimaciones iniciales de la tasa de error del seleccionado un subárbol. El proceso se repite 10 veces usando diferentes combinaciones de los nueve subconjuntos de datos restantes y un subconjunto de datos 1/10 diferente para probar el árbol resultante. Los resultados de las 10 pruebas se combinan luego para calcular las tasas de error para árboles de cada tamaño posible y es aplicado para podar el árbol completo.

Li y Rapkin (2009) proponen utilizar el método de CRT con particiones recursivas para explorar los cambios cognitivos y así explicar el fenómeno de *Response Shift* en la CVRS de manera que fueran interpretables y consistente, demostrando que, a pesar

de que las particiones se realizan utilizando el método CTR, estas son arbitrarias provocando errores que caerían como cascadas resaltando la importancia de la poda de árboles mediante validación cruzadas (Radespiel-Troger et al., 2003). Los autores establecen que el método permite obtener una descripción del fenómeno de *Response Shift*.

La Tabla 22 muestra las investigaciones que han sido publicada en la última década utilizando este método CTR para la detección del *Response Shift*.

Tabla 22. *Publicaciones más relevantes de calidad de vida relacionada con la salud para la detección del Response Shift, utilizando el Método de Clasificación y Regresión de Árbol (2010-2020)*

No.	Año de Publicación	Autor Principal	Título de la investigación
1	(2011)	Li & Schwartz	Data mining for response shift patterns in multiple sclerosis patients using recursive partitioning tree analysis
2	(2011)	Schwartz et al.	Response shift in patients with multiple sclerosis: an application of three statistical techniques
4	(2017)	Machuca et al.	Using classification and regression tree modelling to investigate response shift patterns in dentine hypersensitivity
5	(2020)	Machuca et al.	How peoples' ratings of dental implant treatment change over time?

3.4.5. Método STATIS Dual como herramienta para evaluar *Response Shift*

Vicente-Galindo (2003) propone por primera vez la utilización de los métodos de la familia STATIS (*Structuration de Tableaux A Trois Indices de la Statistique*) (Escoufier, 1976a; L'Hermier des Plantes, 1976) como herramienta para detectar *Response Shift* en estudios de autopercepción, concretamente en estudios de CVRS, partiendo del hecho de que estos métodos llevan a cabo el análisis estadístico de las intra-estructuras definidas como productos escalares de matrices, para filtrar el “ruido” contenido en los datos y capturar la información estadísticamente significativa.

En su trabajo afirma que, si el objetivo fundamental del estudio fuese analizar las similitudes de los individuos a lo largo del tiempo, o cuando éstos están sometidos a distintos tratamientos, debe utilizarse el método STATIS; y si el objetivo está en analizar la estabilidad, o la evolución, de las estructuras de covariación entre los ítems, se utilizará el STATIS Dual.

Vicente-Galindo (2003) prueba que los datos de CVRS longitudinales, pueden ser entendidos como datos de Tres Vías, el primer modo se refiere a los pacientes, el segundo a los ítems del cuestionario utilizado para evaluar CV, y el tercero se refiere a las distintas ocasiones en las que la CV se evalúa.

Prueba, asimismo, que el *Response Shift* puede ser evaluado comparando configuraciones de covariación (o de similitud) a partir del producto de Hilbert-Schmidt, ya que a partir de este producto interno es posible definir una norma y una distancia de Hilbert-Schmidt, entre configuraciones. Si la correlación vectorial entre dos matrices es igual a la unidad, eso significa que ambas matrices son equivalentes en el sentido de que ambas estructuras son congruentes y comparables; es decir, no aparecerían discrepancias entre las estructuras factoriales de los dos (o más) tiempos. Por tanto, si la correlación vectorial entre las matrices factoriales de dos tiempos (en el sentido de Hilbert Schmidt), es igual a la unidad ambas estructuras serían coincidentes; cuanto más próximo a uno, más similares las estructuras.

El STATIS Dual (también el STATIS) está basado en el producto escalar de Hilbert-Schmidt (HS) para inducir la distancia entre las configuraciones, por tanto, puede ser utilizado para detectar estabilidad o cambio (*Response Shift*).

Una vez calculada la correlación vectorial de Hilbert-Schmidt, entre los operadores que capturan las matrices de covariaciones entre las variables, en cada matriz de datos, se somete esa matriz de correlaciones Hilbert-Schmidt a un Análisis de Componentes Principales. De esta forma, cada matriz de datos de partida correspondiente al estudio de CV percibida en un tiempo puede ser representada sobre el subespacio de máxima inercia como un punto en un espacio vectorial de baja dimensión en el cual la proximidad entre puntos se interpreta en términos de similitud entre las estructuras de covarianza de las variables y, por tanto, en términos de congruencia entre las estructuras

factoriales. Así se consigue reducir la dimensionalidad y representar cada matriz de datos como un punto en el subespacio euclídeo de baja dimensión en donde se reproducen las interdistancias del hiperespacio que estaba estructurado con la distancia de Hilbert-Schmidt.

Es posible construir la matriz compromiso que pueda proyectar los ítems de las matrices originales para analizar sus trayectorias a lo largo del tiempo. Esto puede ayudar a detectar los ítems implicados en la reconceptualización y en la recalibración.

Siguiendo a Vicente-Galindo y Galindo-Villardón, en su artículo publicado en 2009 en la Revista de *Matemática: Teoría y Aplicaciones*, podemos afirmar que:

- Si en la imagen euclídea se observa que los ángulos que forman los vectores que unen el origen con los puntos que representan a las matrices son muy agudos, eso significa que las estructuras factoriales de las diferentes matrices son muy similares, por tanto, hay razones para suponer que no se ha producido una reconceptualización con el paso del tiempo. Dicho de otra manera, si se ha producido un cambio Gamma a lo largo del tratamiento, la representación euclídea de las matrices debe exhibir algún ángulo grande, al menos alguno de los puntos de la configuración euclídea debe estar alejado del resto.
- Si además en la representación euclídea, las normas de los vectores son iguales, tampoco se ha producido una recalibración. Esto supone una alternativa al *Then-Test* propuesto por [Howard y Dailey \(1979\)](#) para evaluar la componente de recalibración del *Response Shift*, basadas en el producto escalar y la distancia de Hilbert-Schmidt.
- Si en la representación euclídea permite corroborar la estabilidad en el tiempo, se puede calcular una matriz compromiso (media) que nos dé una visión global de las dimensiones latentes del constructo Calidad de Vida, a lo largo de todo el tratamiento.
- Si la representación euclídea pone de manifiesto que alguna de las estructuras es diferente y, por tanto, que cabe suponer que la concepción del constructo ha cambiado, no podremos comparar los resultados antes y después del

tratamiento para cuantificar el cambio, Alpha ya que está enmascarado por el cambio Gamma.

- Tampoco es posible si se ha detectado recalibración ya que en ese caso tampoco será interpretable el cambio entre el pre y el post tratamiento, ya que estará sesgado por el cambio Beta.
- Sobre la matriz compromiso se pueden proyectar los ítems que conforman el cuestionario para la evaluación de la Calidad de Vida, de las matrices originales para analizar sus trayectorias a lo largo del tiempo. Esto nos puede ayudar a detectar los ítems implicados en la reconceptualización y/o en la recalibración.

Para contrastar las alternativas propuestas, [Vicente-Galindo \(2003\)](#) estudia la Calidad de Vida en 504 pacientes Osteoporóticos en un estudio multicéntrico de atención primaria, a través de la información recogida con un cuestionario genérico, el EUROQOL-5D y otro específico, el QUALEFFO. Más detalles sobre el método y sobre la aplicación pueden verse en [Vicente-Galindo \(2003\)](#) y en [Vicente-Galindo y Galindo-Villardón \(2009\)](#)

CAPÍTULO IV

METABILOT: UNA PROPUESTA PARA LA DETECCIÓN DE *RESPONSE SHIFT*

4. METABILOT: UNA PROPUESTA PARA LA DETECCIÓN DEL *RESPONSE SHIFT*

4.1. Métodos para detectar cambios en espacios consenso

Para la identificación del cambio [Golembiewski et al. \(1976\)](#) señala que la congruencia de las estructuras factoriales en los diferentes momentos en que se evalúan los pacientes podría definirse como la ausencia de un cambio, es decir, que existe los mismos factores latentes y existe el mismo patrón de ítems sobre los factores.

Según el autor, la ausencia de cambio se puede constatar como congruencia de las estructuras factoriales en los diferentes tiempos del estudio. Sin embargo, esta afirmación debe ser revisada ya que existen diversas formas en las cuales dos estructuras factoriales son congruentes. Para ello surgen las siguientes inquietudes con respecto a las estructuras factoriales en dos tiempos: ¿ambas estructuras requieren el mismo número de factores?, ¿requieren disponer el mismo orden de factores?, ¿los niveles de absorción de inercia requieren ser las mismas?, ¿los factores de carga de cada eje requieren ser los mismos y poseer la misma magnitud?

Para [Golembiewski et al. \(1976\)](#) dos estructuras son iguales cuando tienen el mismo número de factores, y los factores están configurados por los mismos factores de carga, pero no es necesario que la magnitud de los factores de carga se mantenga. Por tanto, para constatar la ausencia de cambio debe existir los mismos factores latentes y existir el mismo patrón de ítem sobre los factores.

[Krzanowski \(1979\)](#) propone un procedimiento para la comparación de subespacios obtenidos por componentes principales los que se generalizan en encontrar los componentes principales de cada espacio, comparar cada componente de uno de los espacios con todas las del espacio correspondiente a la segunda matriz, seguido de localizar la más similar y calcular el coseno del ángulo que forman esas dos componentes más similares y luego hacer el mismo procedimiento para cada una. El propósito es lograr la comparación de forma simultánea del conjunto de matrices y determinar las congruencias entre las estructuras factoriales.

El estudio de la congruencia de varias matrices de datos es similar y el procedimiento no posee limitaciones en el número de matrices que son comparadas. Si las soluciones son congruentes, es condición suficiente para afirmar que no se ha producido una reconceptualización ya que las dimensiones y los patrones del constructo son los mismos.

De esta manera, si las soluciones multidimensionales son similares, es posible calcular una solución consenso que permita representar las matrices y evaluar la estructura multifactorial para todo el periodo de estudio. En caso contrario, si las soluciones no son globalmente congruentes, se podrá analizar la discrepancia y conocer los ítems en los que se ha producido cambio.

En sentido, se presenta la propuesta de [Vicente-Galindo \(2003\)](#) en el contexto del STATIS Dual, pero para que la misma proporcione resultados óptimos, las estructuras latentes a comparar deben tener una primera dimensión dominante en todos los momentos del tiempo y esa dimensión debe tener el mismo patrón factorial en todos los casos. Si esto no ocurre, cosa bastante frecuente en estudios de CVRS, donde a menudo se encuentra el llamado efecto techo y/o efecto suelo, las configuraciones serían declaradas como incongruentes con esta alternativa.

Los elementos presentados por [Golembiewski et al. \(1976\)](#), ampliados por [Krzanowski \(1979\)](#) y con una nueva propuesta por [Vicente-Galindo \(2003\)](#) plantean que la detección del cambio mediante el análisis de estructuras factoriales y su comportamiento dentro de un espacio consenso con fundamentos para poder demostrar la capacidad de detección de *Response Shift* con datos de CVRS. La incorporación de un método ya desarrollado de tipo multivariante con tablas de tres vías, pero no explorado en este contexto, propone una oportunidad para superar las limitaciones de los modelos anteriores.

4.2. MetaBiplot en la detección de *Response Shift*

El MetaBiplot propuesto por [Martín-Rodríguez \(1996\)](#) y [Martín-Rodríguez, et al. \(2002\)](#) es una técnica útil para el tratamiento de información proveniente de varias matrices de datos, es un métodos de análisis de datos que han sido desarrollados para

extraer la información relevante que se almacena en tablas de datos de tres vías y permite la obtención de un subespacio consenso que integra las configuraciones resultantes de varios análisis, siendo el MetaBiplot una técnica que compara e integra varios análisis Biplot.

La obtención del subespacio consenso tiene como propósito identificar el conjunto de ejes ortogonales que mejor capture la congruencia entre las subdimensiones latentes. Dicho de otra manera, un subespacio consenso permitirá identificar diferencias y similitudes entre los elementos o variables estudiadas en un mismo espacio de representación.

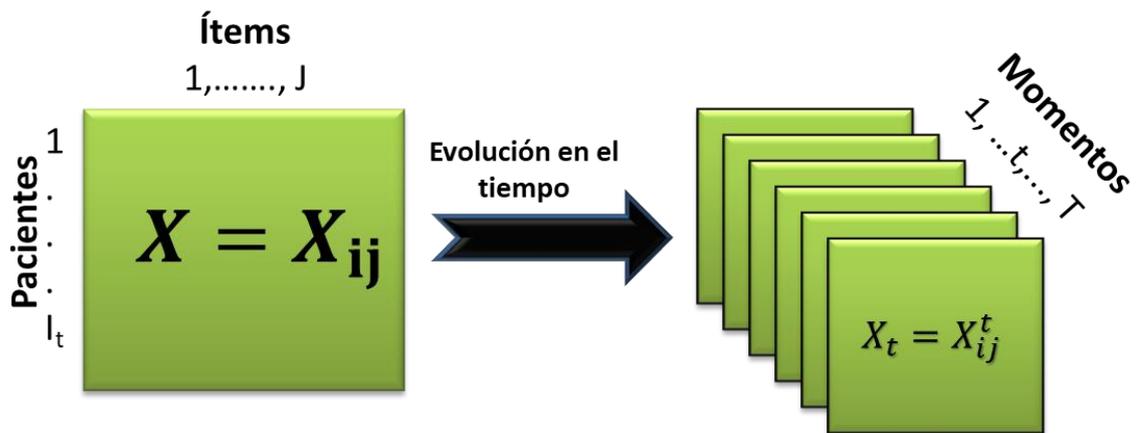
El MetaBiplot ha resultado ser una técnica utilizada en diferentes áreas de estudio. Se han identificado diferentes investigaciones como las de [Martín-Rodríguez y Vicente-Villardón \(2004\)](#) enfocada a la biología marina; [Nunes y Galindo-Villardón \(2006\)](#) en estudios de calidad del agua, y [Caballero-Juliá et al. \(2017\)](#) en ciencias sociales manejando datos textuales y análisis cualitativos. En consecuencia, la versatilidad del MetaBiplot sugiere una oportunidad para explorar criterios que permitan ser considerados en el análisis de datos de calidad de vida.

4.2.1. MetaBiplot en el contexto de Calidad de Vida

El MetaBiplot es un método multivariante para tablas de datos de tres vías que puede ser utilizado como método para la detección del *Response Shift* basado en los planteamientos de [Golembiewski et al. \(1976\)](#), ya que los datos de calidad de vida relacionado con la salud pueden ser considerados como datos de tres vías, por lo cual se aplican los siguientes principios:

- Los datos de calidad de vida relacionados con la salud son datos de tres vías, ya que constan de tres elementos donde el primer modo representa los pacientes, el segundo modo los ítems y el tercer modo los distintos momentos en que se evalúa la calidad de vida, tal como se muestra en la Figura 19.

Figura 19. Representación de los datos de Calidad de Vida: Pacientes (**I**), Ítems (**J**) y momentos (**T**) como datos de Tres Vías



Siendo X_t representa la matriz de datos en el momento específico t , la cual tendrá **I** filas para cada pacientes y **J** columnas una para cada ítem y la **T** representará el número de momentos en que será evaluado el estudio.

- La información para cada momento puede ser establecida según objetivo. Si el objetivo es evaluar las similitudes entre los pacientes, el análisis estará dado por el JK-Biplot o RMP-Biplot (*Row Metric Preserving*), donde se obtiene una alta calidad de representación para las pacientes.

Mientras que, si el objetivo es evaluar la evolución y estabilidad de las estructuras factoriales entre los ítems, el análisis estará dado por el GH-Biplot o CMP-Biplot (*Column Metric Preserving*), y se obtiene una alta calidad de representación para los ítems.

Como el objetivo de nuestra investigación es probar que el método MetaBiplot permite detectar el *Response Shift*, se utilizará el GH-Biplot o CMP Biplot ya que permitirá comparar las estructuras factoriales y las correlaciones en los diferentes momentos del estudio en un subespacio consenso.

- *Response Shift* puede ser evaluado a través de las correlaciones de las matrices y la variabilidad explicada por el consenso. Para comparar las correlaciones se utiliza el índice de similitud de los subespacios de los marcadores.

El ángulo de cada grupo con las direcciones medias se mide por:

$$\partial_{tj} = \cos^{-1} \left\{ (b_j' U_t U_t' b_j)^{1/2} \right\}$$

Una medida de la disimilaridad entre el vector b_1 y los subespacios de k -dimensión de cada grupo es:

$$\partial_t = \cos^{-1} \left\{ (b_1' U_t U_t' b_1)^{1/2} \right\}$$

La similaridad global de los subespacios se mide por:

$$\frac{\sum_{i=1}^k \lambda_i}{gk}$$

La similaridad de todos los subespacios con cada uno de los nuevos vectores principales b_t es medido por el siguiente índice:

$$\frac{\lambda_i}{g}$$

El porcentaje de variabilidad en cada componente explicado por cada uno de los grupos están en la matriz \mathbf{C} con $c_{ij} = \cos^2 \partial_{tj}$

El coseno de los ángulos entre los marcadores de filas de los subespacios originales y los subespacios de grupos se pueden interpretar como la similaridad entre ellos (coeficiente de correlación).

La proporción de variabilidad de cada grupo explicada por las componentes consenso, estará dada por:

$$\frac{b_j' X_t X_t' b_j}{\text{traza}(X_t X_t')}$$

La variabilidad de los grupos explicada por todas las componentes del consenso:

$$\frac{\text{traza}(B' X_t X_t' B)}{\text{traza}(X_t X_t')}$$

La variabilidad total explicada por el consenso:

$$\frac{\sum_{t=1}^g b_j' X_t X_t' b_j}{\sum_{t=1}^g \text{traza}(X_t X_t')}$$

Las contribuciones relativas del factor al elemento estarán contenidas en la matriz:

$$CF = \{r_{ij}^2\}$$

Por lo que se considera que, si la correlación entre dos matrices es igual a la unidad, se puede decir que ambas matrices son comparables, lo que indica que no hay diferencia entre las estructuras factoriales de las matrices en el tiempo.

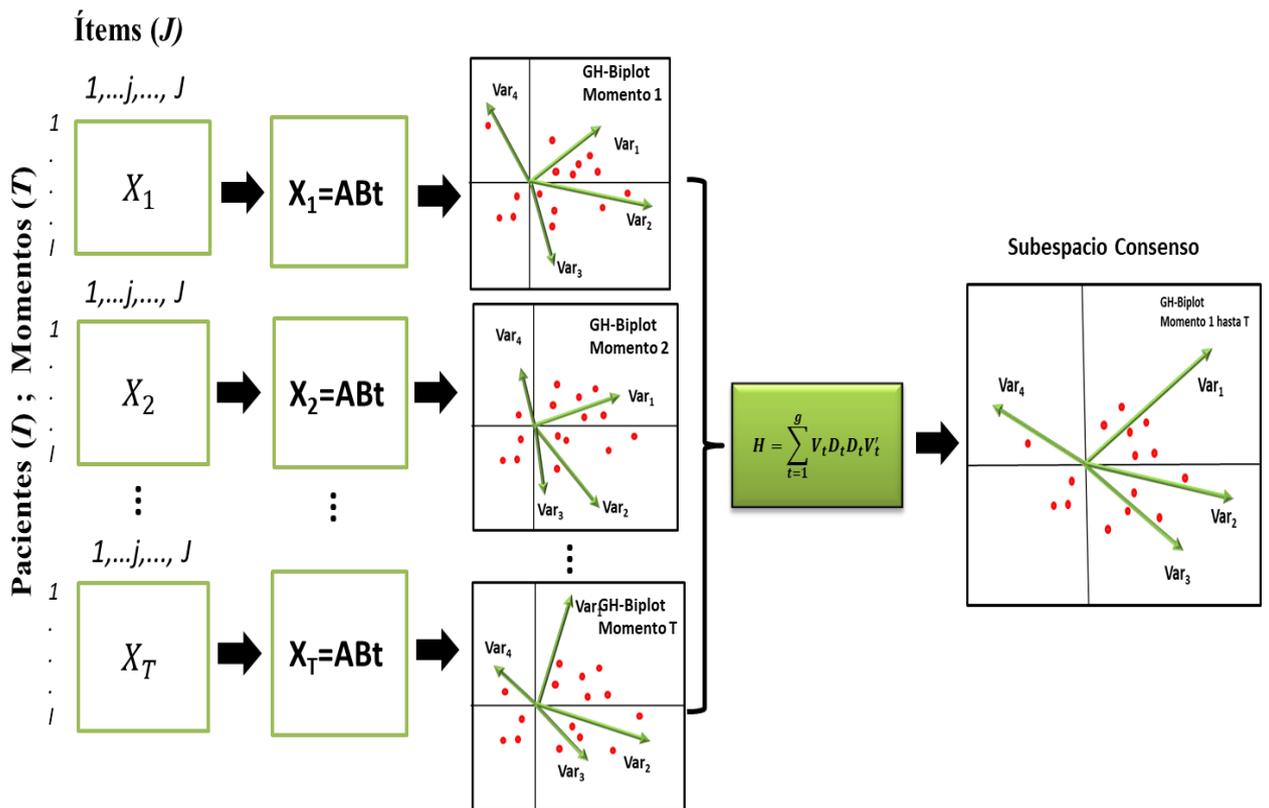
4.2.2. Método MetaBiplot

El MetaBiplot es un método que consiste en representar cada matriz de datos correspondiente a los diferentes momentos del estudio en un subespacio consenso de máxima calidad de representación que integre las configuraciones resultantes de varios

análisis con el objetivo de identificar el conjunto de ejes ortogonales que mejor capture la congruencia entre las subdimensiones latentes, y así interpretar las similitudes entre las estructuras de covariables de las variables y por tanto en términos de congruencia entre las estructuras factoriales. Al lograr una proximidad entre los puntos se podría decir que las matrices tienen la misma estructura de covariación entre los ítems.

Al evidenciar que las estructuras son comparables se puede obtener un subespacio consenso que represente todas las matrices correspondientes a cada momento evaluado en el estudio y sobre ella representar los datos originales permitiendo analizar el comportamiento de los ítems y su evaluación a lo largo del tiempo.

Figura 20. Proceso de generación del subespacio consenso, basado en el MetaBiplot



Fuente: Elaboración propia

La Figura 20 presenta el proceso de generación del subespacio consenso obtenido del MetaBiplot. El mismo inicia con la identificación de las matrices de variables e individuos en cada uno de los momentos, generando así la representación euclídea de cada uno de ellos mediante un Biplot. Con ello se obtiene la configuración consenso que integra los diferentes resultados de los Biplot, en un solo plano de máxima calidad de representación, denominado subespacio consenso.

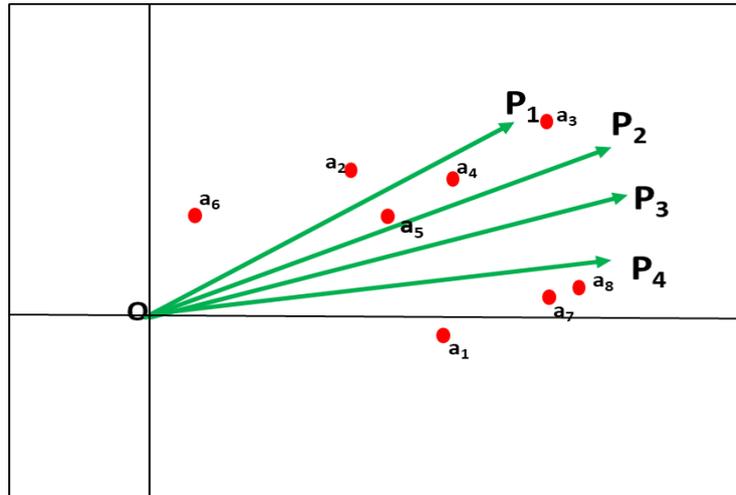
4.2.3. Evaluación del subespacio consenso

Tal como se muestra en la figura anterior (Figura 20), al contar con más de dos momentos, \mathbf{X}_T con $T > 2$, a los que se evalúa el mismo número de pacientes (\mathbf{I}) y los mismos ítems (\mathbf{J}) en cada uno de los (\mathbf{T}) momentos y aplicando la misma transformación en cada uno, se realiza un análisis Biplot que se obtiene del Análisis de Componentes Principales a partir de la Descomposición en Valores Singulares, logrando así reducir la dimensionalidad y representar cada matriz de dato en un subespacio euclídeo para luego integrarla en subespacio consenso.

Para interpretar la representación obtenida de un subespacio consenso, presentamos cuatro ejemplos hipotéticos de una matriz $\mathbf{X}_{(8 \times 4)}^4$ que recoge la información de los pacientes ($I=8$) a través de 4 ítems de un cuestionario de CVRS en diferentes momentos de evaluación ($T=4$). Considerando que \mathbf{a}_i representa los puntos y \mathbf{P}_j representa los vectores (Ítem), capturado en un subespacio consenso que recoge la información de los cuatros momentos evaluados.

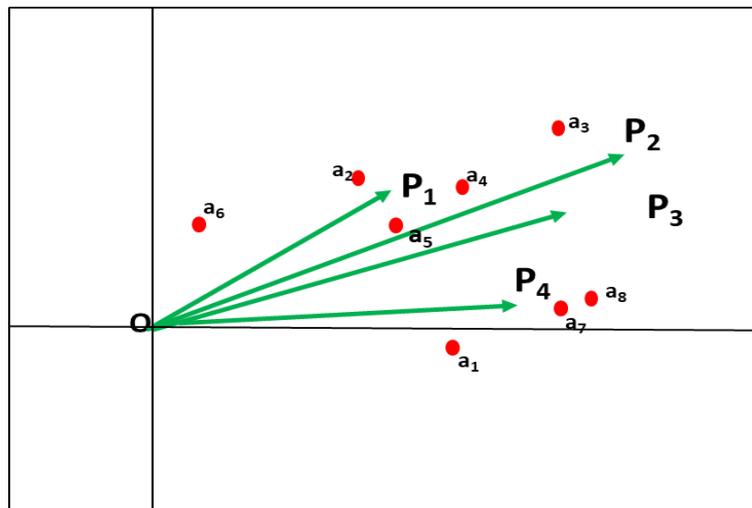
Se ha unido cada punto con el origen de coordenadas para obtener una estimación gráfica de las correlaciones entre las estructuras a través del ángulo que forman las líneas que unen los puntos que representan las matrices con el origen de coordenadas. Las Figuras 21, 22, 23, y 24 presentan diferentes escenarios sobre el comportamiento de las variables dentro del subespacio consenso.

Figura 21. *Caso 1: Existencia de una estructura común*



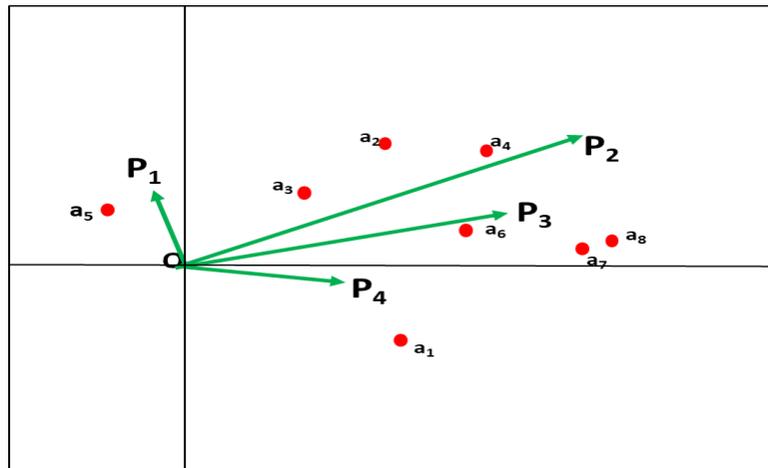
En el *caso 1*, se observa como los cuatro ítems evaluados están bien representado en el subespacio consenso, donde los vectores tienen similares normas y próximos entre sí formando ángulos pequeños, presentando elevados coeficientes de asociación. Por lo que se podría considerar que existe una estructura de covariación estable.

Figura 22. *Caso 2: Estructura común con normas diferente*



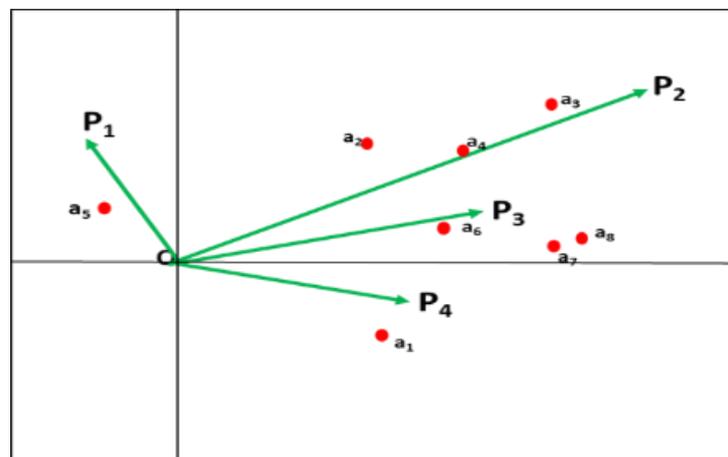
En el *caso 2*, se puede observar que los vectores forman ángulos pequeños sugiriendo elevados coeficientes de asociación, lo que es indicativo que se presenta las mismas estructuras de covariables, aunque no se presente igual normas.

Figura 23. *Caso 3: Vectores mal representados*



En el *caso 3*, se presenta vectores que no quedan bien representados (P_1 y P_4) en el subespacio consenso además de presentar normas diferentes y formar ángulos grandes entre ellos, por lo que se puede indicar la existencia de incongruencia en sus estructuras factoriales.

Figura 24. *Caso 4: Ausencia de una estructura común*



En el *caso 4*, los cuatro ítems son diferentes entre sí, ya que presentan diferentes normas y ángulos grandes, por lo que se puede concluir que no existe una estructura común entre las matrices.

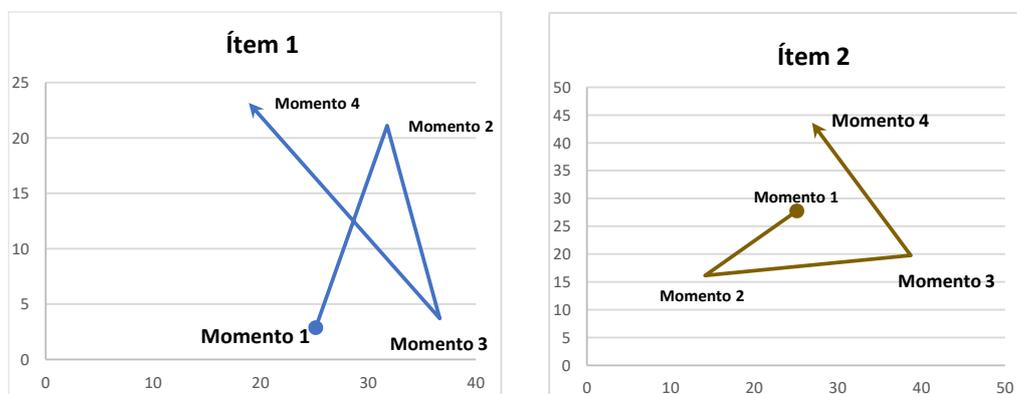
4.2.4. Evaluación de las trayectorias de los ítems

Una vez obtenido el subespacio consenso se somete a un análisis de componentes principales para reducir la dimensionalidad y poder presentar los resultados en un subespacio de baja dimensión, es decir, un plano de máxima inercia en el cual se puede proyectar la información original para analizar las trayectorias de los ítems.

Una trayectoria se define como el cambio en la posición de un ítem a lo largo del tiempo representado sobre el subespacio consenso. Una trayectoria con poca variación indica un ítem estable en el tiempo; de lo contrario, una variación destacable en el patrón de una trayectoria sugiere que el individuo ha reflejado cambios en sus resultados durante los diferentes momentos de evaluación. La variación de las trayectorias en los ítems está relacionada con la existencia de una posible recalibración o reconceptualización.

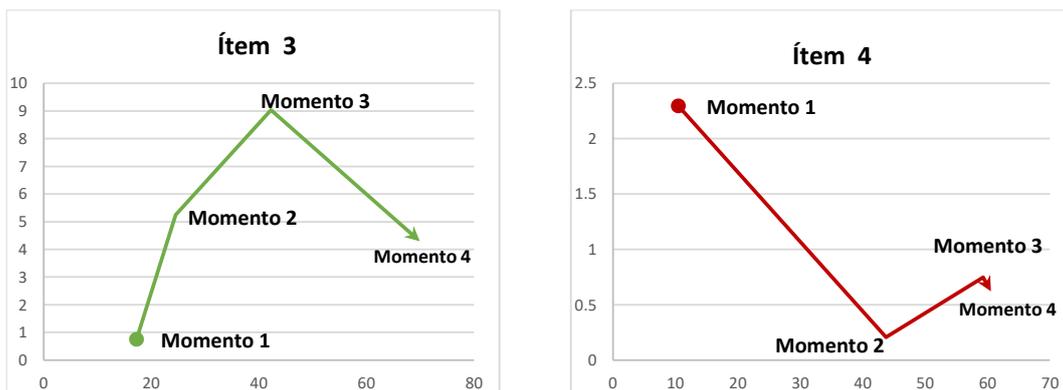
De acuerdo con el comportamiento de los ítems, las trayectorias pueden ser de dos tipos: trayectorias envolventes y trayectorias excéntricas. Las trayectorias envolventes corresponden a ítems que siguen una evolución media, es decir, la diferencia entre el valor del ítem y la media es regular en el tiempo, tal como se muestran la Figura 25.

Figura 25. *Trayectorias envolventes de un ítem*



Mientras tanto, las trayectorias excéntricas o de gran amplitud corresponden a un cambio en la estructura del ítem en el transcurso del tiempo diferente de la media. Este tipo de trayectorias se refleja en la Figura 26.

Figura 26. *Trayectorias excéntricas de un ítem*



4.2.5. MetaBiplot y el *Response Shift*: criterios para su detección

De acuerdo con nuestra revisión conceptual presentada en el Capítulo I y según los planteamientos de [Schwartz y Spranger \(1999\)](#), los estudios de CVRS y la detección *Response Shift* se presentan en tres tipos: Recalibración, Repriorización y Reconceptualización. Por otro lado, investigaciones de CVRS han estudiado los diferentes cambios en la línea de [Golembiewski et al. \(1976\)](#) donde la clasificación del cambio la hacen en términos del Cambio Gamma para referirse a la Repriorización y Reconceptualización, y con el cambio Beta para indicar Recalibración. Además, el tipo de *Response Shift* de Reconceptualización es estudiado en función de las congruencias de las estructuras factoriales.

En consecuencia, realizar un MetaBiplot y analizar la imagen euclídea generada por el espacio consenso permite la detección de *Response Shift* y sus tipos. Para ello, se establecen los siguientes criterios:

Criterio 1: Si en la representación euclídea todas las matrices tienen la misma estructura factorial, se puede calcular un subespacio consenso que permite una visión global de los ítems del instrumento de calidad de vida evaluado en todos los momentos del tratamiento.

Criterio 2: Si los ángulos que forman los vectores que representan a los ítems en los diferentes tiempos o momentos son muy agudos, es decir están próximos a cero (0) grados, eso implica que las subdimensiones del constructo latente son estables a lo largo del tiempo, y por tanto no se detecta **Reconceptualización**.

Criterio 3: Si las normas de los vectores que representan a los ítems son iguales, esto evidencia ausencia de **Recalibración**.

Criterio 4: Si al menos uno de los vectores es diferente, en términos de longitud, se establece que se ha producido una **Recalibración**.

Criterio 5: Sobre el subespacio consenso se pueden proyectar los ítems de las matrices originales para analizar las trayectorias a lo largo del tiempo. La proyección de los ítems permite detectar aquellos ítems implicados en la **Reconceptualización** y/o en la **Recalibración**.

CAPÍTULO V

EVALUACIÓN DEL *RESPONSE SHIFT* EN LA PERCEPCIÓN DE LA CALIDAD DE VIDA DE PACIENTES MIOPE SOMETIDOS A CIRUGÍA REFRACTIVA CON LÁSER EXCIMER UTILIZANDO EL METABILOT

5. EVALUACIÓN DEL *RESPONSE SHIFT* EN LA PERCEPCIÓN DE CALIDAD DE VIDA DE PACIENTES MIOPE SOMETIDOS A CIRUGÍA REFRACTIVA CON LÁSER EXCIMER UTILIZANDO EL METABILOT

Debido a la estructura de los datos de calidad de vida, los métodos de Tablas de Tres Vías representan una alternativa potencial para el análisis del *Response Shift* y sus diferentes tipos. De igual forma, corresponden a un grupo de metodologías sencillas pero robustas que permiten ser interpretables de manera gráfica y puede ser utilizada por investigadores que no son necesariamente estadísticos.

El MetaBiplot corresponde a este grupo de métodos que ha sido utilizados en otras áreas de estudios, pero no ha sido explorado en el contexto de datos de calidad de vida. Al haber establecido criterios teóricos para la detección de *Response Shift* en el Capítulo IV, en esta sección se aplicará el método MetaBiplot a un caso de estudio correspondiente a una muestra de pacientes que fueron sometidos a cirugía ocular y a su vez evaluada su autopercepción en cuanto a su calidad de vida relacionada con la visión en diferentes momentos.

5.1. Descripción del caso de estudio

Esta investigación utilizará como caso el estudio de [Perdomo-Arguello \(2015\)](#) el cual trata de un estudio descriptivo, observacional y prospectivo para una muestra de 218 pacientes miopes sometidos a cirugía refractiva con Láser Excimer del tipo LASIK de la Clínica de Cirugía Ocular y Láser en Costa Rica. Como tal, este será un estudio relacionado a la calidad de vida relacionada con la visión (CVRV).

El tamaño de la muestra se calculó para una población finita donde se entiende que la prevalencia de la miopía en Costa Rica es alrededor del 16.8% de la población total ([Tarczy-Hornoch et al., 2006](#)). Se utilizó como método de muestreo la selección de forma consecutiva de los pacientes para la consulta oftalmológica y que solicitaban valoración para la cirugía LASIK con la finalidad de corregir su condición de miopía.

Los participantes cumplieron los siguientes criterios de inclusión: edades entre 15 y 75 años, personas latinoamericanas de habla hispana, ser miopes con o sin astigmatismo y presbicia en rango entre -0.75 y hasta -25.00 dioptría, agudeza visual preoperatoria no corregida peor que 20/40, capacidad visual preoperatoria corregida mejor que 20/40.

La estimación de la muestra consideró un nivel de confianza del 95.5% y un nivel de significancia no superior al 5%, y se recolectó entre noviembre 2012 a mayo 2013. Según el estudio, todos los pacientes leyeron, entendieron y firmaron el consentimiento informado previo a la cirugía, y con ello aceptaron proporcionar sus datos primarios para la presentación y publicación de resultados.

Las características de los pacientes atendidos y sometidos a cirugía refractiva con Láser Excimer fueron heterogéneas, mayoritariamente adultos (Tabla 23). De la muestra de 218 personas por miopía (432 ojos operados), 134 (61.5%) eran mujeres. La edad promedio de la muestra fue de 31.8 ± 9.8 años, mayoritariamente costarricenses (97.3%). Según otras características, 131 (60.1%) son solteros, 114 (52.3%) tenían estudios universitarios y 137 (62.8%) tenían alguna profesión.

El perfil de los pacientes que acceden a este tipo de cirugías muestra ser personas con un alto nivel sociocultural y económico, mayoritariamente adultos menores de 60 años, económicamente activos con un nivel educativo superior y por tanto con mayor formación y calificación, trabajos mejor remunerados que generan mayores ingresos.

Para la evaluación de la CVRV, se utilizó el NEI-VFQ-25 (National Eye Institute Visual Function Questionnaire) en su versión reducida en español, que es un instrumento que mide el impacto de la patología ocular en la CVRV (Grubbs et al., 2014; Kay & Ferreira, 2014) a través de 23 ítems dentro de 12 subescalas: salud general, visión general, dolor ocular, actividades de visión cercana, actividades de visión distante, funcionamiento social, salud mental, dificultades de rol específico, dependencia visual específica, dificultades de conducción, visión de colores y visión periférica. Las ponderaciones van de 0 a 100, donde 100 representa la no existencia de discapacidad relacionada con la visión.

El cuestionario NEI-VFQ-25 tiene una validez adecuada y altamente confiable, y proporciona resultados muy similares a los obtenidos con la versión larga. Es ampliamente utilizado e internamente consistente (Quaranta et al., 2016). Se utilizó el NEI-VFQ-25, pero excluyeron las dimensiones salud general y dificultades de conducción, ya que el autor señaló que no aportaban significancia a su investigación, tal como se muestra en la Tabla 24.

Tabla 23. *Características Demográficas de los pacientes sometidos a la Cirugía Excimer Laser Corneal. Agosto 2012 a septiembre 2013*

Variables	Categories	Casos	
		n	%
Sexo	Femenino	134	61.5
	Masculino	84	38.5
Grupo de Edad	Menores (≤ 19)	8	3.7
	Adultos (20-59)	208	95.4
	Mayores (≥ 60)	2	0.9
Educación	Primaria	4	1.8
	Pre-Media	21	9.6
	Media	79	36.3
	Licenciatura	80	36.7
	Doctorado	21	9.6
Residencia	Técnico	13	6.0
	San José	111	50.9
	Alajuela	43	19.7
	Cartago	17	7.8
	Heredia	35	16.1
	Puntarenas	5	2.3
Estado civil	Limón	6	2.7
	Managua, Nicaragua	1	0.5
	Soltero	131	60.1
	Casado	74	33.9
	Separado/Divorciado	11	5.1
Profesión	Viudo(a)	2	0.9
	Profesional	137	62.8
	Técnicos	45	20.6
	Estudiante	31	14.2
Nacionalidad	Ama de casa	5	2.4
	Costa Rica	212	97.2
	Otros	6	2.8

Tabla 24. Descripción de las Variables del NEI-VFQ-25

Dimensiones	Ítems por dimensión	Descripción
Visión General	A1	Se refiere a la valoración por la persona del tipo de vista que tiene
Actividad de visión cercana	A4	Se refiere a la dificultad que tienen para leer la letra de los periódicos
	A5	Valora la dificultad para hacer trabajo o pasatiempos que requieran ver bien de cerca
	A6	Evalúa la dificultad para encontrar algo en una repisa
Actividad de visión distante	A7	Dificultad en leer carteles
	A8	Dificultad en bajar escalones, escaleras o el borde de la acera cuando hay poca luz o es de noche
	A13	Dificultad para salir al cine, teatro o ver eventos deportivos
Dolor ocular	A3	Dolor o malestar que ha sentido en los ojos o alrededor por ardor, picazón o dolor
	A17	Se refiere a cada que tiempo no puede hacer lo que quisiera a causa del dolor o malestar en los ojos o alrededor de los ojos como picazón, ardor o dolor
Función específica social	A10	Dificultad para ver cómo reacciona la persona cuando le dicen algo
	A12	Dificulta para ir de visita
Visión – Rol específico de dificultades	A15	Se refiere a cada cuanto tiempo realiza menos trabajo a causa de la vista
	A16	Valora cada cuanto tiempo está limitado para trabajar o hacer otras cosas por la vista
Visión – Salud mental específica	A2	Hace referencia a los seguido que se preocupa por la vista
	A19	Valora la frustración que siente a causa de la vista
	A20	Se refiere al menor control que podría o no producirle a la persona la vista
	A23	Preocupación de hacer cosas que le causan vergüenza por la vista
Visión – Dependencia visual específica	A18	Se queda en casa por la vista
	A21	Dependencia que tiene por la vista
	A22	Necesidad de ayuda de otras personas a causa de la vista
Visión de color	A11	Dificultad para escoger y coordinar su propia ropa
Visión periférica	A9	Dificultad para notar objetos a los lados cuando va caminando

Para validar el contenido del tamaño de la muestra, [Kuhn et al. \(1996\)](#) sugieren por lo menos de 5 a 10 sujetos por ítem o unidad. Un rango entre 125 y 250 sujetos es aceptable para aplicar el NEI-VFQ-25, lo que indica que una muestra de 218 personas (432 ojos operados) utilizados en ese estudio fue representativa. Todos los procedimientos cumplieron con los estándares éticos del país de origen en el momento de estudio.

El instrumento se aplicó en cinco etapas diferentes: el *Pre-Test*, que se aplica durante la evaluación preoperatoria y establece la condición inicial de CV, nivel de miopía y capacidad visual; el *Post-Test # 1*, aplicado al día siguiente después de la cirugía LASIK; el *Post-Test # 2*, al mes siguiente de la cirugía; y el *Post-Test # 3*, tres meses después. En el mismo momento se aplica un *Then-Test* de forma retrospectiva con respecto a su estado antes de la cirugía LASIK. Para cada paciente, el cuestionario tuvo una duración de 5 a 10 minutos y se aplicó entre noviembre de 2012 a septiembre de 2013, comparando los valores de agudeza visual y equivalente esférico en el pre y postoperatorio.

Para detectar el cambio, se compararon los resultados de *Pre-Test* y *Post-Test* apoyado por el análisis de diferencias estadísticamente significativas ($\alpha: 0.05$). El método *Then-Test* se utilizó para identificar un cambio de tipo recalibración. Para valorar si existe o no reconceptualización, se utilizó el esquema de [Vicente-Galindo \(2003\)](#), y [Vicente-Galindo y Galindo-Villardón \(2009\)](#) por medio del Método STATIS Dual ([H. L'Hermier Des Plantes, 1976](#)).

Los resultados obtenidos en las puntuaciones de CV antes y después de la cirugía muestran un cambio real y un cambio del tipo reconceptualización en CVRV. La existencia de recalibración en la percepción del paciente se demuestra con una diferencia entre *Then-Test* y *Pre-Test* de $p = 0.018$, que es estadísticamente significativa. Además, el cambio real que comparó el *Post-Test* y el *Then-Test*, registró puntuaciones por encima del 83.1% al tercer mes de operado con $p \leq 0.001$.

Para la identificación de reconceptualización, se analizaron las estructuras de las matrices en los cuatro momentos utilizando el coeficiente de correlación de Hilbert-Schmidt. Al ser la correlación vectorial entre las matrices tendiente a la unidad, y

mostrando una fuerte correlación entre 0.940 (max) y 0.857 (min), no se identificó la presencia de este tipo de cambio (reconceptualización).

5.2. Análisis sobre la calidad de las escalas de medida

Para evaluar la calidad de las escalas de medida, se realiza un análisis de tipo univariante de los resultados de cada ítem del instrumento NEI-VFQ-25 evaluados en cada momento. Este análisis tiene como propósito examinar si los datos de cada ítem poseen escalas de medidas diferentes que requieran ser sometidos a algún tipo de transformación para su estandarización general, o no.

Para ello, se calcularon las medias y las desviaciones típicas para cada uno de los momentos evaluados y para cada uno de los ítems del instrumento. La Tabla 25 muestra los estadísticos de todos ítem estudiados en los cuatro momentos (preoperatorio, al día siguiente, al mes siguiente, tres meses después).

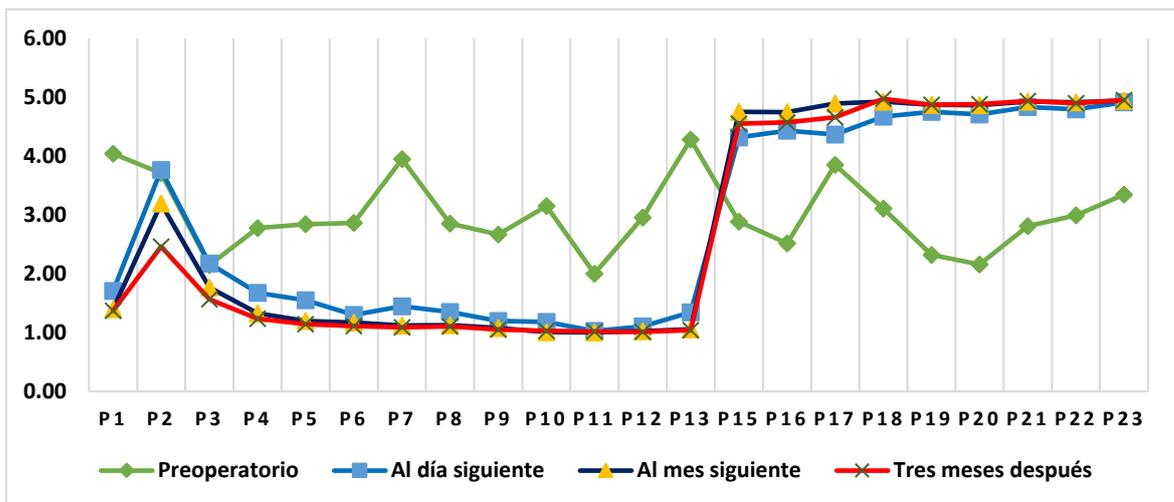
Tabla 25. Estadísticos de cada ítem del cuestionario NEI-VFQ-25 en cada uno de los momentos evaluados

Ítems	Preoperatorio		Al día siguiente		Al mes siguiente		Tres meses después	
	Media	Desv. Tip.	Media	Desv. Tip.	Media	Desv. Tip.	Media	Desv. Tip.
Ítem 1	4.05	0.88	1.71	0.70	1.40	0.62	1.37	0.59
Ítem 2	3.71	1.12	3.77	1.24	3.20	1.45	2.46	1.37
Ítem 3	2.15	1.04	2.17	0.84	1.77	0.76	1.57	0.67
Ítem 4	2.78	1.37	1.67	0.99	1.33	0.77	1.24	0.75
Ítem 5	2.84	1.35	1.55	0.92	1.20	0.64	1.14	0.51
Ítem 6	2.87	1.25	1.30	0.67	1.17	0.56	1.11	0.41
Ítem 7	3.95	1.04	1.45	0.66	1.12	0.34	1.09	0.39
Ítem 8	2.85	1.24	1.35	0.67	1.13	0.43	1.11	0.46
Ítem 9	2.67	1.16	1.20	0.49	1.08	0.36	1.05	0.26
Ítem 10	3.15	1.31	1.18	0.49	1.01	0.10	1.03	0.22
Ítem 11	2.00	1.19	1.03	0.19	1.00	0.07	1.02	0.13
Ítem 12	2.96	1.44	1.11	0.44	1.02	0.22	1.01	0.12
Ítem 13	4.28	0.93	1.34	0.66	1.06	0.31	1.04	0.32
Ítem 15	2.89	1.39	4.32	1.10	4.76	0.67	4.56	1.19
Ítem 16	2.52	1.30	4.44	1.04	4.75	0.66	4.58	1.14
Ítem 17	3.85	1.28	4.37	1.00	4.89	0.34	4.66	0.97

Ítem 18	3.11	1.64	4.67	0.90	4.93	0.35	4.98	0.15
Ítem 19	2.32	1.54	4.76	0.77	4.88	0.55	4.88	0.51
Ítem 20	2.16	1.43	4.71	0.80	4.87	0.59	4.89	0.51
Ítem 21	2.81	1.59	4.83	0.61	4.93	0.45	4.94	0.41
Ítem 22	3.00	1.54	4.80	0.66	4.92	0.46	4.90	0.54
Ítem 23	3.35	1.66	4.92	0.39	4.94	0.34	4.95	0.33

El análisis anterior sobre la calidad de las escalas de medida muestra que los valores medios de los ítems no son todos homogéneos en los cuatro momentos (Figura 27). Por eso se sugiere que los valores de los ítems sean sometidos a un proceso de transformación. Dicho proceso será considerado durante la aplicación del método propuesto.

Figura 27. Valores de las medias de los ítems en los diferentes momentos (preoperatorio, al día siguiente, al mes siguiente, tres meses después)



El análisis univariante requiere también examinar las diferentes dimensiones del instrumento, por ello se analiza los valores de las medias y errores estándar de cada dimensión para los distintos momentos. La Tabla 26 muestra los estadísticos por cada dimensión del instrumento y la significancia del contraste de t de Student en los diferentes momentos de la evaluación. Se debe entender que mientras mayor sea la puntuación, mejor será la condición de calidad de vida de los pacientes según la dimensión.

Tabla 26. Análisis de las medias de las dimensiones en los diferentes momentos (preoperatorio, al día siguiente, al mes siguiente, tres meses después)

Dimensiones	n	Preoperatorio		Al día siguiente		Al mes siguiente		Tres meses después		p-valor
		Media	S.E.	Media	S.E.	Media	S.E.	Media	S.E.	
Actividad de visión cercana	218	3.22	0.090	1.51	0.059	1.16	0.036	1.11	0.029	<0.0001
Actividad de visión distante	218	3.57	0.089	1.28	0.040	1.09	0.026	1.07	0.027	<0.0001
Dolor Ocular	218	3.00	0.088	4.10	0.105	3.33	0.052	3.11	0.076	<0.0001
Función específica social	218	3.06	0.093	1.11	0.026	1.02	0.012	1.02	0.012	<0.0001
Rol específico de dificultad	218	2.70	0.092	2.83	0.118	4.75	0.045	4.57	0.079	<0.0001
Salud mental específica	218	2.89	0.108	3.99	0.103	4.47	0.076	4.29	0.089	<0.0001
Dependencia visual específica	218	2.73	0.106	3.78	0.114	3.95	0.118	3.96	0.117	<0.0001

El análisis muestra que las dimensiones Actividad de visión cercana, Actividad de visión distante y Rol específico de dificultad son similares en comportamiento ya que sus valores medios son altos en el preoperatorio en comparación con los demás momentos. Se destaca que las diferencias encontradas en las puntuaciones de las muestras son estadísticamente significativas.

El valor medio más elevado de la dimensión Dolor ocular se presenta al día siguiente de haberse realizado la cirugía, por lo que muestra valores y las diferencias son estadísticamente significativas.

Las dimensiones Función específica social, Salud mental y Dependencia visual específica presentaron puntuaciones más altas a los tres meses después de haberse realizado la cirugía, y evidencian estadísticamente diferencias significativas en los cuatros momentos con un p-valor <0.0001.

De forma general, se evidencia una mejora en la calidad de vida de los pacientes en las dimensiones Dolor ocular, Rol específico de dificultad, Salud mental, Dependencia Visual, y esta mejoría se manifiesta en todos los momentos evaluados después de la cirugía.

5.3. El MetaBiplot como herramienta de inspección de *Response Shift* en pacientes operados de miopía

Como el objetivo de nuestra investigación es identificar si se ha producido cambio en los pacientes miopes que fueron sometidos a una cirugía correctiva y aplicado un instrumento de autopercepción para medir su calidad de vida relacionada con la visión (CVRV), para configurar el análisis de tipo MetaBiplot, primero se dispone de una matriz de datos donde se cuenta con $n = 218$ pacientes y $p = 23$ ítems evaluados en 4 momentos. Luego, se toma en consideración la unidad de interés que será clasificada, para este nuestro estudio serán los ítems del instrumento NEI-VFQ 25.

Seguido, se realizará un análisis MetaBiplot utilizando el GH-Biplot para la comparación e integración global de los resultados de todos los momentos estudiados permitiendo obtener un subespacio consenso que integre los elementos y se pueda identificar las similitudes o diferencias entre ellos. Finalmente, se realiza el análisis de las diferentes estructuras factoriales permitiendo evaluar los criterios establecidos para la detección del cambio.

Tomando en cuenta que se cuenta con cuatro matrices, una correspondiente a cada momento, se le aplicará a cada matriz la transformación de estandarización por columna, ya que los elementos que se utilizarán serán las puntuaciones de cada ítem, siendo ésta la transformación que mejor resultados produce. La aplicación del método MetaBiplot se ha trabajado el programa MultiBiplot desarrollado por [Vicente-Villardón \(2010\)](#).

El análisis inicia identificando el conjunto de ejes ortogonales que mejor permita la congruencia de las principales direcciones de la inercia. La estructura de los componentes comunes (ejes) en el subespacio consenso se muestra en la Tabla 27 de los valores propios en el plano 1-2, donde la absorción de la inercia obtiene un 77.4% de la información total al considerar los dos primeros ejes; es decir, recoge un alto porcentaje de información.

Con respecto a la similaridad global, esta nos indica que al ser el coeficiente 0.9725 bastante próximo a uno (1), los diferentes momentos son similares a la estructura del espacio consenso en la primera componente y menos similar en la segunda, con un

valor de 0.5755. Esto nos indica las estructuras de los cuatro momentos evaluados son bastante congruente en el subespacio de la Componente 1.

Tabla 27. *Valores propios del MetaBiplot*

Eje	Valor Propio	% Inercia	% Acumulado	Similaridad Global
Eje 1	3.890	48.627	48.627	0.9725
Eje 2	2.302	28.770	77.397	0.5755

La Tabla 28 presenta los cosenos cuadrados de cada momento con relación a las componentes del subespacio consenso. En términos general, los cuatro momentos evaluados muestran una alta relación con la primera componente, ya que todos sus valores son próximos a uno (1).

Tabla 28. *Cosenos cuadrados del subespacio consenso según los momentos*

Momento	Componente 1	Componente 2
Preoperatorio	0.975	0.394
Al día siguiente	0.981	0.535
Al mes siguiente	0.951	0.701
Tres meses después	0.983	0.671

En la tabla anterior se puede observar que, para la primera componente, todos los momentos son muy similares a la estructura del subespacio consenso y que presentan valores próximos a la unidad en todos los casos. Por otro lado, la segunda componente muestra en el momento Preoperatorio un coeficiente algo distante con respecto a los presentados en los otros tres momentos, sugiriendo que los vectores asociados al momento Preoperatorio estarían más relacionados con la primera componente.

La proporción de variabilidad acumulada de cada momento absorbida por cada uno de los componentes de consenso se muestra en la Tabla 29. En la misma se puede observar que las dos componentes son muy similares a los obtenidos en los análisis individuales, indicando que la pérdida de información es casi mínima.

Tabla 29. *Porcentaje de inercia absorbida por el consenso según momento*

Momentos	Componente 1	Componente 2
Preoperatorio	39.531	45.223
Al día siguiente	33.437	40.292
Al mes siguiente	30.409	41.048
Tres meses después	30.861	41.264

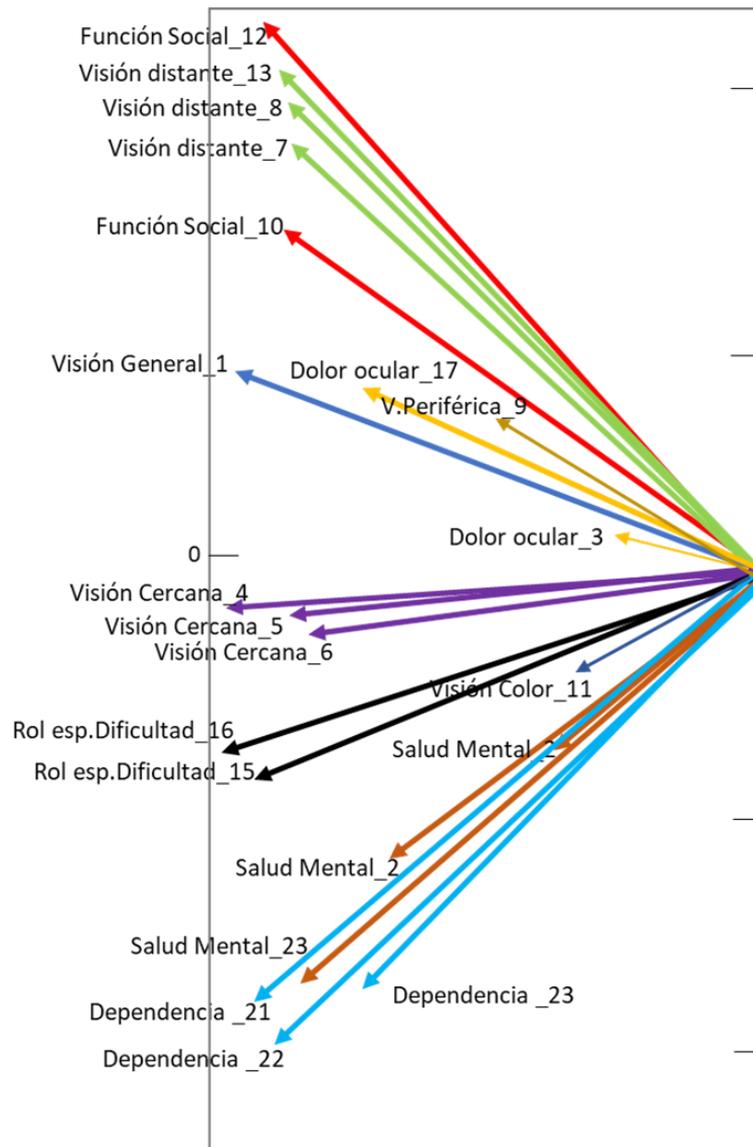
Para evaluar la similitud entre los componentes originales de cada momento evaluado y las nuevas componentes del subespacio consenso, se calculan los cosenos de los ángulos cuyos resultados se presentan en la Tabla 30. En la misma se muestran altas relaciones entre todos los ejes 1 de todos los momentos con respecto al componente 1; mientras que similar comportamiento se manifiesta entre todos los ejes 2 de todos los momentos con respecto a la segunda componente.

Tabla 30. *Cosenos de los ángulos entre componentes comunes según momento*

Momentos	Ejes	Componente 1	Componente 2
Preoperatorio	Eje 1	0.987	-0.018
	Eje 2	0.003	0.627
Al día siguiente	Eje 1	-0.990	0.008
	Eje 2	-0.019	0.732
Al mes siguiente	Eje 1	0.953	0.226
	Eje 2	-0.206	0.806
Tres meses después	Eje 1	0.987	0.067
	Eje 2	0.095	-0.817

De acuerdo con los resultados obtenidos, se puede señalar que todas las matrices recogen un alto porcentaje de la información, y podrán ser visualizadas en un solo subespacio consenso, conformando así un MetaBiplot, permitiendo representar en un espacio de baja dimensión la mayor información que explique el comportamiento de las variables, tal como se muestra en la Figura 28.

Figura 28. Representación de las variables (ítems) en el subespacio consenso (MetaBiplot)



La representación del subespacio consenso, o sea el MetaBiplot, de la figura anterior muestra un conjunto de vectores que representan los diferentes ítems del instrumento utilizado en el caso de estudio. Según el mismo, corresponde a un total de 22 ítems y 10 dimensiones que capturan en los diferentes momentos las percepciones de los pacientes a lo largo del tratamiento médico.

En esta representación se pueden realizar varias observaciones generales. Primero, se distingue que los vectores de una misma dimensión muestran una fuerte agrupación

entre ellos, sugiriendo que las opiniones de los pacientes con respecto a los ítems son consistentes en todos los momentos. Estos vectores se han identificado con el mismo color según dimensión para lograr una mejor visualización de cada conjunto.

En adición, los ítems por dimensión muestran una fuerte correlación con ítems de otras dimensiones. Los vectores de las dimensiones Función Social, Visión Distante, Visión General y Dolor Ocular y Visión Periférica se encuentran más relacionados entre sí como un subgrupo; mientras que las dimensiones Visión Cercana, Rol Específico de la Dificultad, Salud Mental, Visión Color y Dependencia Específica, se encuentran más relacionados dentro de un segundo subgrupo.

Con respecto a los criterios establecidos para el análisis del MetaBiplot conducentes a la detección de *Response Shift*, y basado en los resultados del MetaBiplot, se señalan las siguientes consideraciones:

- **Criterio 1:** Si en la representación euclídea todas las matrices tienen la misma estructura factorial, se puede calcular un subespacio consenso que permite una visión global de los ítems del instrumento de calidad de vida evaluado en todos los momentos del tratamiento.

La Figura 28 muestra una clara representación de todas las variables, dado por los 22 ítems del instrumento NEI-VFQ 25, en el subespacio. Por tanto, al representarse todos los ítems de manera clara y consistente, se logra una configuración consenso indicando que las estructuras en los cuatro momentos resultaron ser congruentes, sugiriendo que existe una fuerte representatividad de todas las variables

- **Criterio 2:** Si los ángulos que forman los vectores que representan a los ítems en los diferentes tiempos o momentos son muy agudos, es decir están próximos a cero (0) grados, eso implica que las subdimensiones del constructo latente son estables a lo largo del tiempo, y por tanto no se detecta **Reconceptualización.**

Bajo el criterio de que los ángulos de los ítems dentro de una misma dimensión sean próximos a cero grados, la Figura 28 muestra que, para todas las 10 dimensiones, las distancias entre los vectores son mínimas y próximas a cero, dada las altas correlaciones entre las variables. Por tanto, el cumplimiento de este criterio sugiere que no se ha producido Reconceptualización.

- **Criterio 3:** Si las normas de los vectores que representan a los ítems son iguales, esto evidencia ausencia de **Recalibración**.

Mediante la representación de las variables en el subespacio consenso se puede identificar normas de vectores que presentan diferentes magnitudes. A pesar de que se muestran normas de los vectores con magnitudes similares, como por ejemplo Función Social (12), Visión Distante (13), Visión Distante (8), Visión Distante (7), Dependencia (21), Dependencia (22), Rol de Dificultad (15) y Rol de Dificultad (16); también se muestran normas de los vectores con magnitudes diferentes, como son el resto de los ítems.

Por tanto, dado que todas las normas de los vectores dentro de este subespacio consenso no son iguales, se evidencia la presencia de *Response Shift* del tipo Recalibración.

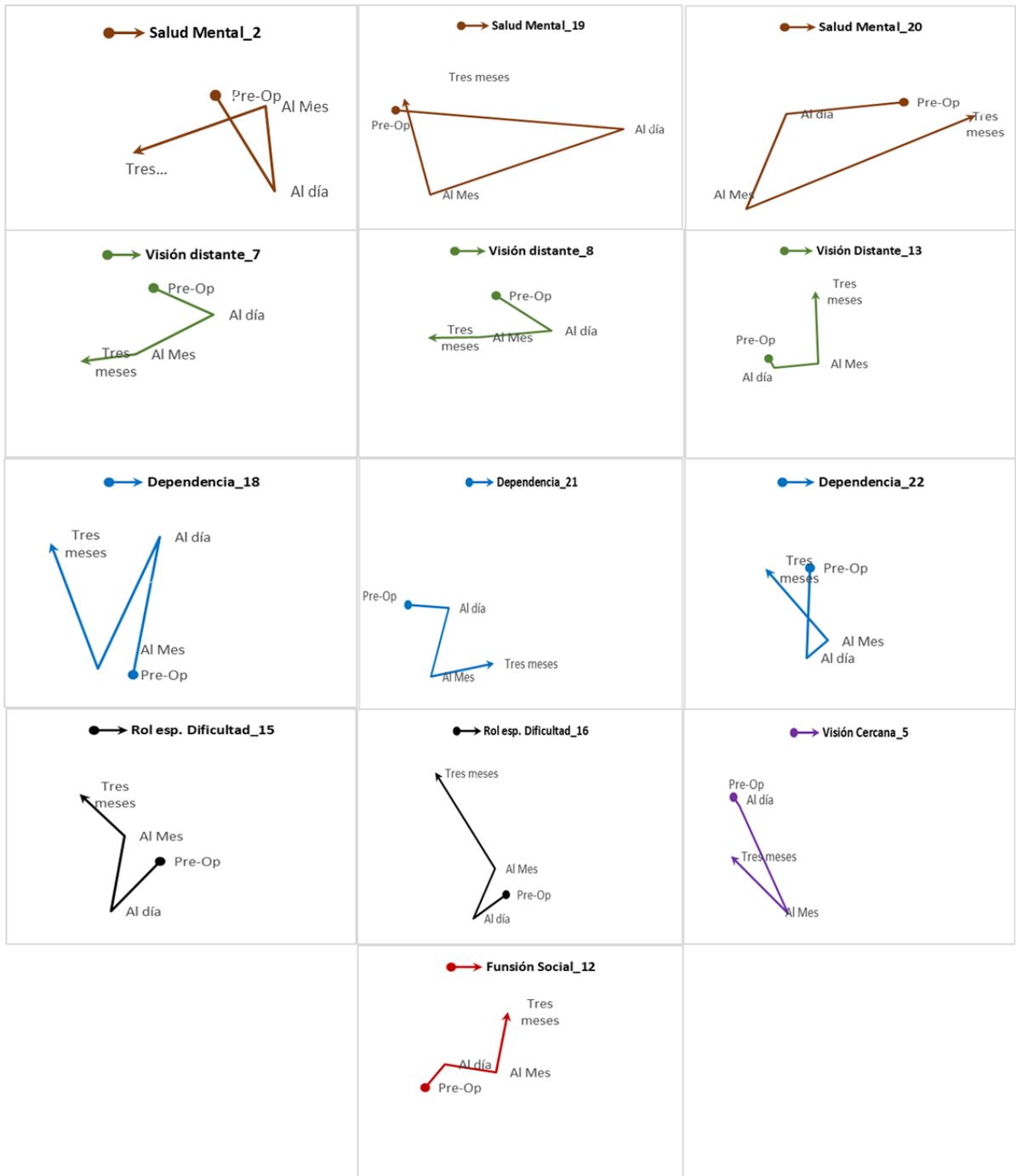
- **Criterio 4:** Si al menos uno de los vectores es diferente, en términos de longitud, se establece que se ha producido una **Recalibración**.

Al verificar la longitud de los vectores, se identifica que las variables Visión Periférica (9), Dolor Ocular (3), Visión Color (11) y Salud Mental (19) poseen longitudes menores y muy marcadas con respecto al resto de las variables. Esto es evidencia que, en término de longitud, se ha producido un *Response Shift* del tipo Recalibración.

- **Criterio 5:** Sobre el subespacio consenso se pueden proyectar los ítems de las matrices originales para analizar las trayectorias a lo largo del tiempo. La proyección de los ítems permite detectar aquellos ítems implicados en la **Reconceptualización** y/o en la **Recalibración**.

Al proyectar la trayectoria de cada uno de los 22 ítems del caso de estudio, se pudo identificar la existencia de 13 ítems con trayectorias envolventes y 9 ítems con trayectorias excéntricas. Las primeras proyecciones sugieren que, al no mostrar mayor variabilidad en sus trayectorias, estos ítems son relativamente estables en el tiempo y no han influenciado en la existencia de Recalibración, tal como muestra la Figura 29.

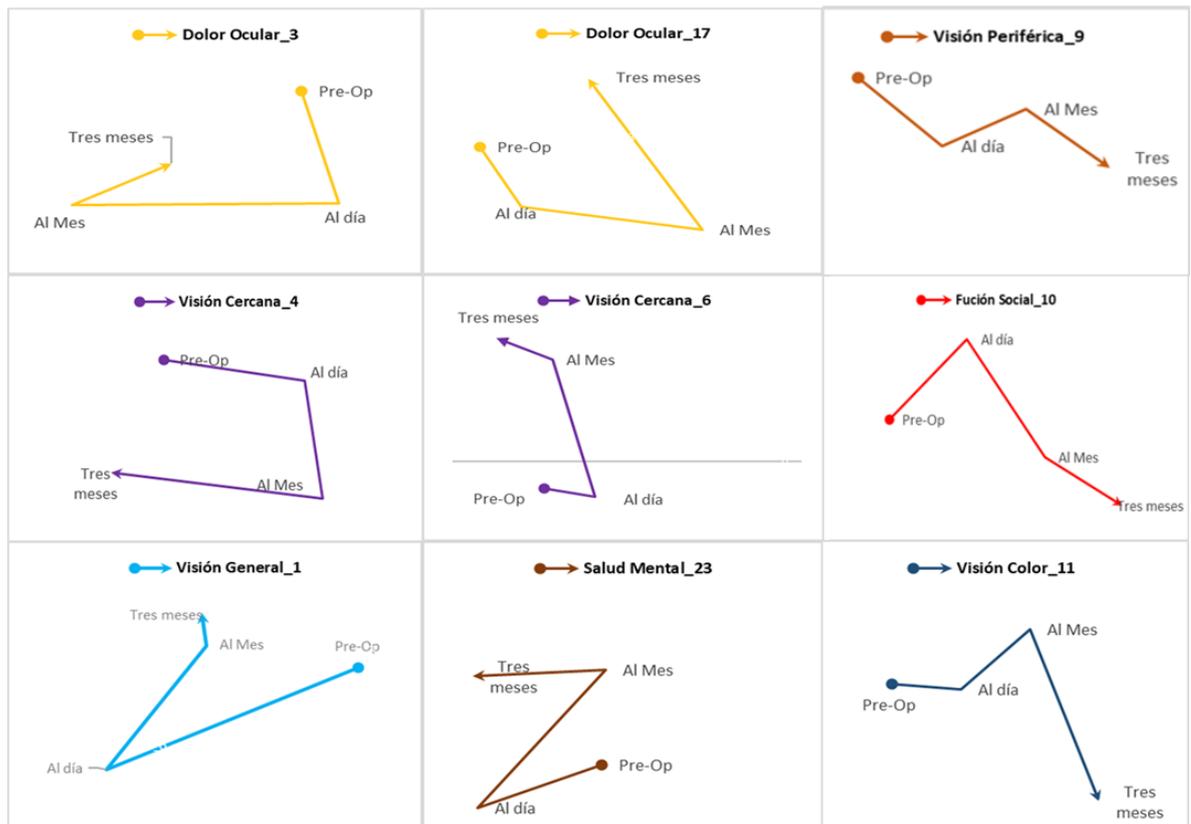
Figura 29. Ítems con trayectorias envolventes



Por el otro lado, las siguientes proyecciones con características de trayectorias excéntricas, y que se presentan en los ítems Dolor Ocular (3), Dolor Ocular (17), Visión Periférica (9), Visión Cercana (4), Visión Cercana (6), Función Social (10), Visión General (1), Salud Mental (23) y Visión Color (11), siguen trayectorias con alta

variación y reflejan un cambio en la estructura del ítem en el transcurso del tiempo. Por tanto, es evidencia suficiente para señalar que estos ítems están implicados en la existencia de *Response Shift* del tipo Recalibración, tal como muestra la Figura 30.

Figura 30. Ítems con trayectorias excéntricas



Con base a la identificación de las trayectorias excéntricas, se puede señalar que los 9 vectores presentados en la figura anterior representan a aquellos ítems que mayor influyen en la existencia de Recalibración en los pacientes.

5.4. Consideraciones de los resultados

El MetaBiplot es un método estadístico multivariante riguroso que permite una fácil interpretación gráfica que, junto a criterios teóricos establecidos en esta investigación, facilita la evaluación del comportamiento de las variables.

La aplicación del método MetaBiplot a los datos de CVRV permitió identificar un subespacio consenso que representa las variables en un espacio de baja dimensión y una alta calidad de representación. En este espacio euclídeo, se logró evaluar las relaciones entre los vectores, representados por los ítems, y con ello se examinaron sus ángulos, normas, longitudes y trayectorias individuales para cada ítem.

Mediante el método MetaBiplot y la aplicación de los criterios al caso del estudio se obtuvo evidencia suficiente para detectar la existencia de un cambio, o sea *Response Shift*, en la autopercepción de los pacientes sometidos al procedimiento médico, y a su vez permitió identificar la existencia de un *Response Shift* de tipo Recalibración. Sin embargo, los resultados no mostraron evidencia de una Reconceptualización.

El método MetaBiplot es la primera vez que se utiliza dentro del contexto de calidad de vida relacionado con la salud, y la revisión de la literatura así lo demuestra. Como método estadístico multivariante, y mediante el análisis de los datos de tres vías, el MetaBiplot ofrece la oportunidad de examinar individuos y variables a lo largo de varios momentos, permitiendo capturar en una sola configuración consenso el comportamiento de los elementos estudiados.

Con la ayuda de este método, la contribución de nuestra investigación demostró que se puede evaluar la existencia de estabilidad o de cambio (*Response Shift*) a nivel de variables y dimensiones en donde se manifiesten dichas variaciones ante un evento positivo.

A su vez, el MetaBiplot demostró su capacidad para la detección del cambio de tipo Recalibración y Reconceptualización, aun cuando en nuestros datos se señaló la no existencia de este último.

En cuanto a la detección del *Response Shift*, se puede señalar que mediante el método MetaBiplot se evidenció un cambio ante un evento positivo en los pacientes después

de haberse sometido al procedimiento quirúrgico de corrección visual. Este resultado se demostró al evaluar la trayectoria de los ítems de las dimensiones Visión General, Dolor Ocular, Visión Cercana, Visión Periférica, Visión Color, Función Social y Salud Mental, las cuales demostraron variaciones con altas amplitudes.

Lo anterior demostrado en nuestra investigación mediante la aplicación del método MetaBiplot es consistente con los resultados presentados por [Perdomo-Arguello \(2015\)](#) y que es la fuente de datos originales para este estudio, al señalar que los datos tienen capacidad de ser interpretados y comparables en el tiempo, logró la detección del cambio de tipo Recalibración y evidenció la no existencia de Reconceptualización.

Para los efectos de los objetivos de esta investigación, solamente se abordó la detección de *Response Shift* y en análisis para la detección del tipo Recalibración y Reconceptualización. En este punto, se abre el espacio para realizar posteriores investigaciones que permitan evaluar la capacidad del MetaBiplot para identificar la Repriorización en la CVRS.

CONCLUSIONES

CONCLUSIONES

Luego de realizar los análisis de resultados, se presentan las siguientes conclusiones:

1. La calidad de vida relacionada con la salud se entiende como el nivel de bienestar derivado de la evaluación que la persona realiza en diversos dominios de su vida considerando el impacto en su estado de salud. Mientras tanto, *Response Shift* es el cambio en el significado de una autoevaluación aplicada a un sujeto con respecto a su calidad de vida. Cuando el cambio se manifiesta en los estándares internos de la escala de medición del individuo, se establece una Recalibración; si se manifiesta en la escala de valores, existe una Repriorización; y si se manifiesta una redefinición del constructo, se denomina Reconceptualización.
2. El estudio de la calidad de vida relacionada con la salud ha sido sujeto a la aplicación de un número diverso de metodologías para su evaluación. La literatura señala que entre las técnicas de mayor difusión se encuentran el *Then-Test* y los métodos multivariantes. De este último grupo, los Modelos de Ecuaciones Estructurales han sido los más empleados por su fortaleza metodológica. Sin embargo, otros modelos se han explorado para encontrar nuevas metodologías para examinar el fenómeno: Teoría de Respuesta al Ítem, Efectos Mixtos, Medidas de Importancia Relativa y el Método de Clasificación y Regresión de Árbol.
3. La revisión sistemática desarrollada en esta investigación para estudiar los diferentes métodos estadísticos utilizados para la detección de *Response Shift* durante los años 2010 a 2020 identificó un total 107 artículos, de los cuales el 31.8% utilizaron el método *Then-Test* y 27.1% adoptaron los Modelos de Ecuaciones Estructurales de Oort. Otros modelos utilizados fueron la Regresión Lineal Múltiple (7), Regresión de Efectos Mixtos (6), Análisis de Trayectoria Latente (6), Teoría de Respuesta al Ítem (6), Regresión Logística (5), Regresión y Clasificación de Árbol (4) y los Métodos de Importancia

Relativa (4). Se utilizaron otros métodos, pero en un número menor de estudios.

4. El método tradicional para la detección de *Response Shift* es el *Then-Test* y registró el mayor número de *Response Shift* del tipo recalibración (27), mientras que los Modelos de Ecuaciones Estructurales de Oort registraron similar número de estudios detectando recalibración (24), pero el mayor registro de repriorización (13) y reconceptualización (7) de toda la revisión sistemática. Los Modelos de Ecuaciones Estructurales de Schmidt detectaron cambios solamente en los estándares internos en dos artículos. El resto de los métodos también demostraron ser utilizados para la detección de los tres tipos de *Response Shift*, con excepción de los Métodos de Importancia Relativa, Regresión de Efectos Mixtos, Regresión de Bosques Aleatorios y el Análisis de Trayectoria Latente que no identificaron la presencia de reconceptualización.
5. Los resultados muestran que no todos los métodos logran detectar *Response Shift* en similares proporciones, la mayoría logró identificar recalibración, y en menor proporción repriorización y en pocos casos reconceptualización. Sin embargo, nuestra revisión sistemática pone de manifiesto la necesidad de explorar otros métodos que permitan igual capacidad de detección, con metodologías robustas y alta precisión.
6. Debido a la estructura de los datos de calidad de vida relacionada con la salud, los métodos de Tablas de Tres Vías representan una alternativa potencial para el análisis del *Response Shift* y sus diferentes tipos. De igual forma, corresponden a un grupo de metodologías con solidez suficiente y que permiten ser interpretables de manera gráfica y utilizada por investigadores que no son necesariamente estadísticos. Dentro de estos métodos se encuentran el MetaBiplot. La revisión de la literatura destacó el MetaBiplot ha sido utilizado en otras áreas de estudio, pero que por su naturaleza puede ser explorado dentro del contexto de datos de calidad de vida.

7. En nuestra investigación se aplicó el método MetaBiplot y con el mismo se detectó la existencia del cambio y la identificación de cambios en los estándares internos. Tomando en consideración el fundamento conceptual de la teoría de *Response Shift* y las bases de los criterios aportados por Vicente-Galindo (2003), el método MetaBiplot permitió demostrar la detección del *Response Shift*, la identificación de *Response Shift* del tipo Recalibración, y a su vez, demostró la no existencia de *Response Shift* del tipo Reconceptualización.
8. El análisis individual de las trayectorias de los ítems permitió identificar cuáles de ellos están más implicados en la existencia de cambios en la escala de medidas de la autopercepción de los pacientes, o sea Recalibración. Con ello, se identificó que los ítems Visión General (1), Dolor Ocular (3), Visión Cercana (4), Visión Cercana (6), Visión Periférica (9), Función Social (10), Visión Color (11), Dolor Ocular (17) y Salud Mental (23) fueron aquellos que presentaron mayor variabilidad en sus proyecciones, cuyas trayectorias fueron excéntricas.
9. En cuanto a la detección del *Response Shift*, se puede señalar que mediante el método MetaBiplot se evidenció un cambio ante un evento positivo en los pacientes después de haberse sometido al procedimiento quirúrgico de corrección visual. Este resultado se demostró al evaluar la trayectoria de los ítems de las dimensiones Visión General, Dolor Ocular, Visión Cercana, Visión Periférica, Visión Color, Función Social y Salud Mental, las cuales demostraron variaciones con altas amplitudes.
10. La contribución de nuestra investigación concluye que el método MetaBiplot es una técnica estadística multivariante de visualización fundamentada en una metodología robusta y con capacidad suficiente para ser considerada como un método alternativo la detección de *Response Shift* y la identificación del tipo de cambio Recalibración y Reconceptualización.

REFERENCIAS BIBLIOGRÁFICAS

REFERENCIAS BIBLIOGRÁFICAS

- Aburub, A. S., Gagnon, B., Ahmed, S., Rodríguez, A. M., & Mayo, N. E. (2018). Impact of reconceptualization response shift on rating of quality of life over time among people with advanced cancer. *Supportive Care in Cancer*, 26(9), 3063–3071. <https://doi.org/10.1007/s00520-018-4156-7>
- Afifi, A., & Clark, V. (1990). *Computer-aided multivariate analysis* (Chapman &).
- Ahmavaara, Y. (1954). Transformation analysis of factorial data. *Annals of the Academy Science Fennicae*, 881(2), 54–59.
- Ahmed, S., Mayo, N., Scott, S., Kuspinar, A., & Schwartz, C. (2011). Using latent trajectory analysis of residuals to detect response shift in general health among patients with multiple sclerosis article. *Quality of Life Research*, 20(10), 1555–1560. <https://doi.org/10.1007/s11136-011-0005-6>
- Ahmed, S., Sawatzky, R., Levesque, J. F., Ehrmann-Feldman, D., & Schwartz, C. E. (2014). Minimal evidence of response shift in the absence of a catalyst. *Quality of Life Research*, 23(9), 2421–2430. <https://doi.org/10.1007/s11136-014-0699-3>
- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52(3), 317–332. <https://doi.org/10.1007/BF02294359>
- Albouy, V., Godefroy, P., & Lollivier, S. (2010). Measuring quality of life. *Institut National de La Statistique et Des Études Économiques (INSEE) France Portrait*. http://www.insee.fr/en/ffc/docs_ffc/ref/FPORSOC10E.pdf
- Altman, D. (1991). *Practical Statistics for medical Research* (CHAPMAN &).
- Anota, A., Bascoul-Mollevi, C., Conroy, T., Guillemin, F., Velten, M., Jolly, D., Mercier, M., Causeret, S., Cuisenier, J., Graesslin, O., Hamidou, Z., & Bonnetain, F. (2014). Item response theory and factor analysis as a mean to characterize occurrence of response shift in a longitudinal quality of life study in breast cancer patients. *Health and Quality of Life Outcomes*, 12(1). <https://doi.org/10.1186/1477-7525-12-32>

- Ardila, R. (2003). Calidad de vida: Una definición integradora. *Revista Latinoamericana de Psicología*, 35(2).
- Arons, A. M. M., Krabbe, P. F. M., Van Der Wilt, G. J., Olde Rikkert, M. G. M., & Adang, E. M. M. (2013). Visual analogue scales: Scale recalibration by patients with dementia and their proxies. *Quality of Life Research*, 22(5), 979–986. <https://doi.org/10.1007/s11136-012-0226-3>
- Artavia-Mora, L., Wagner, N., Thiombiano, B. A., & Bedi, A. (2020). Adaptation and biomedical transition of people living with HIV to antiretroviral treatment in Burkina Faso. *Global Public Health*, 15(5), 638–653. <https://doi.org/10.1080/17441692.2019.1704819>
- Arthur, J., Watts, T., Davies, R., Manchaia, V., & Slater, J. (2016). An exploratory study identifying a possible response shift phenomena of the Glasgow hearing aid benefit profile. *Audiology Research*, 6(2), 44–48. <https://doi.org/10.4081/audiores.2016.152>
- Attorresi, H. F., Lozzia, G. S., Abal, F. J. P., Galibert, M. S., & Aguerri, M. E. (2009). Teoría de respuesta al ítem: Conceptos básicos y aplicaciones para la medición de constructos psicológicos. *Revista Argentina de Clínica Psicológica*, 18(2), 179–188.
- Baltes, P. B., & Nesselroade, J. R. (1973). The developmental analysis of individual differences on multiple measures. In J. R. Nesselroade & H. W. Reese, *Life-Span Developmental Psychology: Methodological Issues*.
- Barclay-Goddard, R., Lix, L. M., Tate, R., Weinberg, L., & Mayo, N. E. (2011). Health-related quality of life after stroke: Does response shift occur in self-perceived physical function? *Archives of Physical Medicine and Rehabilitation*, 92(11), 1762–1769. <https://doi.org/10.1016/j.apmr.2011.06.013>
- Barclay-Goddard, R., & Tate, R. B. (2014). Response shift recalibration and reprioritization in health-related quality of life was identified prospectively in older men with and without stroke. *Journal of Clinical Epidemiology*, 67(5), 500–507. <https://doi.org/10.1016/j.jclinepi.2013.12.003>

- Bauer, R. (1966). *Social Indicators*. Cambridge: MIT Press.
<https://www.worldcat.org/title/social-indicators/oclc/165838>
- Bentler, P., & Bonett, D. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In *Statistical theories of mental test scores* (pp. 453–479). Reading, MA: Addison Wesley.
- Blanchin, M., Sébille, V., Guilleux, A., & Hardouin, J. B. (2016). The Guttman errors as a tool for response shift detection at subgroup and item levels. *Quality of Life Research*, 25(6), 1385–1393. <https://doi.org/10.1007/s11136-016-1268-8>
- Blanco-Abarca, A.; Chacón, F. (1985). La evaluación de la calidad de vida. In *Psicología social aplicada* (pp. 183–210).
<https://dialnet.unirioja.es/servlet/articulo?codigo=4594993>
- Blockeel, H., & Struyf, J. (2003). Efficient algorithms for decision tree cross-validation. *Journal of Machine Learning Research*, 3(4–5), 621–650.
- Bloom, B. L. (1978). Social Indicators and health care policy. *Second National Needs Assessment*.
- Bock, R. (1997). A brief history of item response theory. *Educational Measurement*, 16, 21–32.
- Bollen, K. A., & Stine, R. A. (1992). Bootstrapping Goodness-of-Fit Measures in Structural Equation Models. *Sociological Methods & Research*, 21(2), 205–229.
<https://doi.org/10.1177/0049124192021002004>
- Boucekine, M., Loundou, A., Baumstarck, K., Minaya-Flores, P., Pelletier, J., Ghattas, B., & Auquier, P. (2013). Using the random forest method to detect a response shift in the quality of life of multiple sclerosis patients: A cohort study. *BMC Medical Research Methodology*, 13(1). <https://doi.org/10.1186/1471-2288-13-20>
- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and*

- Regression Trees*. Chapman Hall/CRC.
- Brinksma, A., Tissing, W. J. E., Sulkers, E., Kamps, W. A., Roodbol, P. F., & Sanderman, R. (2014). Exploring the response shift phenomenon in childhood patients with cancer and its effect on health-related quality of life. *Oncology Nursing Forum*, *41*(1), 48–56. <https://doi.org/10.1188/14.ONF.41-01AP>
- Burckhart, C., & Anderson, K. (2003). The Quality of Life Scale (QOLS): Reliability, Validity, and Utilization. *Health and Quality of Life Outcomes*. *BioMed Central*, *1*, 60.
- Caballero-Juliá, D., Villardón, M. P. G., & García, M.-C. (2017). JK-Meta-Biplot y STATIS Dual como herramientas de análisis de tablas textuales múltiples. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, *25*, 18–33. <https://doi.org/10.17013/risti.25.18-33>
- Carlier, I. V. E., van Eeden, W. A., de Jong, K., Giltay, E. J., van Noorden, M. S., van der Feltz-Cornelis, C., Zitman, F. G., Kelderman, H., & van Hemert, A. M. (2019). Testing for response shift in treatment evaluation of change in self-reported psychopathology amongst secondary psychiatric care outpatients. *International Journal of Methods in Psychiatric Research*, *28*(3), 1–14. <https://doi.org/10.1002/mpr.1785>
- Carrizo, D., & Moller, C. (2018). Estructuras metodológicas de revisiones sistemáticas de literatura en Ingeniería de Software: un estudio de mapeo sistemático. *Ingeniare. Revista Chilena de Ingeniería*, *26*, 45–54. <https://doi.org/10.4067/s0718-33052018000500045>
- Chen, P. Y., Jan, Y. W., & Yang, C. M. (2017). Are the Insomnia Severity Index and Pittsburgh Sleep Quality Index valid outcome measures for Cognitive Behavioral Therapy for Insomnia? Inquiry from the perspective of response shifts and longitudinal measurement invariance in their Chinese versions. *Sleep Medicine*, *35*, 35–40. <https://doi.org/10.1016/j.sleep.2017.04.003>
- Cook, D., Mulrow, C., & Haynes, R. (1997). Systematic Reviews: Synthesis of Best Evidence for Clinical Decisions. *Annals of Internal Medicine*, *126*(5), 376. <https://doi.org/10.7326/0003-4819-126-5-199703010-00006>

- Cruz-Benito, J. (2016). *Systematic Literature Review & Mapping*. Grupo GRIAL.
- Cummins, R. (2000). Objective and subjective quality of life: An interactive model. *Social Indicators Research*, 52(1), 55–72.
- Curran, P., Bauer, D., & Willoughby, M. (2004). Testing and probing main effects and interactions in latent curve analysis. *Psychological Methods*, 9(2), 220–237.
- Curran, P., & Hussong, A. (2001). Structural equation modeling of repeated measures data. In *Modeling Intraindividual Variability with Repeated Measures Data: Methods and Applications* (pp. 59–86).
- Curran, P., & Hussong, A. (2003). The use of latent trajectory models in psychopathology research. *Journal of Abnormal Psychology*, 112, 526–544.
- Dabakuyo, T. S., Guillemin, F., Conroy, T., Velten, M., Jolly, D., Mercier, M., Causeret, S., Cuisenier, J., Graesslin, O., Gauthier, M., & Bonnetain, F. (2013). Response shift effects on measuring post-operative quality of life among breast cancer patients: A multicenter cohort study. *Quality of Life Research*, 22(1), 1–11. <https://doi.org/10.1007/s11136-012-0135-5>
- DeConde, A. S., Bodner, T. E., Mace, J. C., & Smith, T. L. (2014). Response shift in quality of life after endoscopic sinus surgery for chronic rhinosinusitis. *JAMA Otolaryngology - Head and Neck Surgery*, 140(8), 712–719. <https://doi.org/10.1001/jamaoto.2014.1045>
- Delgado, M., & Llorca, J. (2004). Estudios longitudinales: concepto y particularidades. *Revista Española de Salud Pública*, 78(2), 141–148.
- Dempster, M., Carney, R., & McClements, R. (2010). Response shift in the assessment of quality of life among people attending cardiac rehabilitation. *British Journal of Health Psychology*, 15(2), 307–319. <https://doi.org/10.1348/135910709X464443>
- Dunn, O. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56, 52–64.
- Ebrahim, S. (1995). Clinical and public health perspectives and applications of

- health-related quality of life measurement. *Social Science & Medicine*, 41(10), 1383–1394. <http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=2891195>
- Edelaar-Peeters, Y., & Stiggelbout, A. M. (2013). Anticipated adaptation or scale recalibration? *Health and Quality of Life Outcomes*, 11(1), 1–10. <https://doi.org/10.1186/1477-7525-11-171>
- Elliott, B. A., Gessert, C. E., Larson, P. M., & Russ, T. E. (2014). Shifting responses in quality of life: People living with dialysis. *Quality of Life Research*, 23(5), 1497–1504. <https://doi.org/10.1007/s11136-013-0600-9>
- Emmerich, W. (1968). Personality Development and Concepts of Structure. *Child Development*, 39(3), 671. <https://doi.org/10.2307/1126978>
- Escoufier, Y. (1976a). Opérateur associé à un tableau de données. *Annales de l'insée*, 22/23, 165. <https://doi.org/10.2307/20075217>
- Escoufier, Y. (1976b). Opérateur associé à un tableau de données. *Annales de l'insée*, 29(22/23), 165. <https://doi.org/10.2307/20075217>
- Eton, D. T. (2010). Why we need response shift: an appeal to functionalism. *Quality of Life Research*, 19(6), 929–930. <https://doi.org/10.1007/s11136-010-9684-7>
- Felix, J., Becker, C., Vogl, M., Buschner, P., Plötz, W., & Leidl, R. (2019). Patient characteristics and valuation changes impact quality of life and satisfaction in total knee arthroplasty - Results from a German prospective cohort study. *Health and Quality of Life Outcomes*, 17(1), 1–10. <https://doi.org/10.1186/s12955-019-1237-3>
- Fernández-Ballesteros, R. (1998). Quality of life: The differential conditions. *Psychology in Spain*, 2, 57–65.
- Ferreira-González, I., Urrútia, G., & Alonso-Coello, P. (2011). Systematic reviews and meta-analysis: Scientific rationale and interpretation. *Revista Espanola de Cardiologia*, 64(8), 688–696. <https://doi.org/10.1016/j.recesp.2011.03.029>
- Flynn, T. N., Peters, T. J., & Coast, J. (2013). Quantifying response shift or

- adaptation effects in quality of life by synthesising best-worst scaling and discrete choice data. *Journal of Choice Modelling*, 6, 34–43.
<https://doi.org/10.1016/j.jocm.2013.04.004>
- Fokkema, M., Smits, N., Kelderman, H., & Cuijpers, P. (2013). Response shifts in mental health interventions: An illustration of longitudinal measurement invariance. *Psychological Assessment*, 25(2), 520–531.
<https://doi.org/10.1037/a0031669>
- Friedrich, M., Karoff, J., & Hinz, A. (2019). Response shift effects in patients' assessments of their quality of life after cardiac rehabilitation. *Quality of Life Research*, 28(9), 2609–2620. <https://doi.org/10.1007/s11136-019-02195-9>
- Friedrich, M., Zenger, M., & Hinz, A. (2018). Response shift effects of quality of life assessments in breast cancer survivors. *European Journal of Cancer Care*, 28(2), 1–9. <https://doi.org/10.1111/ecc.12979>
- Galenkamp, H., Huisman, M., Braam, A. W., & Deeg, D. J. H. (2012). Estimates of prospective change in self-rated health in older people were biased owing to potential recalibration response shift. *Journal of Clinical Epidemiology*, 65(9), 978–988. <https://doi.org/10.1016/j.jclinepi.2012.03.010>
- Galindo-Villardón, M., & Vicente-Galindo, M. (2004). Calidad de vida relacionada con la Salud: un constructo multidimensional dinámico. *Metodología de Las Ciencias Del Comportamiento, Suplemento*(January 2004), 219–227.
- Gandhi, P. K., Ried, L. D., Kimberlin, C. L., Kauf, T. L., & Huang, I. C. (2013). Influence of explanatory and confounding variables on HRQoL after controlling for measurement bias and response shift in measurement. *Expert Review of Pharmacoeconomics and Outcomes Research*, 13(6), 841–851.
<https://doi.org/10.1586/14737167.2013.852959>
- Gandhi, P. K., Schwartz, C. E., Reeve, B. B., DeWalt, D. A., Gross, H. E., & Huang, I. C. (2016). An item-level response shift study on the change of health state with the rating of asthma-specific quality of life: a report from the PROMIS®Pediatric Asthma Study. *Quality of Life Research*, 25(6), 1349–1359.
<https://doi.org/10.1007/s11136-016-1290-x>

- Gandhi, P., Ried, L., Huang, I., Kimberlin, C., & Kauf, T. (2013). Assessment of response shift using two structural equation modeling techniques. *Quality of Life Research*, 22(3), 461–471. <https://doi.org/10.1007/s11136-012-0171-1>
- Gerlich, C., Schuler, M., Jelitte, M., Neuderth, S., Flentje, M., Graefen, M., Krüger, A., Mehnert, A., & Faller, H. (2016). Prostate cancer patients' quality of life assessments across the primary treatment trajectory: 'True' change or response shift? *Acta Oncologica*, 55(7), 814–820. <https://doi.org/10.3109/0284186X.2015.1136749>
- Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*, 73(1), 43–56. <https://doi.org/10.1093/biomet/73.1.43>
- Golembiewski, R., Billingsley, K., & Yeager, S. (1976). Measuring Change and Persistence in Human Affairs: Types of Change Generated by OD Designs. *The Journal of Applied Behavioral Science*, 12(2), 133–157. <https://doi.org/10.1177/002188637601200201>
- González, U. (2002). El concepto de calidad de vida y la evolución de los paradigmas de las ciencias de la salud. *Revista Cubana de Salud Publica*, 28(2), 157–175. http://aprendeenlinea.udea.edu.co/lms/moodle/file.php/531/Memoria_textos_linea_I/Unidad_1/el_concepto_de_calidad_de_vida_y_la_evolucion_de_los_paradigmas.pdf
- Grubbs, J. R., Tolleson-Rinehart, S., Huynh, K., & Davis, R. M. (2014). A Review of Quality of Life Measures in Dry Eye Questionnaires. *Cornea*, 33(2), 215–218. <https://doi.org/10.1097/ICO.0000000000000038>
- Guilleux, A., Blanchin, M., Vanier, A., Guillemin, F., Falissard, B., Schwartz, C. E., Hardouin, J. B., & Sébille, V. (2015). RespOnse Shift ALgorithm in Item response theory (ROSALI) for response shift detection with missing data in longitudinal patient-reported outcome studies. *Quality of Life Research*, 24(3), 553–564. <https://doi.org/10.1007/s11136-014-0876-4>
- H. L'Hermier Des Plantes. (1976). *Structuration des tableaux à trois indices de la*

Statistique : théorie et application d'une méthode d'analyse conjointe.

Université des Sciences et Techniques du Languedoc (USTL).

Haagsma, J. A., Spronk, I., De Jongh, M. A. C., Bonsel, G. J., & Polinder, S. (2020). Conventional and retrospective change in health-related quality of life of trauma patients: An explorative observational follow-up study. *Health and Quality of Life Outcomes*, *18*(1), 1–13. <https://doi.org/10.1186/s12955-020-01404-1>

Haas, B. K. (1999). A Multidisciplinary Concept Analysis of Quality of Life. *Western Journal of Nursing Research*, *21*(6), 728–742. <https://doi.org/10.1177/01939459922044153>

Hair, J., Anderson, R., Tatham, R., & Black, W. (1999). *Análisis Multivariante*. Prentice Hall Iberia.

Hamidou, Z., Dabakuyo-Yonli, T. S., Guillemin, F., Conroy, T., Velten, M., Jolly, D., Causeret, S., Graesslin, O., Gauthier, M., Mercier, M., & Bonnetain, F. (2014). Impact of response shift on time to deterioration in quality of life scores in breast cancer patients. *PLoS ONE*, *9*(5). <https://doi.org/10.1371/journal.pone.0096848>

Hamidou, Z., Dabakuyo, T. S., & Bonnetain, F. (2011). Impact of response shift on longitudinal quality-of-life assessment in cancer clinical trials. *Expert Review of Pharmacoeconomics & Outcomes Research*, *11*(5), 549–559. <https://doi.org/10.1586/erp.11.57>

Hamilton, M. (1967). Comparison of factors by Ahmavaara's method. *British Journal of Mathematical and Statistical Psychology*, *20*, 107–110.

Hill, L. G., & Betz, D. L. (2005). Revisiting the retrospective pretest. *American Journal of Evaluation*, *26*(4), 501–517. <https://doi.org/10.1177/1098214005281356>

Hinds, A., Sajobi, T., Sebille, V., Sawatzky, R., & Lix, L. (2018). A systematic review of the quality of reporting of simulation studies about methods for the analysis of complex longitudinal patient-reported outcomes data. *Quality of Life Research*, *27*(10), 2507–2516. <https://doi.org/10.1007/s11136-018-1861-0>

- Hinz, A., Finck Barboza, C., Zenger, M., Singer, S., Schwalenberg, T., & Stolzenburg, J. U. (2011). Response shift in the assessment of anxiety, depression and perceived health in urologic cancer patients: An individual perspective. *European Journal of Cancer Care*, *20*(5), 601–609. <https://doi.org/10.1111/j.1365-2354.2011.01256.x>
- Hinz, A., Mehnert, A., Dégi, C., Reissmann, D. R., Schotte, D., & Schulte, T. (2017). The relationship between global and specific components of quality of life, assessed with the EORTC QLQ-C30 in a sample of 2019 cancer patients. *European Journal of Cancer Care*, *26*(2), 1–8. <https://doi.org/10.1111/ecc.12416>
- Hochberg, Y. (1988). A Sharper Bonferroni produce for multiple tests of significance. *Biometrika*, *75*, 800–802.
- Höfer, S., Pfaffenberger, N., Renn, D., Platter, M., & Ring, L. (2011). Coronary Intervention Improves Disease Specific Health-Related Quality of Life but Not Individualised Quality of Life: A Potential Response Shift Effect? *Applied Research in Quality of Life*, *6*(1), 81–90. <https://doi.org/10.1007/s11482-010-9117-3>
- Hollman, F., Wessel, R. N., & Wolterbeek, N. (2016). Response shift of the Western Ontario Rotator Cuff index in patients undergoing arthroscopic rotator cuff repair. *Journal of Shoulder and Elbow Surgery*, *25*(12), 2011–2018. <https://doi.org/10.1016/j.jse.2016.05.012>
- Hosseini, B., Nedjat, S., Zendehdel, K., Majdzadeh, R., Nourmohammadi, A., & Montazeri, A. (2017). Response shift in quality of life assessment among cancer patients: A study from Iran. *Medical Journal of the Islamic Republic of Iran*, *31*(1). <https://doi.org/10.14196/mjiri.31.120>
- Hosseini, K., Omorou, A. Y., Hubert, J., Ngueyon Sime, W., Ladrière, M., & Guillemin, F. (2017). Nephrectomy Complication Is a Risk Factor of Clinically Meaningful Decrease in Health Utility among Living Kidney Donors. *Value in Health*, *20*(10), 1376–1382. <https://doi.org/10.1016/j.jval.2017.05.022>
- Howard, G., & Dailey, P. (1979). Response-shift bias: A source of contamination of

- self-report measures. *Journal of Applied Psychology*, 64(2), 144–150.
<https://doi.org/10.1037/0021-9010.64.2.144>
- Howard, J., Mattacola, C., Howell, D., & Lattermann, C. (2011). Response shift theory: An application for health-related quality of life in rehabilitation research and practice. *Journal of Allied Health*, 40(1), 31–38.
- Howard, J. S., Mattacola, C. G., Mullineaux, D. R., English, R. A., & Lattermann, C. (2014). Influence of response shift on early patient-reported outcomes following autologous chondrocyte implantation. *Knee Surgery, Sports Traumatology, Arthroscopy*, 22(9), 2163–2171. <https://doi.org/10.1007/s00167-013-2654-1>
- Huberty, C., & Wisenbaker, J. (1992). Variable importance in multivariate group comparisons. *Journal of Educational Statistics*, 17, 75–91.
- Huppert, F. A., & Whittington, J. E. (2003). Evidence for the independence of positive and negative well-being: Implications for quality of life assessment. *British Journal of Health Psychology*, 8(1), 107–122.
<https://doi.org/10.1348/135910703762879246>
- Ilie, G., Bradfield, J., Moodie, L., Lawen, T., Ilie, A., Lawen, Z., Blackman, C., Gainer, R., & Rutledge, R. (2019a). The Role of Response-Shift in Studies Assessing Quality of Life Outcomes Among Cancer Patients: A Systematic Review. *Frontiers in Oncology*, 9. <https://doi.org/10.3389/fonc.2019.00783>
- Ilie, G., Bradfield, J., Moodie, L., Lawen, T., Ilie, A., Lawen, Z., Blackman, C., Gainer, R., & Rutledge, R. D. H. (2019b). The role of response-shift in studies assessing quality of life outcomes among cancer patients: A systematic review. *Frontiers in Oncology*, 9(AUG). <https://doi.org/10.3389/fonc.2019.00783>
- INCMSZ. (2013). ¿Qué es Calidad de Vida? *Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán*.
- Ito, N., Ishiguro, M., Tanaka, M., Tokunaga, K., Sugihara, K., & Kazuma, K. (2010). Response shift in quality-of-life assessment in patients undergoing curative surgery with permanent colostomy: A preliminary study. *Gastroenterology Nursing*, 33(6), 408–412. <https://doi.org/10.1097/SGA.0b013e318200bf30>

- Jabrayilov, R., Emons, W. H. M., de Jong, K., & Sijtsma, K. (2017). Longitudinal measurement invariance of the Dutch Outcome Questionnaire-45 in a clinical sample. *Quality of Life Research*, 26(6), 1473–1481. <https://doi.org/10.1007/s11136-017-1500-1>
- Jakola, A. S., Solheim, O., Gulati, S., & Sagberg, L. M. (2017). Is there a response shift in generic health-related quality of life 6 months after glioma surgery? *Acta Neurochirurgica*, 159(2), 377–384. <https://doi.org/10.1007/s00701-016-3040-9>
- Johnson, J. W., & Lebreton, J. M. (2004). History and Use of Relative Importance Indices in Organizational Research. *Organizational Research Methods*, 7(3), 238–257. <https://doi.org/10.1177/1094428104266510>
- Jöreskog, K. (1993). *Testing structural equation models*.
- Jöreskog, K., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago: Scientific Software International.
- Kagan, J. (1980). *Perspectives on continuity*. In O.G. Bin, Jr. & J. Kagan (Eds.) *Constancy and change in human development*. Harvard University Press.
- Kay, S., & Ferreira, A. (2014). Mapping the 25-item National Eye Institute Visual Functioning Questionnaire (NEI VFQ-25) to EQ-5D Utility Scores. *Ophthalmic Epidemiology*, 21(2), 66–78. <https://doi.org/10.3109/09286586.2014.888456>
- Kievit, W., Hendriks, J., Stalmeier, P. F. M., Van De Laar, M. A. F. J., Van Riel, P. L. C. M., & Adang, E. M. (2010). The relationship between change in subjective outcome and change in disease: A potential paradox. *Quality of Life Research*, 19(7), 985–994. <https://doi.org/10.1007/s11136-010-9665-x>
- Kimura, A., Arakawa, H., Noda, K., Yamazaki, S., Hara, E. S., Mino, T., Matsuka, Y., Mulligan, R., & Kuboki, T. (2012). Response shift in oral health-related quality of life measurement in patients with partial edentulism. *Journal of Oral Rehabilitation*, 39(1), 44–54. <https://doi.org/10.1111/j.1365-2842.2011.02241.x>
- King-Kallimanis, B. L., Oort, F. J., Nolte, S., Schwartz, C. E., & Sprangers, M. A. G. (2011). Using structural equation modeling to detect response shift in

- performance and health-related quality of life scores of multiple sclerosis patients. *Quality of Life Research*, 20(10), 1527–1540.
<https://doi.org/10.1007/s11136-010-9844-9>
- King-kallimanis, B., & Oort, F. (2010). *Using structural equation modelling to detect measurement bias and response shift in longitudinal*. 139–156.
<https://doi.org/10.1007/s10182-010-0129-y>
- Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, 51(1), 7–15.
<https://doi.org/10.1016/j.infsof.2008.09.009>
- Kline, R. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford Press.
- Krägeloh, C. U., Bergomi, C., Siegert, R. J., & Medvedev, O. N. (2018). Response Shift After a Mindfulness-Based Intervention: Measurement Invariance Testing of the Comprehensive Inventory of Mindfulness Experiences. *Mindfulness*, 9(1), 212–220. <https://doi.org/10.1007/s12671-017-0764-4>
- Kruskal, W., & Majors, R. (1989). Concepts of relative importance in recent scientific literature. *American Statistician*, 42, 2–6.
- Krzanowski, W. J. (1979). Between-Groups Comparison of Principal Components. *Journal of the American Statistical Association*, 74(367), 703–707.
<https://doi.org/10.1080/01621459.1979.10481674>
- Kuhn, F., Morris, R., Witherspoon, C. D., Heimann, K., Jeffers, J. B., & Treister, G. (1996). A Standardized Classification of Ocular Trauma. *Ophthalmology*, 103(2), 240–243. [https://doi.org/10.1016/S0161-6420\(96\)30710-0](https://doi.org/10.1016/S0161-6420(96)30710-0)
- Kvam, A. K., Wisløff, F., & Fayers, P. M. (2010). Minimal important differences and response shift in health-related quality of life; a longitudinal study in patients with multiple myeloma. *Health and Quality of Life Outcomes*, 8, 1–8.
<https://doi.org/10.1186/1477-7525-8-79>
- L’Hermier des Plantes, H. (1976). *Structuration des tableaux à trois indices de la*

statistique. Université des Montpellier.

- Lacey, H. P., Loewenstein, G., & Ubel, P. A. (2011). Compared to what? A joint evaluation method for assessing quality of life. *Quality of Life Research, 20*(8), 1169–1177. <https://doi.org/10.1007/s11136-011-9856-0>
- Lemon, S., Roy, J., Clark, M., Friedmann, P., & Rakowski, W. (2003). Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med., 26*(3), 172–181.
- Lepore, S. J., & Eton, D. T. (2000). Response shifts in prostate cancer patients: An evaluation of suppressor and buffer models. In *Adaptation to changing health: Response shift in quality-of-life research*. (pp. 37–51). American Psychological Association. <https://doi.org/10.1037/10382-003>
- Lévy, J. (2003). *Modelización y análisis con ecuaciones estructurales*. Prentice Hall.
- Li, Y., & Rapkin, B. (2009). HIV/AIDS Patients' Quality of Life Appraisal Depends on Their Personal Meaning Of Quality of Life and Frame of Reference. *Quality of Life Research, A-36*.
- Li, Y., & Schwartz, C. (2011). Data mining for response shift patterns in multiple sclerosis patients using recursive partitioning tree analysis. *Quality of Life Research, 20*(10), 1543–1553. <https://doi.org/10.1007/s11136-011-0004-7>
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ, 339*(jul21 1), b2700–b2700. <https://doi.org/10.1136/bmj.b2700>
- Liu, J. J., & Davis, G. E. (2015). The significance of response shift in sinus surgery outcomes. *International Forum of Allergy and Rhinology, 5*(1), 55–59. <https://doi.org/10.1002/alr.21420>
- Lix, L. M., Chan, E. K. H., Sawatzky, R., Sajobi, T. T., Liu, J., Hopman, W., & Mayo, N. (2016). Response shift and disease activity in inflammatory bowel disease. *Quality of Life Research, 25*(7), 1751–1760.

<https://doi.org/10.1007/s11136-015-1188-z>

- Lix, L. M., Sajobi, T. T., Sawatzky, R., Liu, J., Mayo, N. E., Huang, Y., Graff, L. A., Walker, J. R., Ediger, J., Clara, I., Sexton, K., Carr, R., & Bernstein, C. N. (2013). Relative importance measures for reprioritization response shift. *Quality of Life Research*, 22(4), 695–703. <https://doi.org/10.1007/s11136-012-0198-3>
- MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, 100(1), 107–120. <https://doi.org/10.1037/0033-2909.100.1.107>
- Machuca, C., Vettore, M. V., Krasuska, M., Baker, S. R., & Robinson, P. G. (2017). Using classification and regression tree modelling to investigate response shift patterns in dentine hypersensitivity. *BMC Medical Research Methodology*, 17(1), 120. <https://doi.org/10.1186/s12874-017-0396-3>
- Machuca, C., Vettore, M. V., & Robinson, P. G. (2020). How peoples' ratings of dental implant treatment change over time? *Quality of Life Research*, 29(5), 1323–1334. <https://doi.org/10.1007/s11136-019-02408-1>
- Manchikanti, L., Datta, S., H., S., & Hirsch, J. (2009). Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: part 6. Systematic reviews and meta-analyses of observational studies. *Pain Physician*, 12(5), 819–850.
- Martín-Rodero, H. (2014). *La búsqueda bibliográfica, pilar fundamental de la medicina basada en la evidencia: evaluación multivariante de las enfermedades nutricionales y metabólicas*. Universidad Miguel Hernandez, España.
- Martín-Rodríguez, J. (1996). *Contribuciones a la Integración de Subespacios desde una Perspectiva Biplot*. Universidad de Salamanca.
- Martín-Rodríguez, J., Galindo-Villardón, M. P., & Vicente-Villardón, J. L. (2002). Comparison and integration of subspaces from a biplot perspective. *Journal of Statistical Planning and Inference*, 102(2), 411–423. [https://doi.org/10.1016/S0378-3758\(01\)00101-X](https://doi.org/10.1016/S0378-3758(01)00101-X)
- Martín-Rodríguez, J., & Vicente-Villardón, J. L. (2004). Metabiplot: Application to

- the study of a Lake. *Advance in Multivariate Analysis*, 93–96.
- Martínez-Arias, M. (1995). *Psicometría: Teoría de los Tests Psicológicos y Educativos*.
- Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 42, 149–174.
- Mayo, N. E., Scott, S. C., Bernstein, C. N., & Lix, L. M. (2015). How are you? Do people with inflammatory bowel disease experience response shift on this question? *Health and Quality of Life Outcomes*, 13(1), 1–11.
<https://doi.org/10.1186/s12955-015-0232-6>
- Mayo, N. E., Scott, S. C., Dendukuri, N., Ahmed, S., & Wood-Dauphinee, S. (2008). Identifying response shift statistically at the individual level. *Quality of Life Research*, 17(4), 627–639. <https://doi.org/10.1007/s11136-008-9329-2>
- McArdle, J., & Epstein, D. (1998). Latent growth curves within developmental structural equation models. *Child Development*, 58(1), 110–133.
- McPhail, S., & Haines, T. (2010). Response shift, recall bias and their effect on measuring change in health-related quality of life amongst older hospital patients. *Health and Quality of Life Outcomes*, 8, 1–9.
<https://doi.org/10.1186/1477-7525-8-65>
- Meade, M., & Richardson, W. (1997). Selecting and appraising studies for a systematic review. *Ann Intern Med*, 127(7), 531–537.
- Meeberg, G. A. (1993). Quality of life: a concept analysis. *Journal of Advanced Nursing*, 18(1), 32–38. <https://doi.org/10.1046/j.1365-2648.1993.18010032.x>
- Menard, S. (2004). Six approaches to calculating standardized logistic regression coefficients. *The American Statistician*, 58, 218–223.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55(1), 107–122. <https://doi.org/10.1007/BF02294746>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of*

- Clinical Epidemiology*, 62(10), 1006–1012.
<https://doi.org/10.1016/j.jclinepi.2009.06.005>
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1), 1. <https://doi.org/10.1186/2046-4053-4-1>
- Mollerup, A., & Johansen, J. D. (2015). Response shift in severity assessment of hand eczema with visual analogue scales. *Contact Dermatitis*, 72(3), 178–183. <https://doi.org/10.1111/cod.12335>
- Montori, V., Swiontkowski, M., & Cook, D. (2003). Methodologic Issues in Systematic Reviews and Meta-Analyses. *Clinical Orthopaedics & Related Research*, 413, 43–54. <https://doi.org/10.1097/01.blo.0000079322.41006.5b>
- Moreno, B., & Ximenez, C. (1996). *Evaluación de la calidad de vida* (Madrid: Siglo Veintiuno (ed.)). <https://dialnet.unirioja.es/servlet/libro?codigo=563965>
- Mortimer, J., Finch, M., & Kunka, D. (1982). Persistence and change in development: The multidimensional self-concept. *Life Span Development and Behavior*, 4, 263–313.
- Muraki, E. (1992). A generalized partial credit model. Application of on Em algorithm. *Applied Psychological Measurement*, 16, 351–363.
- Murata, T., Suzukamo, Y., Shiroywa, T., Taira, N., & Shimosuma, K. (2018). Pcn377 - Response Shift Effect on Quality of Life in Randomized Control Trial of Taxane Versus S-1 for Metastatic Breast Cancer Patients (Select-Bc): Structural Equation Modeling. *Value in Health*, 21, S78. <https://doi.org/10.1016/j.jval.2018.09.459>
- Murata, T., Suzukamo, Y., Shiroywa, T., Taira, N., Shimosuma, K., Ohashi, Y., & Mukai, H. (2020a). Response Shift–Adjusted Treatment Effect on Health-Related Quality of Life in a Randomized Controlled Trial of Taxane Versus S-1 for Metastatic Breast Cancer: Structural Equation Modeling. *Value in Health*, 23(6), 768–774. <https://doi.org/10.1016/j.jval.2020.02.003>

- Murata, T., Suzukamo, Y., Shiroywa, T., Taira, N., Shimozuma, K., Ohashi, Y., & Mukai, H. (2020b). Response Shift–Adjusted Treatment Effect on Health-Related Quality of Life in a Randomized Controlled Trial of Taxane Versus S-1 for Metastatic Breast Cancer: Structural Equation Modeling. *Value in Health*, 23(6), 768–774. <https://doi.org/10.1016/j.jval.2020.02.003>
- Muthén, B. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 1–17). American Psychological Association.
- Nakamura, K., Shimozuma, K., Suzukamo, Y., Taira, N., Shiroywa, T., Shibahara, H., & Saito, S. (2013). Using Structural Equation Modeling to Detect Response Shift and True Change in Health-Related Quality-Of-Life Scores of Breast Cancer Patients After Surgery. *Value in Health*, 16(7), A602. <https://doi.org/10.1016/j.jval.2013.08.1709>
- Neuman, H. B., Park, J., Fuzesi, S., & Temple, L. K. (2012). Rectal cancer patients' quality of life with a temporary stoma: Shifting perspectives. *Diseases of the Colon and Rectum*, 55(11), 1117–1124. <https://doi.org/10.1097/DCR.0b013e3182686213>
- Nichols, G. A. L., Antoun, J. S., Fowler, P. V., Al-Ani, A. H., & Farella, M. (2018). Long-term changes in oral health-related quality of life of standard, cleft, and surgery patients after orthodontic treatment: A longitudinal study. *American Journal of Orthodontics and Dentofacial Orthopedics*, 153(2), 224–231. <https://doi.org/10.1016/j.ajodo.2017.06.018>
- Nielsen, L., Jarden, M., Andersen, C., Frederiksen, H., & Abildgaard, N. (2017). A systematic review of health-related quality of life in longitudinal studies of myeloma patients. *European Journal of Haematology*, 99(1), 3–17. <https://doi.org/10.1111/ejh.12882>
- Nolte, S., Mierke, A., Fischer, H. F., & Rose, M. (2016). On the validity of measuring change over time in routine clinical assessment: a close examination of item-level response shifts in psychosomatic inpatients. *Quality of Life*

- Research*, 25(6), 1339–1347. <https://doi.org/10.1007/s11136-015-1123-3>
- Nunes, S. M. (2009). *Evaluación del cambio en pacientes oncológicos: un enfoque estadístico* [Universidad de Salamanca. España].
<https://dialnet.unirioja.es/servlet/tesis?codigo=22210>
- Nunes, S. M., & Galindo-Villardón, M. P. (2006). Statis e Metabiplot. Un estudio comparativo. *XIII Congreso Anual Da SPE*.
- Oliver, J., Rosel, J., & Jara, P. (2000). Modelos de regresión multinivel: aplicación en psicología escolar. *Psicothema*, 12(3), 487–494.
- Oort, F. J. (2005). Using structural equation modeling to detect response shifts and true change. *Quality of Life Research*, 14(3), 587–598.
<https://doi.org/10.1007/s11136-004-0830-y>
- Ousmen, A., Conroy, T., Guillemin, F., Velten, M., Jolly, D., Mercier, M., Causeret, S., Cuisenier, J., Graesslin, O., Hamidou, Z., Bonnetain, F., & Aota, A. (2016a). Impact of the occurrence of a response shift on the determination of the minimal important difference in a health-related quality of life score over time. *Health and Quality of Life Outcomes*, 14(1), 1–15.
<https://doi.org/10.1186/s12955-016-0569-5>
- Ousmen, A., Conroy, T., Guillemin, F., Velten, M., Jolly, D., Mercier, M., Causeret, S., Cuisenier, J., Graesslin, O., Hamidou, Z., Bonnetain, F., & Aota, A. (2016b). Impact of the occurrence of a response shift on the determination of the minimal important difference in a health-related quality of life score over time. *Health and Quality of Life Outcomes*, 14(1), 167.
<https://doi.org/10.1186/s12955-016-0569-5>
- Palomba, R. (2002). Calidad de Vida : Conceptos y medidas Concepto de calidad de vida. *Calidad de Vida: Conceptos y Medidas*, 1–12.
http://www.cepal.org/celade/agenda/2/10592/envejecimientoRP1_ppt.pdf
- Perdomo-Arguello, F. (2015). *Evaluación del cambio en la calidad de vida de pacientes miopes operados con cirugía de Láser Excimer corneal*. Universidad de Panamá.

- Phillips, R., Gandhi, M., Cheung, Y. B., Findlay, M. P., Win, K. M., Hai, H. H., Yang, J. M., Lobo, R. R., Soo, K. C., & Chow, P. K. H. (2015). Summary scores captured changes in subjects' QoL as measured by the multiple scales of the EORTC QLQ-C30. *Journal of Clinical Epidemiology*, *68*(8), 895–902. <https://doi.org/10.1016/j.jclinepi.2015.02.011>
- Powden, C. J., Hoch, M. C., & Hoch, J. M. (2018). Examination of Response Shift After Rehabilitation for Orthopedic Conditions: A Systematic Review. *Journal of Sport Rehabilitation*, *27*(5), 469–479. <https://doi.org/10.1123/jsr.2017-0104>
- Powden, C. J., Hoch, M. C., Jamali, B. E., & Hoch, J. M. (2019). Response shift after a 4-week multimodal intervention for chronic ankle instability. *Journal of Athletic Training*, *54*(4), 397–402. <https://doi.org/10.4085/1062-6050-345-17>
- Powell, G. A., Adair, C. E., Streiner, D. L., Mayo, N., & Latimer, E. (2017). Changes in quality of life from a homelessness intervention: true change, response shift, or random variation. *Quality of Life Research*, *26*(7), 1853–1864. <https://doi.org/10.1007/s11136-017-1522-8>
- Preiß, M., Friedrich, M., Stolzenburg, J. U., Zenger, M., & Hinz, A. (2019a). Response shift effects in the assessment of urologic cancer patients' quality of life. *European Journal of Cancer Care*, *28*(4), 1–8. <https://doi.org/10.1111/ecc.13027>
- Preiß, M., Friedrich, M., Stolzenburg, J., Zenger, M., & Hinz, A. (2019b). Response shift effects in the assessment of urologic cancer patients' quality of life. *European Journal of Cancer Care*, *28*(4). <https://doi.org/10.1111/ecc.13027>
- Quaranta, L., Riva, I., Gerardi, C., Oddone, F., Floriano, I., & Konstas, A. G. P. (2016). Quality of Life in Glaucoma: A Review of the Literature. *Advances in Therapy*, *33*(6), 959–981. <https://doi.org/10.1007/s12325-016-0333-6>
- Radespiel-Troger, M., Rabenstein, T., Schneider, H., & Lausen, B. (2003). Comparison of tree-based methods for prognostic stratification of survival data. *Artificial Intelligence*, *28*, 323–341.
- Ramírez, R. (2007). Calidad de vida relacionada con la salud como medida de

- resultados en salud: revisión sistemática de la literatura. *Revista Colombiana de Cardiología*, 14(4), 207–222.
- Rapkin, B. D., & Schwartz, C. E. (2004). Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift. *Health and Quality of Life Outcomes*, 2(May 2014). <https://doi.org/10.1186/1477-7525-2-14>
- Rasch, G. (1960). *Probabilistic Models for some intelligence and attainment test*. Copenhagen:Danish Institute for Educational Research.
- Razmjou, H., Schwartz, C. E., & Holtby, R. (2010). The Impact of Response Shift on Perceived Disability Two Years Following Rotator Cuff Surgery. *The Journal of Bone and Joint Surgery-American Volume*, 92(12), 2178–2186. <https://doi.org/10.2106/JBJS.I.00990>
- Reeve, B. B. (2010). An opportunity to refine our understanding of “response shift” and to educate researchers on designing quality research studies: response to Ubel, Peeters, and Smith. *Quality of Life Research*, 19(4), 473–475. <https://doi.org/10.1007/s11136-010-9612-x>
- Reissmann, D. R., Erler, A., Hirsch, C., Sierwald, I., Machuca, C., & Schierz, O. (2018a). Bias in retrospective assessment of perceived dental treatment effects when using the Oral Health Impact Profile. *Quality of Life Research*, 27(3), 775–782. <https://doi.org/10.1007/s11136-017-1725-z>
- Reissmann, D. R., Erler, A., Hirsch, C., Sierwald, I., Machuca, C., & Schierz, O. (2018b). Bias in retrospective assessment of perceived dental treatment effects when using the Oral Health Impact Profile. *Quality of Life Research*, 27(3), 775–782. <https://doi.org/10.1007/s11136-017-1725-z>
- Reissmann, D. R., John, M. T., Feuerstahler, L., Baba, K., Szabó, G., Čelebić, A., & Waller, N. (2016). Longitudinal measurement invariance in prospective oral health-related quality of life assessment. *Health and Quality of Life Outcomes*, 14(1), 1–9. <https://doi.org/10.1186/s12955-016-0492-9>
- Reissmann, D. R., Remmler, A., John, M. T., Schierz, O., & Hirsch, C. (2012).

- Impact of response shift on the assessment of treatment effects using the Oral Health Impact Profile. *European Journal of Oral Sciences*, 120(6), 520–525.
<https://doi.org/10.1111/j.1600-0722.2012.00998.x>
- Rios, R. (2006). *Aportaciones al Análisis Dinámico de Tablas de Tres Vías*.
Universidad de Salamanca, España.
- Rohn, E. J., Tate, D. G., Forchheimer, M., & DiPonio, L. (2019). Contextualizing the lived experience of quality of life for persons with spinal cord injury: A mixed-methods application of the response shift model. *Journal of Spinal Cord Medicine*, 42(4), 469–477. <https://doi.org/10.1080/10790268.2018.1517471>
- Rojas, M. (2014). *The Measurement of Quality of Life : Conceptualization Comes First A Four- Qualities-of-Life Conceptual Framework and an Illustration to Latin America ISQOLS Conference : Measures and goals for the progress of societies Satellite meeting Measuring subject. August 2013*.
- Ruiz, M., Pardo, A., & San Martín, R. (2010). MODELOS DE ECUACIONES ESTRUCTURALES. *Papeles Del Psicólogo*, 31(1), 34–45.
- Rutgers, M., Creemers, L. B., Yang, K. G. A., Raijmakers, N. J. H., Dhert, W. J. A., & Saris, D. B. F. (2015a). Osteoarthritis treatment using autologous conditioned serum after placebo: Patient considerations and clinical response in a non-randomized case series. *Acta Orthopaedica*, 86(1), 114–118.
<https://doi.org/10.3109/17453674.2014.950467>
- Rutgers, M., Creemers, L. B., Yang, K. G. A., Raijmakers, N. J. H., Dhert, W. J. A., & Saris, D. B. F. (2015b). Osteoarthritis treatment using autologous conditioned serum after placebo. *Acta Orthopaedica*, 86(1), 114–118.
<https://doi.org/10.3109/17453674.2014.950467>
- Sáenz, A. (2001). Leer e interpretar una revisión sistemática. *Boletín de Ls Sociedad de Pediatría de Asturias.*, 215–221.
- Sajobi, T. T., Brahmabatt, R., Lix, L. M., Zumbo, B. D., & Sawatzky, R. (2018). Scoping review of response shift methods: current reporting practices and recommendations. *Quality of Life Research*, 27(5), 1133–1146.

<https://doi.org/10.1007/s11136-017-1751-x>

- Sajobi, T. T., Fiest, K. M., & Wiebe, S. (2014). Changes in quality of life after epilepsy surgery: The role of reprioritization response shift. *Epilepsia*, *55*(9), 1331–1338. <https://doi.org/10.1111/epi.12697>
- Sajobi, T. T., Lix, L. M., Li, L., & Laverty, W. (2011). Discriminant analysis for repeated measures data: Effects of mean and covariance misspecification on bias and error in discriminant function coefficients. *Journal of Modern Applied Statistical Methods*, *10*(2), 571–582. <https://doi.org/10.22237/jmasm/1320120840>
- Sajobi, T. T., Lix, L. M., Singh, G., Lowerison, M., Engbers, J., & Mayo, N. E. (2015). Identifying reprioritization response shift in a stroke caregiver population: a comparison of missing data methods. *Quality of Life Research*, *24*(3), 529–540. <https://doi.org/10.1007/s11136-014-0824-3>
- Sajobi, T. T., Speechley, K. N., Liang, Z., Goodwin, S. W., Ferro, M. A., & Wiebe, S. (2017). Response shift in parents' assessment of health-related quality of life of children with new-onset epilepsy. *Epilepsy and Behavior*, *75*, 97–101. <https://doi.org/10.1016/j.yebeh.2017.07.015>
- Salmon, M., Blanchin, M., Rotonda, C., Guillemin, F., & Sébille, V. (2017a). Identifying patterns of adaptation in breast cancer patients with cancer-related fatigue using response shift analyses at subgroup level. *Cancer Medicine*, *6*(11), 2562–2575. <https://doi.org/10.1002/cam4.1219>
- Salmon, M., Blanchin, M., Rotonda, C., Guillemin, F., & Sébille, V. (2017b). Identifying patterns of adaptation in breast cancer patients with cancer-related fatigue using response shift analyses at subgroup level. *Cancer Medicine*, *6*(11), 2562–2575. <https://doi.org/10.1002/cam4.1219>
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph Supplement* *17*.
- Schmidt, H. (1982). The use of analysis of covariance structures to assess beta and gamma change. *Multivariate Behavioral Research*, *17*, 343–358.

- Schmidt, H., Nordhausen, T., Boese, S., Vordermark, D., Wheelwright, S., Wienke, A., & Johnson, C. D. (2018). Factors influencing global health related quality of life in elderly cancer patients: Results of a secondary data analysis. *Geriatrics (Switzerland)*, 3(1), 1–13. <https://doi.org/10.3390/geriatrics3010005>
- Schmitt, N. (1982). The Use Of Analysis Of Covariance Structures To Assess Beta And Gamma Change. *Multivariate Behavioral Research*, 17(3), 343–358. https://doi.org/10.1207/s15327906mbr1703_3
- Scholten, A. C., Haagsma, J. A., Steyerberg, E. W., van Beeck, E. F., & Polinder, S. (2017). Assessment of pre-injury health-related quality of life: a systematic review. *Population Health Metrics*, 15(1), 10. <https://doi.org/10.1186/s12963-017-0127-3>
- Schwartz, C. (2010). Applications of Response Shift Theory and Methods to Participation Measurement: A Brief History of a Young Field. *Archives of Physical Medicine and Rehabilitation*, 91(9), S38–S43. <https://doi.org/10.1016/j.apmr.2009.11.029>
- Schwartz, C. E., Powell, V. E., & Rapkin, B. D. (2017). When global rating of change contradicts observed change: Examining appraisal processes underlying paradoxical responses over time. *Quality of Life Research*, 26(4), 847–857. <https://doi.org/10.1007/s11136-016-1414-3>
- Schwartz, C. E., Quaranto, B. R., Rapkin, B. D., Healy, B. C., Vollmer, T., & Sprangers, M. A. G. (2014). Fluctuations in appraisal over time in the context of stable versus non-stable health. *Quality of Life Research*, 23(1), 9–19. <https://doi.org/10.1007/s11136-013-0471-0>
- Schwartz, C. E., & Rapkin, B. A. (2012). Understanding appraisal processes underlying the thentest: A mixed methods investigation. *Quality of Life Research*, 21(3), 381–388. <https://doi.org/10.1007/s11136-011-0023-4>
- Schwartz, C. E., Sajobi, T. T., Lix, L. M., Quaranto, B. R., & Finkelstein, J. A. (2013). Changing values, changing outcomes: The influence of reprioritization response shift on outcome assessment after spine surgery. *Quality of Life Research*, 22(9), 2255–2264. <https://doi.org/10.1007/s11136-013-0377-x>

- Schwartz, C. E., & Sprangers, M. A. . (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 48(11), 1531–1548. [https://doi.org/10.1016/S0277-9536\(99\)00047-7](https://doi.org/10.1016/S0277-9536(99)00047-7)
- Schwartz, C. E., Stucky, B. D., Michael, W., & Rapkin, B. D. (2020). Does response shift impact interpretation of change even among scales developed using item response theory? *Journal of Patient-Reported Outcomes*, 4(1). <https://doi.org/10.1186/s41687-019-0162-x>
- Schwartz, C. E., Stucky, B., Rivers, C. S., Noonan, V. K., & Finkelstein, J. A. (2018). Quality of Life and Adaptation in People With Spinal Cord Injury: Response Shift Effects From 1 to 5 Years Postinjury. *Archives of Physical Medicine and Rehabilitation*, 99(8), 1599-1608.e1. <https://doi.org/10.1016/j.apmr.2018.01.028>
- Schwartz, C., Sajobi, T., Verdam, M., Sebille, V., Lix, L., Guilleux, A., & Sprangers, M. (2015). Method variation in the impact of missing data on response shift detection. *Quality of Life Research*, 24(3), 521–528. <https://doi.org/10.1007/s11136-014-0746-0>
- Schwartz, C., & Sprangers, M. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 48(11), 1531–1548. [https://doi.org/10.1016/S0277-9536\(99\)00047-7](https://doi.org/10.1016/S0277-9536(99)00047-7)
- Schwartz, C., & Sprangers, M. (2000). Adaptation to changing health: Response shift in quality-of-life research. In C. E. Schwartz & M. A. G. Sprangers (Eds.), *Quality of Life Research*. American Psychological Association. <https://doi.org/10.1037/10382-000>
- Schwartz, Carolyn, Sprangers, G., Oort, F. J., Ahmed, S., Bode, R., Li, Y., & Vollmer, T. (2011). Response shift in patients with multiple sclerosis: An application of three statistical techniques. *Quality of Life Research*, 20(10), 1561–1572. <https://doi.org/10.1007/s11136-011-0056-8>
- Sébille, V., Hardouin, J. B., Giral, M., Bonnaud-Antignac, A., Tessier, P., Papuchon,

- E., Jobert, A., Faurel-Paul, E., Gentile, S., Cassuto, E., Morélon, E., Rostaing, L., Glotz, D., Sberro-Soussan, R., Foucher, Y., & Meurette, A. (2016). Prospective, multicenter, controlled study of quality of life, psychological adjustment process and medical outcomes of patients receiving a preemptive kidney transplant compared to a similar population of recipients after a dialysis period of less than t. *BMC Nephrology*, *17*(1), 1–9.
<https://doi.org/10.1186/s12882-016-0225-7>
- Shi, H. Y., Lee, K. T., Lee, H. H., Uen, Y. H., & Chiu, C. C. (2011). Response shift effect on gastrointestinal quality of life index after laparoscopic cholecystectomy. *Quality of Life Research*, *20*(3), 335–341.
<https://doi.org/10.1007/s11136-010-9760-z>
- Sprangers, M. A. ., & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research: a theoretical model. *Social Science & Medicine*, *48*(11), 1507–1515. [https://doi.org/10.1016/S0277-9536\(99\)00045-3](https://doi.org/10.1016/S0277-9536(99)00045-3)
- Sprangers, M. A. G., & Schwartz, C. E. (2010). Do not throw out the baby with the bath water: Build on current approaches to realize conceptual clarity. Response to Ubel, Peeters, and Smith. *Quality of Life Research*, *19*(4), 477–479.
<https://doi.org/10.1007/s11136-010-9611-y>
- Spuling, S. M., Wolff, J. K., & Wurm, S. (2017). Response shift in self-rated health after serious health events in old age. *Social Science and Medicine*, *192*, 85–93.
<https://doi.org/10.1016/j.socscimed.2017.09.026>
- Streiner, D., Norman, G., & Cairney, J. (1995). *Health Measurement Scales: A practical guide to their development and use*. Oxford University Press.
- Supo, J. (2012). Seminarios de Investigación Científica Sinopsis del libro 2012. *Seminarios de Investigación Científica*, 34. www.seminariodeinvestigacion.com
- Taminiau-Bloem, E. F., Schwartz, C. E., van Zuuren, F. J., Koeneman, M. A., Visser, M. R. M., Tishelman, C., Koning, C. C. E., & Sprangers, M. A. G. (2016). Using a retrospective pretest instead of a conventional pretest is replacing biases: a qualitative study of cognitive processes underlying responses to thentest items. *Quality of Life Research*, *25*(6), 1327–1337.

<https://doi.org/10.1007/s11136-015-1175-4>

- Taminiau-Bloem, E. F., van Zuuren, F. J., Koeneman, M. A., Rapkin, B. D., Visser, M. R. M., Koning, C. C. E., & Sprangers, M. A. G. (2010). A “short walk” is longer before radiotherapy than afterwards: A qualitative study questioning the baseline and follow-up design. *Health and Quality of Life Outcomes*, *8*, 1–12. <https://doi.org/10.1186/1477-7525-8-69>
- Tarczy-Hornoch, K., Ying-Lai, M., & Varma, R. (2006). Myopic Refractive Error in Adult Latinos: The Los Angeles Latino Eye Study. *Investigative Ophthalmology & Visual Science*, *47*(5), 1845. <https://doi.org/10.1167/iovs.05-1153>
- Taris, T. (2000). *A Primer in Longitudinal Data Analysis*. SAGE Publications Ltd. <https://doi.org/10.4135/9781849208512>
- ten Ham, R. M. T., Broering, J. M., Cooperberg, M. R., Carroll, P., & Wilson, L. S. (2020). Understanding the Major Factors Affecting Response Shift Effects on Health-Related Quality of Life: What the Then-Test Measures in a Longitudinal Prostate Cancer Registry. *Clinical Genitourinary Cancer*, *18*(1), e21–e27. <https://doi.org/10.1016/j.clgc.2019.10.026>
- ten Ham, R. M., Wilson, L. S., Broering, J. M., Cooperberg, M. R., & Carroll, P. (2014). Sustainable Measurement of Response Shift in Prostate Cancer Patients: Adjusting Health Related Quality of Life with the Then-Test. *Value in Health*, *17*(7), A651. <https://doi.org/10.1016/j.jval.2014.08.2366>
- Tessier, P., Blanchin, M., & Sébille, V. (2017). Does the relationship between health-related quality of life and subjective well-being change over time? An exploratory study among breast cancer patients. *Social Science and Medicine*, *174*, 96–103. <https://doi.org/10.1016/j.socscimed.2016.12.021>
- Testa, M. (2000a). Interpretation of quality of life outcomes: Issues that affect magnitude and meaning. *Medical Care*, *38*(9), 166–174. <https://pubmed.ncbi.nlm.nih.gov/10982103/>
- Testa, M. (2000b). Quality-of-Life Assessment in Diabetes Research: Interpreting the Magnitude and Meaning of Treatment Effects. *Diabetes Spectrum*, *13*(1),

29–35.

- Tew, M., Dalziel, K., Clarke, P., Smith, A., Choong, P. F., & Dowsey, M. (2020). Patient-reported outcome measures (PROMs): can they be used to guide patient-centered care and optimize outcomes in total knee replacement? *Quality of Life Research, 29*(12), 3273–3283. <https://doi.org/10.1007/s11136-020-02577-4>
- The Cochrane Collaboration. (2011). *Manual Cochrane de revisiones sistemáticas de intervenciones* (J. P. Higgins & S. Green (eds.)).
- Thomas, D. (1992). Interpreting discriminant functions: A data analytic approach. *Multivariate Behavioral Research, 27*, 323–333.
- Thomas, D., Zumbo, B., Zhu, P., & Dutta, S. (2008). On measuring the relative importance of explanatory variables in a logistic regression. *Journal of Modern Applied Statistical Methods, 7*, 21–38.
- Topp, J., Heesen, C., Augustin, M., Andrees, V., & Blome, C. (2020). Challenges and lessons learned from using anchoring vignettes to explore quality of life response behavior. *Quality of Life Research, 29*(8), 2149–2159. <https://doi.org/10.1007/s11136-020-02488-4>
- Traa, M. J., Braeken, J., De Vries, J., Roukema, J. A., Orsini, R. G., & Den Ouden, B. L. (2015). Evaluating quality of life and response shift from a couple-based perspective: a study among patients with colorectal cancer and their partners. *Quality of Life Research, 24*(6), 1431–1441. <https://doi.org/10.1007/s11136-014-0872-8>
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1–10.
- Tuesca-Molina, R. (2005). La Calidad de Vida, su importancia y cómo medirla. *Salud UNINORTE, 76–86*.
- Ubel, P. A., Peeters, Y., & Smith, D. (2010). Abandoning the language of “response shift”: a plea for conceptual clarity in distinguishing scale recalibration from true changes in quality of life. *Quality of Life Research, 19*(4), 465–471. <https://doi.org/10.1007/s11136-010-9592-x>

- Urzúa, A., & Caqueo-Urizar, A. (2012). Calidad de vida: Una revisión teórica del concepto Quality of life: A theoretical review. *Terapia Psicológica*, 30(1), 718–4808.
- Urzúa, Alfonso. (2010). Calidad de vida relacionada con la salud: Elementos conceptuales. *Revista Médica de Chile*, 138, 358–365.
<https://scielo.conicyt.cl/pdf/rmc/v138n3/art17.pdf>
- van der Ploeg, T., Peter, A., & Steyerberg, E. (2014). Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. *Medical Research Methodology*, 14(137), 1–13. <https://doi.org/1471-2288/14/137>
- Veenhoven, R. (2001). Are the Russians as unhappy as they say they are? Comparability of self-reports across nations. *Journal of Happiness Studies*, 2, 111–136.
- Veenhoven, R. (2006). The Four Qualities of Life: Ordering concepts and measures of the good life. *Published in Mark McGillivray & Mathew Clark (Eds) "Understanding Human Well-Being" United Nations University Press*, 74–100.
- Velarde-Jurado, E., & Avila-Figueroa, A. (2002). Consideraciones metodológicas para evaluar la calidad de vida. *Salud Pública de México*, 44(5), 448–463.
- Verdam, M. G. E., Oort, F. J., & Sprangers, M. A. G. (2016). Using structural equation modeling to detect response shifts and true change in discrete variables: an application to the items of the SF-36. *Quality of Life Research*, 25(6), 1361–1383. <https://doi.org/10.1007/s11136-015-1195-0>
- Verdam, M. G. E., Oort, F. J., & Sprangers, M. A. G. (2017a). Structural equation modeling–based effect-size indices were used to evaluate and interpret the impact of response shift effects. *Journal of Clinical Epidemiology*, 85, 37–44. <https://doi.org/10.1016/j.jclinepi.2017.02.012>
- Verdam, M. G. E., Oort, F. J., & Sprangers, M. A. G. (2017b). Structural equation modeling–based effect-size indices were used to evaluate and interpret the impact of response shift effects. *Journal of Clinical Epidemiology*, 85, 37–44.

<https://doi.org/10.1016/j.jclinepi.2017.02.012>

- Verdam, M. G. E., Oort, F. J., van der Linden, Y. M., & Sprangers, M. A. G. (2015). Taking into account the impact of attrition on the assessment of response shift and true change: a multigroup structural equation modeling approach. *Quality of Life Research*, 24(3), 541–551. <https://doi.org/10.1007/s11136-014-0829-y>
- Vicente-Galindo, M. (2003). *Contribuciones al análisis de datos de calidad de vida relacionada con la salud* [Universidad de Salamanca. España]. <https://dialnet.unirioja.es/servlet/tesis?codigo=220625>
- Vicente-Galindo, M., & Galindo-Villardón, M. (2009). El método STATIS como alternativa para detectar response shift en estudios de calidad de vida relacionada con la salud. *Revista de Matemáticas: Teoría y Aplicaciones*, 16(1), 1–15.
- Vicente-Villardón, J. (2010). *MULTBILOT: A package for Multivariate Analysis using Biplots*. <http://biplot.usal.es/ClassicalBiplot/index.html>
- Visser, M. R. M., Oort, F. J., Van Lanschot, J. J. B., Van Der Velden, J., Kloek, J. J., Gouma, D. J., Schwartz, C. E., & Sprangers, M. A. G. (2013). The role of recalibration response shift in explaining bodily pain in cancer patients undergoing invasive surgery: An empirical investigation of the Sprangers and Schwartz model. *Psycho-Oncology*, 22(3), 515–522. <https://doi.org/10.1002/pon.2114>
- Visser, M. R. M., Smets, E. M. A., Sprangers, M. A. G., & De Haes, H. J. C. J. M. (2000). How response shift may affect the measurement of change in fatigue. *Journal of Pain and Symptom Management*, 20(1), 12–18. [https://doi.org/10.1016/S0885-3924\(00\)00148-2](https://doi.org/10.1016/S0885-3924(00)00148-2)
- Von Blanckenburg, P., Seifart, U., Conrad, N., Exner, C., Rief, W., & Nestoriuc, Y. (2014). Quality of life in cancer rehabilitation: The role of life goal adjustment. *Psycho-Oncology*, 23(10), 1149–1156. <https://doi.org/10.1002/pon.3538>
- Wang, X., Xu, X., Han, H., He, R., Zhou, L., Liang, R., & Yu, H. (2018). Using structural equation modeling to detect response shift in quality of life in patients

- with Alzheimer's disease. *International Psychogeriatrics*, May, 1–10.
<https://doi.org/10.1017/S1041610218000595>
- Wardrop, J. L., Hayduk, L. A., & Bollen, K. A. (1993). Structural Equation Modeling with LISREL: Essentials and Advances. *Journal of Educational Statistics*, 18(1), 115. <https://doi.org/10.2307/1165186>
- WHOQOL, U. M. G. W. H. (1998). *World Health Organization*.
https://apps.who.int/iris/bitstream/handle/10665/77932/WHO_HIS_HSI_Rev.2?sequence=1
- Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, 116(2), 363–381. <https://doi.org/10.1037/0033-2909.116.2.363>
- Wilson, I., & Cleary, P. (1995). Linking clinical variables with related quality of life: A conceptual model of patients outcomes. *Journal of the American Medical Association*, 273(1), 50–65. <https://pubmed.ncbi.nlm.nih.gov/7996652/>
- Wright, B. (1997). The partial credit model. In *Handbook of modern item response theory* (pp. 101–121). Springer.
- Wrosch, C.; Scheier, M. (2003). Personality and quality of life: the importance of optimism and goal adjustment. *Quality of Life Research*, 12, 59–72.
<https://link.springer.com/article/10.1023/A:1023529606137>
- Wu, A. (2000). Quality of life assessment in clinical research: application in diverse populations. *Medical Care*, 38(9), 130–135.
<https://www.jstor.org/stable/3768071>
- Wu, P. (2016). Response Shifts in Depression Intervention for Early Adolescents. *Journal of Clinical Psychology*, 72(7), 663–675.
<https://doi.org/10.1002/jclp.22291>
- Yang, J., Hanna-Pladdy, B., Gruber-Baldini, A. L., Barr, E., von Coelln, R., Armstrong, M. J., Reich, S. G., & Shulman, L. M. (2017). Response shift – The experience of disease progression in Parkinson disease. *Parkinsonism and Related Disorders*, 36, 52–56. <https://doi.org/10.1016/j.parkreldis.2016.12.027>

Zee, B. C. (1998). Growth curve model analysis for quality of life data. *Statistics in Medicine*, 17(5–7), 757–766. [https://doi.org/10.1002/\(SICI\)1097-0258\(19980315/15\)17:5/7<757::AID-SIM819>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1097-0258(19980315/15)17:5/7<757::AID-SIM819>3.0.CO;2-N)

Zhang, X.-H., Li, S.-C., Xie, F., Lo, N.-N., Yang, K.-Y., Yeo, S.-J., Fong, K.-Y., & Thumboo, J. (2012). An Exploratory Study of Response Shift in Health-Related Quality of Life and Utility Assessment Among Patients with Osteoarthritis Undergoing Total Knee Replacement Surgery in a Tertiary Hospital in Singapore. *Value in Health*, 15(1), S72–S78. <https://doi.org/10.1016/j.jval.2011.11.011>

Zhang, X. H., Li, S. C., Xie, F., Lo, N. N., Yang, K. Y., Yeo, S. J., Fong, K. Y., & Thumboo, J. (2012). An exploratory study of response shift in health-related quality of life and utility assessment among patients with osteoarthritis undergoing total knee replacement surgery in a tertiary hospital in Singapore. *Value in Health*, 15(1 SUPPL.), 72–78. <https://doi.org/10.1016/j.jval.2011.11.011>

Zwinderman, A. (1990). The measurement of change of quality of life in clinical trials. *Statistics in Medicine*, 9, 931–942.

ANEXOS

ANEXO 1. Publicación de artículo

“Detection of response shift in health-related quality of life studies: a systematic review”

(2022) 20:20

<https://doi.org/10-1186/s12955-022-01926-w>

Journal: Health and Quality of Life Outcomes

JCR: 2.96

Q1 (2020)

RESEARCH

Open Access

Detection of response shift in health-related quality of life studies: a systematic review



Estelina Ortega-Gómez^{1,2*}, Purificación Vicente-Galindo^{2,3}, Helena Martín-Rodero⁴ and Purificación Galindo-Villardón^{2,5,6}

Abstract

Background: Response Shift (RS) refers to the idea that an individual may undergo changes in its health-related quality of life (HRQOL). If internal standard, values, or reconceptualization of HRQOL change over time, then answer to the same items by the same individuals may not be comparable over time. Traditional measures to evaluate RS is prone to bias and strong methodologies to study the existence of this phenomenon is required. The objective is to systematically identify, analyze, and synthesize the existing and recent evidence of statistical methods used for RS detection in HRQOL studies.

Methods: The analysis of selected studies between January 2010 and July 2020 was performed through a systematic review in MEDLINE/PubMed, Scopus, Web of Science, PsycINFO and Google Scholar databases. The search strategy used the terms "Health-Related Quality of Life" and "Response Shift" using the filters "Humans"; "Journal Article"; "English" and "2010/01/01–2020/07/31". The search was made in August 2020.

Results: After considering the inclusion and exclusion criteria, from the total selected articles (675), 107 (15.9%) of the publications were included in the analysis. From these, 79 (71.0%) detected the existence of RS and 86 (80.4%) only used one detection method. The most used methods were Then Test ($n = 41$) and Oort's Structural Equation Models (SEM) ($n = 35$). Other method used were Multiple Lineal Regression ($n = 7$), Mixed-Effect Regression ($n = 6$), Latent Trajectory Analysis ($n = 6$), Item Response Theory ($n = 6$), Logistics Regression ($n = 5$), Regression and Classification Trees ($n = 4$) and Relative Importance Method ($n = 4$). Most of these detected recalibration, including Then Test ($n = 27$), followed by Oort's SEM that detected the higher combination of RS types: recalibration ($n = 24$), reprioritization ($n = 13$) and reconceptualization ($n = 7$).

Conclusions: There is a continuous interest of studying RS detection. Oort's SEM becomes the most versatile method in its capability for detecting RS in all different types. Despite results from previous systematic reviews, same methods have been used during the last years. We observed the need to explore other alternative methods allowing same detection capacity with robust and highly precise methodology. The investigation on RS detection and types requires more study, therefore new opportunity grows to continue attending this phenomenon through a multidisciplinary perspective.

Keywords: Systematic review, Response shift, Multivariate analysis, Statistical methods

Background

Health-related quality of life (HRQOL) is the affectionation of a patient's physical and mental health, and their correlates, including health risk and conditions, independence level, social support, and socioeconomic status over time [1, 2].

*Correspondence: estelina.ortega@up.ac.pa; estelinao@usal.es
¹ Department of Statistics, University of Panama, Simon Bolívar Ave., Octavio Mendez Pereira Campus, Panama City, Republic of Panama
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

This concept is established as a facet of an individual's state of life while measuring its well-being as a medical patient. HRQOL is analyzed as a functional health state and identifies effective strategies to improve patient's conditions as a result of medical interventions [3, 4].

After a medical treatment, patients may perceive and demonstrate different conditions over time, from the procedure's initial stages up to many years after treatment has finished. The changes in the measurement of individual's perception or internal standard is known as Response Shift (RS) [5, 6]. From a clinical perspective, RS is generated as the change in the meaning of a subject self-administrated assessment [7] as a valid and sensitive mechanism to evaluate the change in different moments in time [8].

Spranger and Schwartz [9] theoretical model explains how RS may affect HRQOL as a result of health state changes. As a baseline model, it presents five components: (1) a catalysts, corresponding to an individual health states or its changes as a result or not of a treatment; (2) antecedents, referring to individual's characteristics influencing catalysts or appraisals mechanisms; (3) mechanism, explained by behavioral, cognitive, or affective processes accommodating changes in catalysts; (4) response shift, representing changes in the meaning of an individual self-evaluation of QOL resulting from changes in internal standards, values, or conceptualization; and (5) perceived QOL. Rapkin and Schwartz [10] propose that QOL appraisal processes must consider how individuals perceive their health status and respond questionnaires about their QOL. The model follows these processes: (1) induction of a frame of reference; (2) sampling based on the frame of reference; (3) judge against standards of comparison; and (4) combine algorithms to formulate a response. The proposal allows dynamic feedback to explain how QOL scores can be stabilized accounting inter-individual and temporary differences despite changes in health status [10].

An individual's self-assessment may demonstrate changes in three contexts: in the internal standards of the measurement scale (recalibration) indicating that the patient has a new scale for measuring its own state of HRQOL; in the scale of values (reprioritization) representing a change in the priority of elements that influence the context of life; and in the definition of the objective construct (reconceptualization) when a patient raises a redefinition of its own concept [9].

The change processes of a patient must be appropriately measured when the effects of a disease or medical treatments is evaluated [11, 12]. The interpretation of HRQOL data represents a challenge because patients self-report their health conditions at a specific time, which can also be influenced by psychological phenomena [13]. This

suggests that HRQOL measurements must consider that the individual reports on its status, at least in two or more moments to detect significant changes over time.

Two main approaches are proposed for the detection of change: methods based on specific study designs and secondary data analysis that includes statistical methods developed to test hypotheses that do not require specific designs [14].

The most commonly used methodology is the retrospective Then Test design [15] that allows evaluating the change of patient's internal standard by comparing the scores with two other moments: pre-test and post-test [5, 7]. However, this method is sensitive to bias and difficult to be used in longitudinal secondary data analysis [15].

Structural Equation Models (SEM) of Schmidt [16] and Oort [17], Item Response Theory (IRT) of Anota et al. [18], and Guilleaux et al. [19], Relative Importance Method of Lix et al. [20], Latent Trajectory Modelling of Ahmed et al. [21], and Classification and Regression Trees of Li & Schwartz [22] are among the most widely used and proven methods for RS detection in both primary and secondary data.

This bibliographic review has permitted to identify the methods traditionally applied in specific clinical studies related to HRQOL or to the analysis of previously elaborated databases, mainly by medical, academic or research institutions. Although other recent publications have carried out exhaustive systematic reviews to address this issue [15, 23], it is necessary to continue exploring if other alternative statistical methods are used as emerging mechanisms in the RS detection. Consequently, the purpose of this investigation is to systematically identify, analyze, and synthesize the existing evidence of new statistical methods used for RS detection in HRQOL studies.

Methods

The systematic review is a structured methodology allowing the identification and integration of different specific studies based on inclusion and exclusion criteria and facilitates the eligibility of relevant or interesting publications [24]. Some critical and determining elements to capture the largest number of eligible publications are the number and orientation of search repositories, and the established inclusion and exclusion criteria.

The PRISMA methodology [25] was applied to develop this systematic review. Through pre-specified eligibility criteria, it allows reducing bias in the identification, selection, synthesis, and summary of results from previously published studies [26]. The precision and reliability of this tool provides important benefits in health-related research [27].

The acceptance of this methodology has been used in recent studies on RS applied to different clinical areas: oncology [23, 28], orthopedic rehabilitation [29], pre-injury [30], as well as patient-reported outcomes (PRO) [31].

Data sources

The selected studies were carried out through an organized review of MEDLINE/PubMed, Scopus, Web of Science Core Collection (SSCI) PsycINFO and Google Scholar databases. The terms considered in MEDLINE/PubMed search are "Quality of Life", "Health-Related Quality of Life", "Response Shift" in descriptors or keywords in the title and/or abstract. The search equation was "Quality of Life"[Mesh] OR "Quality of Life"[Title/Abstract] OR "Life Quality"[Title/Abstract] OR "Health-Related Quality of Life"[Title/Abstract] OR "Health Related Quality of Life"[Title/Abstract] OR "QoL" [Title/Abstract] OR "HRQoL"[Title/Abstract]) AND "Response Shift"[Title/Abstract].

The following filters were used: "Humans", "Journal Articles", "English", "2010/01/01–2020/07/31". This search strategy was adopted for each of the databases consulted. The search was made in August 2020 and completed with an analysis of the selected literature between January 2010 and July 2020.

Articles selection

Articles that qualified for the eligibility review were those in English language that met the following criteria: adapt to the objectives of the search to identify and synthesis of statistical methods used in RS detection in HRQOL studies; be published in peer-reviewed journals and be able to retrieve the full text of the work; and the term "Response Shift" was included in the title, abstract and/or keyword; those studies in a language other than English that did not include the use of statistical methods for the RS detection, as well as studies whose main objective was not the detection of SR and its classification. Conferences, editorial notes, systematic reviews, and conceptual evaluations were excluded from the study.

Article selection was independently made by two authors (EOJ and PVG) who initially review titles, then abstracts, and read the full texts. For article inclusion, a concordance assessment between authors was established to be greater than 80% (Kappa index). In case of discrepancies, the process was repeated until they were resolved by consensus among all the authors. Figure 1 illustrates the literature search and publication selected through the PRISMA flow diagram.

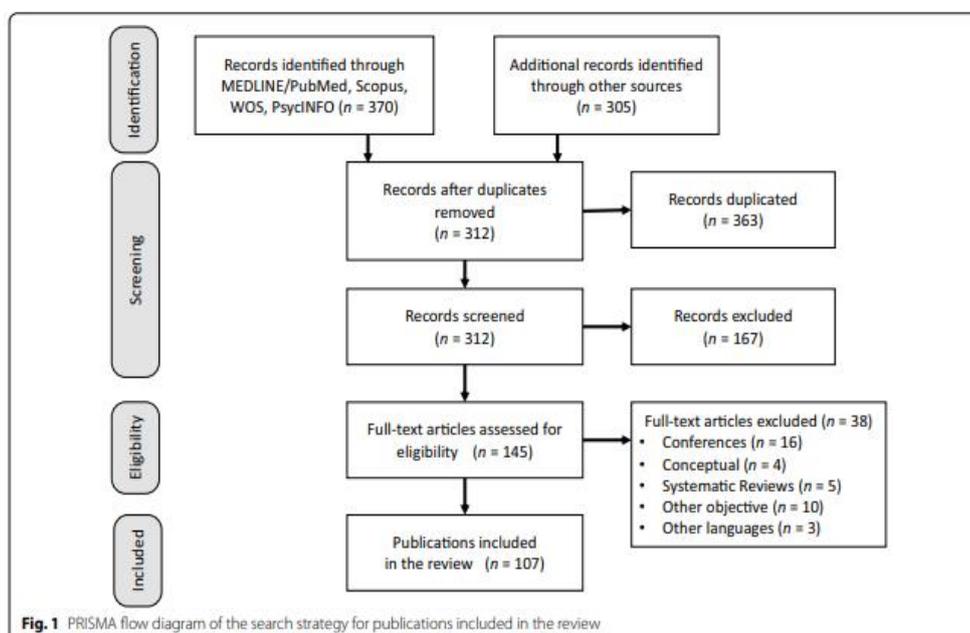




Fig. 2 Number of publications by year

Table 1 Instruments mostly used in analyzed studies

Instrument used	Code	Number
European Organisation for Research and Treatment of Cancer Quality-of-life-Questionnaire Core 30	EORTC QLQ-C30	20
36-Item Short Form Survey	SF-36	22
12-Item Short Form Survey	SF-12	7
EuroQol – 5 Dimensions	EQ-5D	10
Oral Health Impact Profile	OHIP	6

Results

Descriptive characteristics of the study

The search strategy identified a total of 675 articles in the databases of MEDLINE/PubMed (141), Scopus (46), Web of Science Core Collection (114), PsycINFO (69) y Google Scholar (305). After eliminating 363 duplicated records, 167 publications without the term “response shift” in the fields title, abstract and/or keyword, and other 38 articles excluded for not meeting the inclusion criteria, a total $n = 107$ (15.9%) were included in the analysis (See Additional file 1).

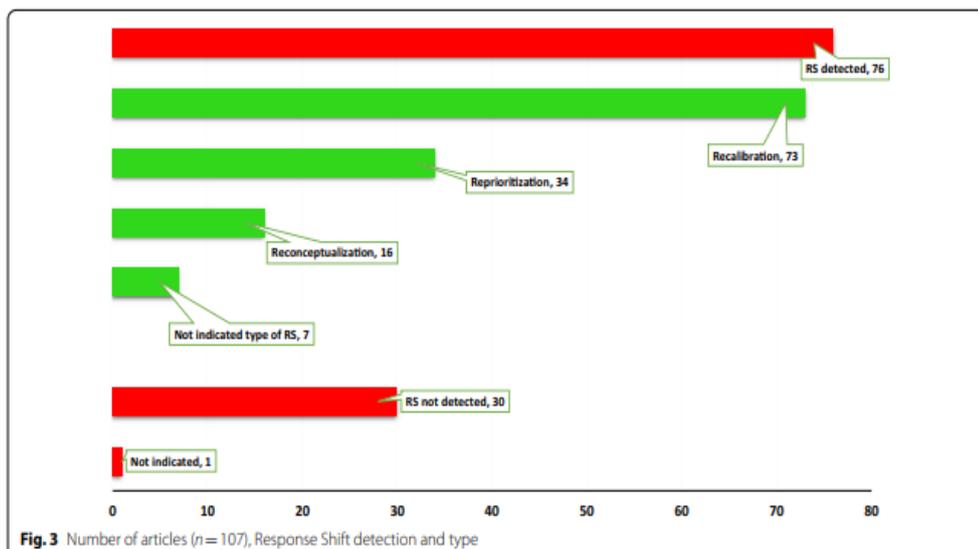


Fig. 3 Number of articles ($n = 107$), Response Shift detection and type

Based in the year of publication, 15 studies were recorded in 2017, followed by 2016 (13) and 2014 (12). An active interest in this research topic was continued during the rest of the years studied (see Fig. 2). A total of 56 journals were identified in this systematic review, where 31.8% of the articles were published in *Quality of Life Research* (34), followed by *Health and Quality of Life Outcomes* (10), *Journal of Clinical Epidemiology* (4) and *European Journal of Cancer Care* (4).

The analysis of the included articles determined that 69.2% of the studies were based on primary data, 29.0% on secondary data, and 1.8% did not specify. Different questionnaires were used to capture patient's information. Some studies concentrated in the use of one instrument, while other indicated diverse mechanisms to collect data (see Table 1). The medical orientations of the publications were mainly in *Oncology* (27), *Neurology* (11), *Psychology/Psychiatrics* (10), *Orthopedics* (9), *Oral health* (7) and *Cardiology* (6).

Table 2 Frequency of methods used for Response Shift detection

Method	Total (n = 129)	Response shift detection		
		Yes	No	Not indicated
Then-test	41	32	9	0
Structural Equation Models (Oort)	35	26	9	0
Structural Equation Models (Schmidt)	2	2	0	0
Multiple Linear Regression	7	4	2	1
Relative Importance Method	4	3	1	0
Mixed-Effects Regression	6	5	1	0
Classification and Regression Trees	4	4	0	0
Random Forest Regression	2	2	0	0
Logistics Regression Model	5	1	4	0
Item Response Theory	6	3	3	0
Latent Trajectory Analysis	6	2	4	0
Other methods	11	7	3	1

About the methods for Response Shift detection

Of the 107 articles analyzed, 76 (71.0%) detected the existence of RS, 30 (28.0%) did not identify it, while one article did not indicate it (see Fig. 3). The review described the methods for RS detection, types of RS, and if studies detected or not its existence. A group of 86 (80.4%) articles used one method for RS detection, 19 (17.8%) articles used two or more methods, and two articles did not specify the method used.

Table 2 describes the frequency of methods used in these publications. 41 (31.8%) articles used the Then-Test method, 35 (27.1%) adopted the Oort's SEM and 2 (1.6%) articles indicated the use of Schmidt's SEM. Other models were Multiple Linear Regression (7), Mixed-Effects Regression (6), Latent Trajectory Analysis (6), Item Response Theory (6), Logistic Regression (5), Classification and Regression Tree (4) and Relative Importance Method (4). Other were used but in a smaller number of studies.

According to the detection methods used in all articles, 91 (70.5%) detected RS, while 36 (27.9%) indicated

Table 3 Number of articles according to the detection method and type of Response Shift

Method	Type of Response Shift			
	Recalibration	Reprioritization	Reconceptualization	Not indicated
Then-Test	27	6	3	3
Structural Equation Models (Oort)	24	13	7	1
Structural Equation Models (Schmidt)	2	0	0	0
Multiple Linear Regression	2	0	1	1
Relative Importance Method	0	3	0	0
Mixed-Effects Regression	3	2	0	1
Classification and Regression Trees	4	2	1	0
Random Forest Regression	1	1	0	0
Logistics Regression Model	1	1	1	0
Item Response Theory	3	3	1	0
Latent Trajectory Analysis	2	1	0	0
Other methods	4	2	2	1

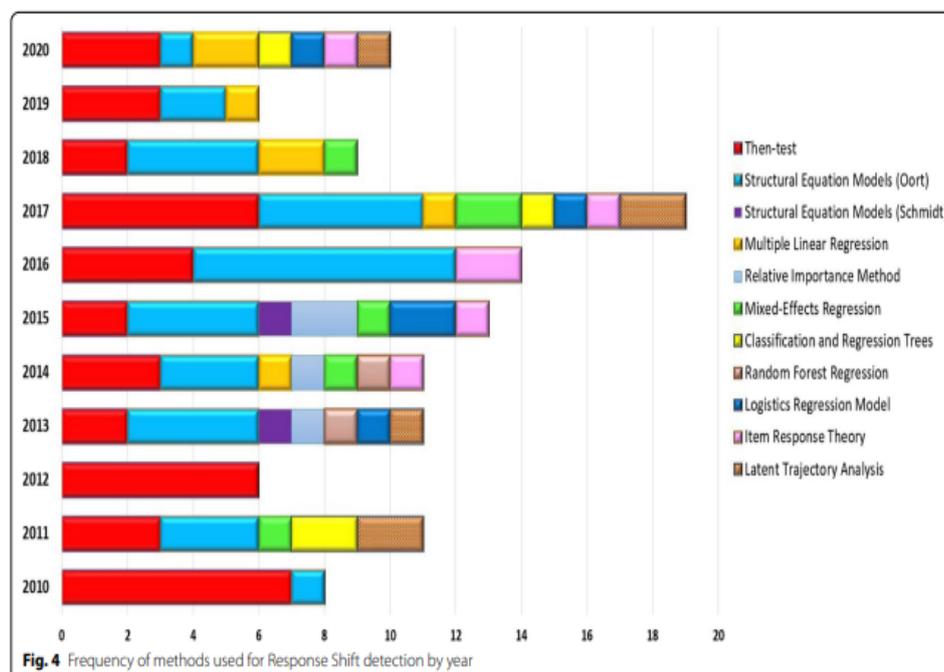


Fig. 4 Frequency of methods used for Response Shift detection by year

its absence. In the review, 2 (1.6%) articles did not identify the presence or absence of this phenomenon. Of the studies that detected RS through different methods, 73 indicated the presence of recalibration, 34 indicated the existence of reprioritization, 16 highlighted reconceptualization, and 7 articles detected the existence of RS, but did not specify the type (see Table 3). From these, 41 articles detected one type of RS, 20 articles registered two types of RS, only 7 studies identified the three types simultaneously. Seven articles did not indicate it.

Of the articles included ($n=107$), the traditional method for RS detection Then Test registered the highest number of RS recalibration type (27), while Oort's SEM registered a lesser number of studies detecting recalibration (24), but the highest record of reprioritization (13) and reconceptualization (7) of the entire systematic review (see Table 3). Schmidt's SEM detected changes in internal standards only in two (2) articles. The rest of the methods also proved to be used for the detection of the three types of RS, except for the Relative Importance Methods, Mixed-Effects Regression, Random Forest Regression, and Latent Trajectory Analysis, which did not identify the presence of reconceptualization.

Despite the effectiveness for detecting RS, this review evidences the growing interest for exploring different

statistical methods for RS studies (see Fig. 4). The most frequently used methods were Then Test and Oort's SEM between 2010 and 2020. However, during the last years, the figure shows that other alternative techniques are being considered for the assessment of this phenomenon.

Discussion

After the systematic review, the study examined the advances on HRQOL research through different methods for the detection of RS. It describes how the detection of this phenomenon is being evaluated in recent years, which methods are used the most, and the most identified type of RS. For physicians and researchers, systematic reviews are useful sources of evidence and scientific advances [32], they provide elements so clinical policymakers can evaluate risks, benefits and effects on health care, as well as new research initiatives [26, 27].

The majority of HRQOL studies and RS detection methods were published during 2016 and 2017. One third of the articles included in the review were studying cancer disease [33–41]. Then Test was the most used method mainly for primary data measuring recalibration [42–48]. Several studies remarked that Then Test may also identify other RS types: reprioritization y reconceptualization [33, 49, 50]. Similar results were presented in Sajobi et al. [15] study corresponding to a systematic review about RS

detecting methods were 54.5% of the articles used Then Test for this purpose. As same, in Ilie et al. [23] a 60.0% of the cancer studies remarked using this methodology, but authors indicate that interpretation must be cautious as it is bias prone.

Then Test has the advantage for easy handling and analysis, but its disadvantage relies in random errors and/or confound and difficult for interpretation. Therefore, this appraisal suggests the utilization of individualized methods with strong statistical rigor [51].

Oort's SEM [17] differs from Then Test, since it is used for studies based on primary and secondary data. This multivariate method evidences the capability for detecting changes in all type of RS: recalibration, reprioritization, and reconceptualization. Several studies reiterate Oort's SEM as an effective method to detect changes in patients, despite the different type of RS or data used [52–54]. This detection capacity might be the reason for a larger interest for researchers of using this method.

Schmidt's model was used in two studies detecting only recalibration. In comparative studies with Oort's method [55, 56], only this type of RS is identified, but conclusions indicated that both approaches use different parameters to identify recalibration: Schmidt defines recalibration as the change in factor variances or factor loadings over time, while Oort by the change in intercepts.

Other methods obtained more records for RS detection: Mixed-Effects and Multiple Linear Regression [57–60], Classification and Regression Trees [22, 61], Item Response Theory [18, 62] and Relative Importance Methods [63–65].

Despite a major frequency of methods to detect recalibration, it calls the attention that all three studies using Relative Importance Method detected reprioritization [63–65]. Mixed-Effect Regression method [57, 58], Classification and Regression Trees [22, 61] and Latent Trajectory Analysis [66] neither provided information on reconceptualization.

During this study we observed that the sample size is related to the type of method used for RS detection. Then Test is more flexible and functional in small samples or in specific studies, while methods based on stronger statistic techniques require larger samples. According to Schwartz et al. [67], for using advanced multivariate methods capable to detect RS, data analysis must include sufficient individuals participating in the study, as well as considering certain number of parameters such a clear the model, loading factor, relationship between the items, data distribution, and processes for estimating these parameters.

The need of exploring new multivariate methodologies for analyzing this phenomenon brings an alternative

proposal for studying HRQOL and RS detection through three-way data: dual STATIS Method [68, 69]. The proposal presented by Vicente Galindo [70] centers in a procedure integrating the dual STATIS methods and the comparison of Krzanowski [71] subspaces, and examines factorial structures from multiple data sets focused to identify the existence of change.

The systematic reviews evaluated in this study [15, 23, 28–31] confirm the continuous use of the same methods for detecting RS and its different types, their advantages, and limitations. The previous proposal is an example of existing opportunities to continue examining other strong statistic methodologies allowing to deepen this line of investigation and reducing bias and analysis ambiguities.

Conclusion

The systematic review has demonstrated to become an adequate and convenient methodology to identify and synthesize advances of specific topics. Results have evidenced a generalized interest for studying HRQOL and RS for different diagnostic groups. RS detection continue attracting researcher's attention that have consolidated a set of methods for its analysis: Then Test, Oort's SEM, Multiple Linear and Mixed-Effects Regressions, Item Response Theory and Latent Trajectory Analysis. Oort's SEM becomes the most versatile method in its capability for detecting RS in all different types.

Results demonstrated that not all the methods achieve RS detection in similar proportions, mostly capable to identify recalibration, in some cases reprioritization, and in few studies reconceptualization. At the same time, previous systematic reviews and the result from this updating research conclude that same methods have been used during the last years and there is no evidence found of alternative statistical techniques proposed for detecting RS. Perhaps, our study states a need for exploring other methods with similar detection capacity, with robust and highly precise methodology either graphical oriented or with simpler methodologies.

Previous systematic reviews and the need to continue investigating about this phenomenon have motivated the update of this search within the last decade (2010–2020). Although these reviews have used a larger number of search terms including different RS types as well as data measurement characteristics, this systematic review has intentionally generalized the terms using RS and HRQOL to focus the results on those studies centered on this topic.

Since RS can be evaluated from different perspectives, and other disciplines such as statistics, psychology, humanities, and other related social sciences can provide significant contributions, it is recommended to continue

studying the RS phenomenon, its detection, and types from a multidisciplinary approach.

Abbreviations

RS: Response Shift; HRQOL: Health-Related Quality of Life; SEM: Structural Equation Model; PRO: Patient-reported outcomes.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12955-022-01926-w>.

Additional file 1: List of articles included in the systematic review: 2010-2020.

Acknowledgements

Not applicable.

Authors' contributions

EO, PV and PG participated in the concept and design; EO and HM participated in the acquisition of data, analysis, interpretation of databases, and drafting of the manuscript; all authors read, revised it critically, and approved the final manuscript. All authors read and approved the final manuscript.

Funding

The authors declare that they have not received special funds, grants or financial support for this research nor publication.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interest.

Author details

¹Department of Statistics, University of Panama, Simon Bolivar Ave., Octavio Mendez Pereira Campus, Panama City, Republic of Panama. ²Department of Statistics, University of Salamanca, 37008 Salamanca, Spain. ³Department of Statistics, IBSAL, University of Salamanca, Salamanca, Spain. ⁴Department of Library and Information, University of Salamanca, Alfonso X El Sabio St. 37007, Miguel de Unamuno Campus, Salamanca, Spain. ⁵Centro de Investigación Institucional (CII), Universidad Bernardo O'Higgins, Avenida Viel 1497, Santiago, Chile. ⁶Universidad Estatal de Milagro, Guayas, Ecuador.

Received: 17 February 2021 Accepted: 25 January 2022

Published online: 05 February 2022

References

- World Health Organization [WHO]. WHOQOL: measuring quality of life. <https://apps.who.int/iris/handle/10665/63482> (1997). Accessed 20 Aug 2020.
- Centers for Disease Control and Prevention [CDC]. Measuring healthy days: population assessment of health-related quality of life. <https://www.cdc.gov/hrqol/pdfs/mhd.pdf> (2000). Accessed 20 Aug 2020.
- Kang K, Gholizadeh L, Inglis SC, Han HR. Interventions that improve health-related quality of life in patients with myocardial infarction. *Qual Life Res.* 2016;25(11):2725–37. <https://doi.org/10.1007/s11136-016-1401-8>.
- Shiroywa T, Fukuda T, Shimozuma K, et al. Long-term health status as measured by EQ-5D among patients with metastatic breast cancer: comparison of first-line oral S-1 and taxane therapies in the randomized phase III SELECT BC trial. *Qual Life Res.* 2017;26(2):445–53. <https://doi.org/10.1007/s11136-016-1388-1>.
- Howard G, Dailey P. Response shift bias: a source of contamination of self-reporting measures. *J Appl Psychol.* 1979;64(2):144–50. <https://doi.org/10.1037/0021-9010.64.2.144>.
- Golembiewski RT, Billingsley K, Yeager S. Measuring change and persistence in human affairs: types of changes generated by OD design. *J Appl Behav Sci.* 1976;12(2):133–57. <https://doi.org/10.1177/002188637601200201>.
- Schwartz CE, Sprangers MA. Methodological approaches for assessing response shift in longitudinal quality of life research. *Soc Sci Med.* 1999;48(11):1531–48. [https://doi.org/10.1016/s0277-9536\(99\)00047-7](https://doi.org/10.1016/s0277-9536(99)00047-7).
- Hamidou Z, Dabakuyo-Yonio T, Bonnetain F. Impact of response shift on longitudinal quality of life assessment in cancer clinical trials. *Expert Rev Pharmacoecon Outcomes Res.* 2011;11(5):549–59. <https://doi.org/10.1586/erp.11.5.7>.
- Sprangers MA, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model. *Soc Sci Med.* 1999;48(11):1507–15. [https://doi.org/10.1016/s0277-9536\(99\)00045-3](https://doi.org/10.1016/s0277-9536(99)00045-3).
- Rapkin BD, Schwartz CE. Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift. *Health Qual Life Outcomes.* 2004;2(14):66. <https://doi.org/10.1186/1477-7525-2-14>.
- McPhail S, Haines T. Response shift, recall bias and their effect on measuring change in health-related quality of life amongst older hospital patients. *Health Qual Life Outcomes.* 2010;8:65. <https://doi.org/10.1186/1477-7525-8-65>.
- Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol.* 2003;56(5):395–407. [https://doi.org/10.1016/s0895-4356\(03\)00044-1](https://doi.org/10.1016/s0895-4356(03)00044-1).
- Kvam AK, Wisloff F, Fayers PM. Minimal important differences and response shift in health-related quality of life; a longitudinal study in patients with multiple myeloma. *Health Qual Life Outcomes.* 2010;8:79. <https://doi.org/10.1186/1477-7525-8-79>.
- Bulteau S, Sauvaet A, Vanier A, et al. Depression reappraisal and treatment effect: will response shift help improve the estimation of treatment efficacy in trials for mood disorders? *Front Psychiatry.* 2019;10(420):1–3. <https://doi.org/10.3389/fpsy.2019.00420>.
- Sajobi TT, Brahmabatt R, Lix L, Zumbo B, Sawatzky R. Scoping review of response shift methods: current reporting practices and recommendations. *Qual Life Res.* 2018;27(5):1133–46. <https://doi.org/10.1007/s11136-017-1751-x>.
- Schmidt N. The use of analysis of covariance structures to assess beta and gamma change. *Multivar Behav Res.* 1982;17(3):343–58. https://doi.org/10.1207/s15327906mbr1703_3.
- Oort FJ. Using structural equation modeling to detect response shifts and true change. *Qual Life Res.* 2005;14(3):587–98. <https://doi.org/10.1007/s11136-004-0830-y>.
- Anota A, Bascoul-Mollevi C, Conroy T, et al. Item response theory and factor analysis as a mean to characterize occurrence of response shift in a longitudinal quality of life study in breast cancer patients. *Health Life Outcomes.* 2014;12(32):2–18. <https://doi.org/10.1186/1477-7525-12-32>.
- Guilleux A, Blanchin M, Vanier A, et al. RespOnse Shift Algorithm in Item response theory (ROSALI) for response shift detection with missing data in longitudinal patient-reported outcome studies. *Qual Life Res.* 2015;24(3):553–64. <https://doi.org/10.1007/s11136-014-0876-4>.
- Lix LM, Sajobi TT, Sawatzky R, et al. Relative importance measures for reprioritization response shift. *Qual Life Res.* 2013;22(4):695–703. <https://doi.org/10.1007/s11136-012-0198-3>.
- Ahmed S, Mayo N, Scott S, Kuspinar A, Schwartz C. Using latent trajectory analysis of residuals to detect response shift in general health among patients with multiple sclerosis article. *Qual Life Res.* 2011;20:1555–60. <https://doi.org/10.1007/s11136-011-0005-6>.

22. Li Y, Schwartz C. Data mining for response shift patterns in multiple sclerosis patients using recursive partitioning tree analysis. *Qual Life Res*. 2011;20:1543–53. <https://doi.org/10.1007/s11136-011-0004-7>.
23. Ilie G, Bradfield J, Moodie L, et al. The role of response-shift in studies assessing quality of life outcomes among cancer patients: a systematic review. *Front Oncol*. 2019;9:783. <https://doi.org/10.3389/fonc.2019.00783>.
24. Montori VM, Swiontkowski MF, Cook DJ. Methodologic issues in systematic reviews and meta-analyses. *Clin Orthop Relat Res*. 2003;413:43–54. <https://doi.org/10.1097/01.blo.0000079322.41006.5b>.
25. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4:1. <https://doi.org/10.1186/2046-4053-4-1>.
26. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. <https://doi.org/10.1371/journal.pmed.1000097>.
27. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700. <https://doi.org/10.1136/bmj.b2700>.
28. Nielsen LK, Jarden M, Andersen CL, Frederiksen H, Abildgaard NA. Systematic review of health-related quality of life in longitudinal studies of myeloma patients. *Eur J Haematol*. 2017;99:3–17. <https://doi.org/10.1111/ejh.12882>.
29. Powden C, Hoch M, Hoch J. Examination of response shift after rehabilitation for orthopedic conditions: a systematic review. *J Sport Rehabil*. 2018;27:469–79. <https://doi.org/10.1123/jsr.2017-0104>.
30. Scholten AC, Haagsma JA, Steyerberg EW, vanBeek EF, Polinder S. Assessment of pre-injury health-related quality of life: a systematic review. *Popul Health Metr*. 2017;15:10. <https://doi.org/10.1186/s12963-017-0127-3>.
31. Hinds AM, Sajobi TT, Sebille V, Sawatzky R, Lix LM. A systematic review of the quality of reporting of simulation studies about methods for the analysis of complex longitudinal patient-reported outcomes data. *Qual Life Res*. 2018;27(10):2507–16. <https://doi.org/10.1007/s11136-018-1861-0>.
32. Manchikanti L, Datta S, Smith HS, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: part 6. Systematic reviews and meta-analyses of observational studies. *Pain Phys*. 2009;12(5):819–50.
33. Taminiau-Bloem E, van Zeuren F, Koeneman M, et al. A 'short walk' is longer before radiotherapy than afterwards: a qualitative study questioning the baseline and follow-up design. *Health Qual Life Outcomes*. 2010;8:69. <https://doi.org/10.1186/1477-7525-8-69>.
34. Neuman H, Park J, Fuzesi S, Temple L. Rectal cancer patients' quality of life with a temporary stoma: shifting perspectives. *Dis Colon Rectum*. 2010;55(11):1117–24. <https://doi.org/10.1097/DCR.0b013e3182685213>.
35. Dabakuyo TS, Guillemin T, Conroy T, et al. Response shift effects on measuring post-operative quality of life among breast cancer patients: a multicenter cohort study. *Qual Life Res*. 2013;22:1–11. <https://doi.org/10.1007/s11136-012-0135-5>.
36. Brinkma A, Tissing WJ, Sulkers E, Kamps WA, Roodbol PF, Sanderman R. Exploring the response shift phenomenon in childhood patients with cancer and its effect on health-related quality of life. *Oncol Nurs Forum*. 2014;41(1):48–56. <https://doi.org/10.1188/14.ONF.41-01AP>.
37. Gerlich C, Schuler M, Jelitte M, et al. Prostate cancer patients' quality of life assessments across the primary treatment trajectory: 'True' change or response shift? *Acta Oncol*. 2016;55(7):814–20. <https://doi.org/10.3109/0284186X.2015.1136749>.
38. Tessier P, Blanchin M, Sébile V. Does the relationship between health-related quality of life and subjective well-being change over time? An exploratory study among breast cancer patients. *Soc Sci Med*. 2017;174:96–103. <https://doi.org/10.1016/j.socscimed.2016.12.021>.
39. Aburub A, Gagnon B, Ahmed S, Rodriguez AM, Mayo N. Impact of reconceptualization response shift on rating of quality of life over time among people with advanced cancer. *Support Care Cancer*. 2018;26(9):3063–71. <https://doi.org/10.1007/s00520-018-4156-7>.
40. Preiß M, Friedrich M, Stolzenburg JU, Zenger M, Hinz A. Response shift effects in the assessment of urologic cancer patients' quality of life. *Eur J Cancer Care*. 2019;28(4):1–8. <https://doi.org/10.1111/ecc.13027>.
41. Murata T, Suzukamo Y, Shirowa T, et al. Response Shift-adjusted treatment effect on health-related quality of life in a randomized controlled trial of Taxane versus S-1 for metastatic breast cancer: structural equation modeling. *Value Health*. 2020;23(6):768–74. <https://doi.org/10.1016/j.jval.2020.02.003>.
42. Dempster M, Carney R, McClements R. Response shift in the assessment of quality of life among people attending cardiac rehabilitation. *Br J Health Psychol*. 2010;15:307–19. <https://doi.org/10.1348/135910709X464443>.
43. Shi HY, Lee KT, Lee HH. Response shift effect on gastrointestinal quality of life index after laparoscopic cholecystectomy. *Qual Life Res*. 2011;20:335–41. <https://doi.org/10.1007/s11136-010-9760-z>.
44. Galenkamp H, Huisman M, Braam A, Deeg D. Estimates of prospective change in self-rated health in older people were biased owing to potential recalibration response shift. *J Clin Epidemiol*. 2012;65:978–88. <https://doi.org/10.1016/j.jclinepi.2012.03.010>.
45. Rutgers M, Creemers L, Yang KG, Rajmakers N, Dhert W, Saris D. Osteoarthritis treatment using autologous conditioned serum after placebo. Patient considerations and clinical response in a non-randomized case series. *Acta Orthop*. 2015;86(1):114–8. <https://doi.org/10.3109/174553674.2014.950467>.
46. Arthur J, Watts T, Davies R, Manchiaiah V, Slater J. An exploratory study identifying a possible response shift phenomenon of the Glasgow hearing aid benefit profile. *Audiol Res*. 2016;6(152):44–8. <https://doi.org/10.4081/audiore.2016.152>.
47. Reissmann D, Erler A, Hirsch C, Sienwald I, Machuca C, Schierz O. Bias in retrospective assessment of perceived dental treatment effects when using the Oral Health Impact Profile. *Qual Life Res*. 2018;27:775–82. <https://doi.org/10.1007/s11136-017-1725-z>.
48. Haagsma J, Spronk I, de Jongh M, Bonsel G, Polinder S. Conventional and retrospective change in health-related quality of life of trauma patients: an explorative observational follow-up study. *Health Qual Life Outcomes*. 2020;18:157. <https://doi.org/10.1186/s12955-020-01404-1>.
49. Höfer S, Pfaffenberger N, Renn D, Platter M, Ring L. Coronary intervention improves disease specific health-related quality of life but not individualised quality of life: a potential response shift effect? *Applied Res Qual Life*. 2011;6:81–90. <https://doi.org/10.1007/s11482-010-9117-3>.
50. Machuca C, Vettore M, Krasuska M, Baker S, Robinson P. Using classification and regression tree modelling to investigate response shift patterns in dentine hypersensitivity. *BMC Med Res Methodol*. 2017;17:120. <https://doi.org/10.1186/s12874-017-0396-3>.
51. Schwartz CE. Applications of response shift theory and methods to participation measurement: a brief history of a young field. *Arch Phys Med Rehabil*. 2010;91(9):S38–43. <https://doi.org/10.1016/j.apmr.2009.11.029>.
52. DeConde A, Bodner T, Mace J, Smith T. Response shift in quality of life after endoscopic sinus surgery for chronic rhinosinusitis. *JAMA Otolaryngol Head Neck Surg*. 2014;140(8):712–9. <https://doi.org/10.1001/jamaoto.2014.1045>.
53. Verdam MG, Oort FJ, Sprangers MAG. Structural equation modeling-based effect-size indices were used to evaluate and interpret the impact of response shift effect. *J Clin Epidemiol*. 2017;66:8537–44. <https://doi.org/10.1016/j.jclinepi.2017.02.012>.
54. Wang X, Xu X, Han H, et al. Using structural equation modeling to detect response shift in quality of life in patients with Alzheimer's disease. *Int Psychogeriatr*. 2018;31(1):123–32. <https://doi.org/10.1017/S1041610218000595>.
55. Gandhi PK, Ried LD, Huang IC, Kimberlin C, Kauf T, Suh DC. Identification of Response Shift among hypertensive patients with coronary artery disease using two structural equation modeling techniques. *Value Health*. 2011;14(3):A52–3. <https://doi.org/10.1016/j.jval.2011.02.298>.
56. Gandhi PK, Ried LD, Huang IC, Kimberlin C, Kauf T. Assessment of response shift using two structural equation modeling techniques. *Qual Life Res*. 2013;22:461–71. <https://doi.org/10.1007/s11136-012-0171-1>.
57. Lacey H, Loewenstein G, Ubel P. Compared to what? A joint evaluation method for assessing quality of life. *Qual Life Res*. 2011;20:1169–77. <https://doi.org/10.1007/s11136-011-9856-0>.
58. Schwartz C, Quaranto B, Rankin B, Healy B, Vollmer T, Sprangers M. Fluctuations in appraisal over time in the context of stable versus non-stable health. *Qual Life Res*. 2014;23:9–19. <https://doi.org/10.1007/s11136-013-0471-0>.
59. Nichols G, Antoun J, Fowler P, Al-Ani A, Farella M. Long-term changes in oral health-related quality of life of standard, cleft, and surgery patients after orthodontic treatment: a longitudinal study. *Am J Orthod Dentofac Orthop*. 2018;153(2):224–31. <https://doi.org/10.1016/j.ajodo.2017.06.018>.

60. Felix J, Becker C, Vogl M, Buschner P, Plötz W, Leidi R. Patient characteristics and valuation changes impact quality of life and satisfaction in total knee arthroplasty – results from a German prospective cohort study. *Health Qual Life Outcomes*. 2019;17:180. <https://doi.org/10.1186/s12955-019-1237-3>.
61. Machuca C, Vettore M, Robinson P. How peoples' ratings of dental implant treatment change over time? *Qual Life Res*. 2020;29:1322–34. <https://doi.org/10.1007/s11136-019-02408-1>.
62. Blanchin M, Sébille V, Guilleux A, Hardouin J. The Guttman errors as a tool for response shift detection at subgroup and item levels. *Qual Life Res*. 2016;25:1385–93. <https://doi.org/10.1007/s11136-016-1268-8>.
63. Schwartz C, Sajobi T, Lix L, Quaranto B, Finkelstein J. Changing values, changing outcomes: the influence of reprioritization response shift on outcome assessment after spine surgery. *Qual Life Res*. 2013;22:2255–64. <https://doi.org/10.1007/s11136-013-0377-x>.
64. Sajobi TT, Fiest KM, Wiebe S. Changes in quality of life after epilepsy surgery: The role of reprioritization response shift. *Epilepsia*. 2014;55(9):1331–8. <https://doi.org/10.1111/epi.12697>.
65. Sajobi TT, Lix L, Singh G, Lowerison M, Engbers J, Mayo N. Identifying reprioritization response shift in a stroke caregiver population: a comparison of missing data methods. *Qual Life Res*. 2015;24:529–40. <https://doi.org/10.1007/s11136-014-0824-3>.
66. Salmon M, Blanchin M, Rotonda C, Guillemain F, Sebille V. Identifying patterns of adaptation in breast cancer patients with cancer-related fatigue using response shift analyses at subgroup level. *Cancer Med*. 2017;6(11):2562–75. <https://doi.org/10.1002/cam4.1219>.
67. Schwartz C, Ahmed S, Sawatzky R, et al. Guidelines for secondary analysis in search of response shift. *Qual Life Res*. 2013;22:2663–73. <https://doi.org/10.1007/s11136-013-0402-0>.
68. L'Hermier des Plantes H. Structuration des tableaux à trois indices de la statistique. Théorie et applications d'une méthode d'analyse conjointe. 1976. Thèse 3^e cycle, USTL, Montpellier.
69. Escoufier Y. Opérateur associé à un tableau de données. *Annales de L'Insee*. 1976;22–23:165–79. <https://doi.org/10.2307/20075217>.
70. Vicente-Galindo P. Contribuciones al análisis de datos de calidad de vida relacionada con la salud. 2003. Doctoral thesis, University of Salamanca, Spain.
71. Krzanowski WJ. Between-groups comparison of principal components. *J Am Stat Assoc*. 1979;74(367):703–7. <https://doi.org/10.1080/01621459.1979.10481674>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



ANEXO 2. Trabajos en proceso de publicar

Producto 1	
Título:	MetaBiplot: Método multivariante como propuesta para la detección del <i>Response Shift</i>
Autores:	Estelina Ortega-Gómez Purificación Galindo-Villardón Purificación Vicente-Galindo
Revista:	Statistical Methods and Applications Web of Science Factor de Impacto: 1.314
Producto 2	
Título:	Exploring response shift in vision-related quality of life data in myopic patients after corneal refractive surgery with excimer laser using a dual STATIS multivariate three-way method
Autores:	Francisco Perdomo-Argüello Estelina Ortega-Gómez Purificación Galindo-Villardón Víctor Leiva Purificación Vicente-Galindo
Revista:	Mathematical Analysis and Applications Web of Science Factor de Impacto: 1.583
Producto 3	
Título:	Application of HJ-Biplot as an alternative to validate the structure of the questionnaire NEI VFQ-25: a case study of vision related quality of life
Autores:	Estelina Ortega-Gómez Purificación Galindo-Villardón Purificación Vicente-Galindo
Revista:	Journal Psychometrika Web of Science Factor de Impacto: 2.500

ANEXO 3. Participación en congresos nacionales e internacionales

Año 2015	
Nombre de Congreso:	XXX Foro Internacional de Estadística. “Estadística Aplicada al turismo y Desarrollo Sostenible” Acapulco, México
Fecha:	14 al 18 de septiembre de 2015
Título de la Ponencia:	Teoría de Respuesta al Ítem: Un análisis sobre la estructura del instrumento VFQ-25 aplicado a la evaluación de Calidad de Vida en pacientes miopes operados, en Costa Rica
Año 2016	
Nombre de Congreso:	XXVI Simposio Internacional de Estadística. Sucre, Colombia
Fecha:	8 al 12 de agosto de 2016
Título de la Ponencia:	Validación del cuestionario NEI-VFQ-25: una evaluación de la validez del constructo y consistencia interna de la fiabilidad
Año 2017	
Nombre de Congreso:	XIV Conferencia Española de Biometría. CEB 2017. Sevilla, España
Fecha:	13 al 15 de septiembre de 2017
Título de la Ponencia:	“STATIS-Dual”: Multivariate technique for identify change in vision related quality of life
Año 2021	
Nombre del Congreso:	V Jornada Científicas de Estudiantes de la Sociedad España de Bioestadística, Valencia, España
Fecha:	27 al 29 de enero de 2021
Título de la Ponencia:	Aplicación de la técnica HJ-Biplot como alternativa para validar la estructura del instrumento NEI-VFQ-25