



Universidad de Valladolid

FACULTAD DE CIENCIAS

TRABAJO FIN DE GRADO

Grado en Estadística

**Desarrollo y validación de un
cuestionario para analizar la
identidad profesional en Optometría**

Autora:

Paula Hernández Puchau

Tutores:

Jesús Alberto Tapia García

Raúl Martín Herranz

Curso 2022-2023

Agradecimientos

Mi más sincero agradecimiento a mi familia, por aguantarme y animarme en todo momento. A mis tutores Jesús y Raúl, por su tiempo, dedicación y confianza en mí. Al equipo investigador por diseñar el cuestionario y recoger las respuestas del estudio piloto para la validación, que detallo a continuación: Diego Arlanzón Lope (Medical Óptica Audición, España), Dr Luisa Simo (University of Portsmouth, Reino Unido), Dr Manbir Nagra (Anglia Ruskin University, Reino Unido), Dr Jenny Child (University of Plymouth, Reino Unido), Dr Ramón Solá i Parés (Universitat Politècnica de Catalunya, España) y Dr. Raúl Martín Herranz (Universidad de Valladolid, España).

Y finalmente a todos los que me han apoyado durante esta etapa.

Resumen

Los cuestionarios se utilizan principalmente para medir variables psicológicas como la ansiedad, la motivación o el conocimiento, pero para poder usarlos es necesario llevar a cabo un proceso de validación. En este trabajo se valida el cuestionario de identidad profesional en optometría con el Modelo de Escala de Clasificación, perteneciente a la Teoría de Respuesta al Ítem. Se realiza la validación con una muestra piloto de 90 individuos tomada en universidades y empresas del sector de la salud visual en Reino Unido. Tras este proceso de validación y corrección, el cuestionario queda preparado para su utilización.

Abstract

Questionnaires are primarily used to measure psychological variables such as anxiety, motivation, or knowledge. However, in order to use them, a validation process needs to be conducted. In this study, the professional identity questionnaire in optometry is validated using an Item Response Theory model named Rating Scale Model. The validation is performed with a pilot sample of 90 individuals taken from universities and companies from the visual health sector in the United Kingdom. After this validation and refinement process, the questionnaire is ready for implementation.

Índice

1. Introducción.....	7
2. Técnicas de validación de cuestionarios	9
2.1. Teoría clásica del test.....	9
2.2. Teoría de Respuesta al Ítem.....	10
2.2.1. Modelo Rasch.....	11
2.2.2. Modelo de Escala de Clasificación	16
3. Caso práctico: ajuste del Modelo de Escala de Clasificación al cuestionario de identidad profesional en optometría	21
Paso 1: Polaridad de los ítems	22
Paso 2: Ajuste de los ítems.....	24
Paso 3: Estructura de las categorías	28
Paso 4: Ajuste de los sujetos.....	33
Paso 5: Unidimensionalidad.....	34
Paso 6: Separación y confiabilidad de los ítems y de las personas.....	35
Paso 7: Independencia local de los ítems.....	36
Paso 8: Funcionamiento diferencial del ítem	37
Paso 9: Objetividad específica	38
4. Conclusiones.....	39
5. Bibliografía	40
Anexo I: cuestionario	41
Anexo II: regresión lineal simple para verificar la objetividad específica.....	46

Índice de figuras

Figura 1. Continuo lineal de la variable latente con la representación de la medida de un ítem β y la medida de una persona θ	12
Figura 2. Curvas de probabilidad de las categorías de una escala Likert de 3 puntos..	17
Figura 3. Representación de los umbrales τ_i de un ítem en el ejemplo del examen de matemáticas.....	18
Figura 4. Representación de los umbrales τ_i de un ítem en el cuestionario de identidad profesional en optometría.....	20
Figura 5. Categorías de respuesta con la escala Likert de 5 puntos.....	21
Figura 6. Respuestas inesperadas de las 26 personas con peor ajuste.....	26
Figura 7. Respuestas esperadas según la medida de la persona.....	26
Figura 8. Medidas medias de cada categoría.....	28
Figura 9. Medida estimada para cada umbral.....	29
Figura 10. Curvas de probabilidad de las categorías.....	30
Figura 11. Medidas medias de cada categoría.....	31
Figura 12. Medida estimada para cada umbral.....	31
Figura 13. Curvas de probabilidad de las categorías.....	32
Figura 14. Categorías de respuesta con la escala Likert de 4 puntos.....	32

Índice de tablas

Tabla 1. Correlación biserial puntual de cada ítem	23
Tabla 2. Estadísticos de diagnóstico del ajuste de los ítems	25
Tabla 3. Resumen de la estructura de las categorías (escala Likert de 5 puntos).....	28
Tabla 4. Resumen de la estructura de las categorías (escala Likert de 4 puntos).....	30
Tabla 5. Estadísticos de diagnóstico del ajuste de los ítems	33
Tabla 6. Estadísticos de diagnóstico de las personas que peor se ajustan al Modelo de Escala de Clasificación	34
Tabla 7. Análisis en Componentes Principales de los residuos estandarizados.....	34
Tabla 8. Separabilidad y confiabilidad de las personas	36
Tabla 9. Separación y confiabilidad de los ítems.....	36
Tabla 10. Correlaciones residuales estandarizadas más fuertes.....	37

1. Introducción

El objetivo de este trabajo fin de grado es obtener un instrumento — cuestionario — válido y fiable para medir la identidad profesional en el ámbito de la optometría.

Los instrumentos científicos como el termómetro o el cronómetro se han empleado ampliamente para medir propiedades físicas, pero ¿cómo se cuantifican las propiedades cognitivas?, ¿qué instrumentos existen para evaluar estas variables intangibles y abstractas? Es en ese contexto donde los cuestionarios desempeñan un papel fundamental para medir este tipo de variables.

El propósito del cuestionario que se trata de validar en este trabajo es medir un rasgo psicológico específico, en concreto, la identidad profesional en el ámbito de la optometría. Este rasgo se considera una variable latente, es decir, no es directamente observable, sino que se infiere a partir de las respuestas del cuestionario – también llamadas ítems –.

La primera versión de este cuestionario fue redactada en inglés por un equipo de profesionales de la salud visual en 2021. Está compuesto por un total de 35 preguntas o ítems. De estos, 18 se miden utilizando una escala Likert de 5 puntos, lo que implica que los participantes deben seleccionar una respuesta de un conjunto de opciones que varían desde “totalmente en desacuerdo” hasta “totalmente de acuerdo”. Por otro lado, las 17 preguntas restantes son principalmente cuestiones categóricas de respuesta abierta, donde los participantes deben elegir entre opciones predefinidas o escribir su propia respuesta si ninguna categoría se ajusta a su situación.

¿Y cómo se sabe si realmente el instrumento está midiendo el rasgo latente?

En el proceso de validación del cuestionario es necesario identificar y eliminar los ítems redundantes. Esto implica identificar aquellos ítems que miden conceptos o aspectos similares, ya que proporcionarían información duplicada. Al mantener únicamente los ítems más representativos y distintos, se asegura una medición más precisa y confiable del rasgo específico de interés.

Además, es importante descartar los ítems que puedan estar midiendo otro rasgo distinto al que se pretende evaluar. Esto se debe a que la inclusión de ítems que miden rasgos no relacionados puede afectar a la validez y la interpretación de los resultados del cuestionario.

Para realizar la validación el equipo investigador, formado por profesores, investigadores y profesionales Optometristas de la Universidad de Valladolid, Universidad Politécnica de Cataluña y de las Universidades de Portsmouth, Plymouth y Anglia Ruskin de Reino Unido, diseñaron el cuestionario piloto para medir la identidad profesional. Este cuestionario se envió a una muestra de profesionales de entre los contactos del equipo investigador de las universidades de Reino Unido, asegurando la representación de distintos tipos de profesionales como optometristas, asistentes de óptica o contactólogos.

El cuestionario se realizó con la plataforma JISC online surveys y la invitación a participar se envió por correo electrónico solicitando un plazo para responder de 2 semanas aproximadamente. Se obtuvo una muestra piloto de 90 individuos.

El protocolo del estudio fue aprobado por el Comité de Ética (Science and Health Faculty Ethics Committee) de la Universidad de Portsmouth el 12 de mayo de 2021 con número de expediente SHFEC 2021-052.

Se comenzará introduciendo las dos teorías que existen para analizar y validar cuestionarios. Después se pasará a explicar las características del Modelo de Rasch y del Modelo de Escala de Clasificación y, finalmente, se ajustará este último modelo a los datos de este estudio.

2. Técnicas de validación de cuestionarios

En este capítulo se introducirán las dos teorías fundamentales que guían la construcción, el análisis y la validación de los cuestionarios.

2.1. Teoría clásica del test

El enfoque clásico del análisis de los cuestionarios se desarrolló a principios del siglo XX. Se basa en el Modelo Lineal Clásico (Spearman, 1904), donde se asume que la puntuación observada en el test — puntuación empírica — es la suma de la puntuación verdadera y un error aleatorio. Este error puede ser debido al propio cuestionario, al individuo que está respondiendo, a causas del ambiente externo, etc. (Stanley 1971). Se encontraron importantes limitaciones en este modelo:

Era el problema de la invarianza de las mediciones y de las propiedades de los instrumentos de medida [...] Pues si se aspira a una medición rigurosa y científica resulta difícil justificar que las mediciones estén en función del instrumento utilizado. Algo así como si la longitud de un objeto dependiese del tipo de regla utilizado para medirlo. Este problema no tenía solución dentro del marco clásico...

La otra invarianza reclamada, era la de las propiedades de los instrumentos de medida respecto de los objetos medidos, a saber, la mayoría de las propiedades de un test, por ejemplo, en su dificultad, dependían del tipo de sujetos utilizado para calcularla, lo cual es inadmisibles a todas luces para un instrumento de medida. (Muñiz & Hambleton, 1992, p. 44)

2.2. Teoría de Respuesta al Ítem

La teoría de respuesta al ítem es una alternativa contemporánea a la Teoría Clásica del Test. A partir de mediados del siglo XX comenzaron a surgir los llamados modelos de rasgos latentes pertenecientes a la Teoría de Respuesta al Ítem. Estos modelos dan solución a algunos de los problemas citados anteriormente de la Teoría Clásica del Test.

Este nuevo enfoque es más complejo computacionalmente, pero como se ha mencionado, aporta importantes cambios necesarios para el estudio de los cuestionarios y es por ello que se ha decidido seguir esta teoría más actual.

2.2.1. Modelo Rasch

Se trata de un modelo logístico propuesto por el matemático danés Georg Rasch en 1960. Está diseñado para analizar datos de estudios donde la respuesta a los ítems sea de tipo binaria. Normalmente se habla de éxito/fracaso, estar de acuerdo/estar en desacuerdo, responder correctamente/incorrectamente un ítem o cualquier par de sucesos mutuamente excluyentes. El modelo pertenece a la Teoría de Respuesta al Ítem mencionada previamente y ha sido extensamente empleado en el análisis de pruebas educativas y psicológicas.

Formulación del modelo:

$$\log \left(\frac{P_{is1}}{P_{is0}} \right) = \theta_s - \beta_i \quad (1)$$
$$i = 1 \dots p \quad s = 1 \dots n$$

Donde P_{is1} es la probabilidad de que la respuesta de la persona s al ítem i sea correcta, P_{is0} es la probabilidad de que sea incorrecta, θ_s es el nivel de la persona s — competencia de la persona —, β_i es el nivel del ítem i — dificultad del ítem —, p es el número de ítems del cuestionario y n es el número de sujetos encuestados.

Teniendo en cuenta que los sucesos de fracaso y éxito son mutuamente excluyentes y que son los únicos sucesos posibles, la suma de sus probabilidades debe ser 1 y por tanto podemos escribir la ecuación (1) como:

$$\log \left(\frac{P_{is1}}{1 - P_{is1}} \right) = \theta_s - \beta_i \quad (2)$$

A la estimación de la dificultad de un ítem se le llama medida del ítem y a la estimación de la competencia de una persona, medida de la persona. Estos parámetros se estiman por el método de máxima verosimilitud, más detalles de estos procedimientos disponibles en Best Test Design (Wright & Stone, 1979). Como se puede observar en la ecuación (2), el logaritmo de las oportunidades de éxito — log odds — se modela como la diferencia entre la competencia de la persona y la dificultad del ítem. Por este motivo, las unidades

de las medidas de los ítems y de las personas son logits — contracción de “log odds units” —. Una de las ventajas de este modelo es el hecho de poder representar estas dos medidas en un mismo continuo lineal y realizar comparaciones entre ítems y personas.

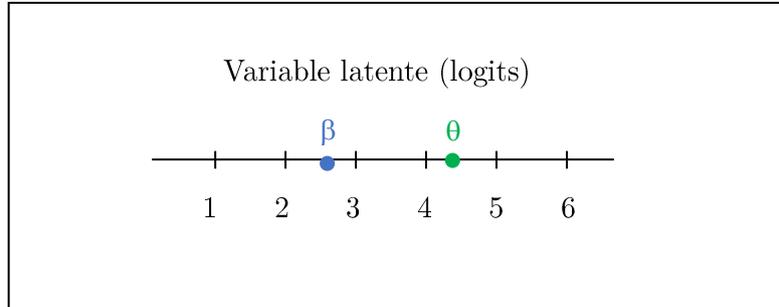


Figura 1. Continuo lineal de la variable latente con la representación de la medida de un ítem β y la medida de una persona θ

En cuanto a la interpretación de las medidas en esta escala, se puede establecer que una medida de ítem mayor indica que ese ítem presenta mayor dificultad en comparación con los demás, mientras que una medida de ítem menor indica que ese ítem es más fácil. Del mismo modo, una medida de persona mayor indica que la persona posee un mayor nivel de competencia en relación con los demás, y una medida menor indica que la persona tiene un nivel de competencia inferior. Estas medidas reflejan las posiciones relativas en la variable latente y por tanto solo se pueden interpretar en comparación con otras medidas del modelo. Además de poder comparar entre ítems y personas, también se pueden realizar comparaciones entre ambos, permitiendo determinar si la competencia de una persona es mayor o menor a lo requerido para responder a un determinado ítem.

Otra formulación equivalente del modelo:

$$P_{is1} = \frac{\exp\{\theta_s - \beta_i\}}{1 + \exp\{\theta_s - \beta_i\}} \quad (3)$$

Cuando la competencia de la persona coincide con la dificultad del ítem es fácil ver que es igual de probable contestar correctamente que incorrectamente.

$$P_{is1} = \frac{\exp\{\theta_s - \beta_i\}}{1 + \exp\{\theta_s - \beta_i\}} = \frac{1}{2} \quad \text{si } \theta_s = \beta_i \quad (4)$$

Cuando la competencia de la persona es mayor que la dificultad del ítem, la probabilidad de contestar correctamente es mayor que la probabilidad de contestar incorrectamente.

$$P_{is1} = \frac{\exp\{\theta_s - \beta_i\}}{1 + \exp\{\theta_s - \beta_i\}} > \frac{1}{2} \quad \text{si } \theta_s > \beta_i \quad (5)$$

Y, por último, cuando la competencia de la persona es menor que la dificultad del ítem, la probabilidad de contestar correctamente es menor que la probabilidad de contestar incorrectamente.

$$P_{is1} = \frac{\exp\{\theta_s - \beta_i\}}{1 + \exp\{\theta_s - \beta_i\}} < \frac{1}{2} \quad \text{si } \theta_s < \beta_i \quad (6)$$

Se debe tener en consideración los tres supuestos que los datos deben cumplir para garantizar una correcta aplicación del modelo:

Unidimensionalidad. Esto es, que los ítems formen una sola dimensión — midan un único rasgo o variable latente —. Si los datos violan este supuesto se podría optar por un Modelo de Rasch multidimensional.

Invarianza de los ítems. La segunda suposición es que existe una relación funcional entre la probabilidad de acertar un ítem y la competencia de la persona y la dificultad del ítem. Se supone que los ítems tienen características invariantes, lo que significa que su función de probabilidad de respuesta no varía en diferentes grupos de individuos con el mismo nivel de habilidad o rasgo. Esto permite que los ítems sean comparables y que las estimaciones de los parámetros sean válidas entre diferentes grupos.

Independencia local. Este supuesto se compone de dos aspectos distintos. El primero es la independencia local de los ítems. Cuando la respuesta a un ítem puede depender de la respuesta a un ítem anterior se está violando el supuesto. Por ejemplo, si se tuviera un cuestionario para medir la ansiedad y dentro de este se tuvieran 2 preguntas sobre situaciones sociales que pueden producir ansiedad, sería probable que, si un individuo responde a la primera que sí le produce ansiedad, a la segunda pregunta respondería igual porque los dos ítems están relacionados con la ansiedad social y no se cumpliría la independencia. El segundo es la independencia local de las personas. Cuando existe un parámetro relacionado con alguna característica de los individuos que relaciona sus respuestas se está violando este aspecto del supuesto. Por ejemplo, si hubiera diferencias significativas entre las respuestas de las mujeres y los hombres, se estaría ante una violación de la independencia local de las personas.

A continuación, se ilustrará la necesidad del Modelo Rasch con un sencillo ejemplo (Boone, 2016).

Cuando se utiliza una escala lineal para medir, como en el caso de medir la altura de 30 niños utilizando un metro, podemos realizar comparaciones directas entre las alturas. Sin embargo, surgen dificultades cuando la escala de medición no es lineal. Tomemos como ejemplo un examen de historia en el que se utiliza un formato de respuestas de opción múltiple para evaluar a estudiantes de 6º grado de primaria. Supongamos que se incluyen 15 preguntas con un nivel adecuado para estos estudiantes y 5 preguntas mucho más avanzadas (nivel de 2º de bachillerato).

En este caso, las puntuaciones obtenidas por los estudiantes en el examen no pueden ser comparadas de la misma manera que las alturas. ¿Por qué ocurre esto? La razón es que la escala de puntuación no es lineal. Si sumamos los puntos obtenidos por cada pregunta para obtener la puntuación final del examen, nos podríamos cuestionar: ¿todas las preguntas deberían tener el mismo valor? ¿Son todas igual de difíciles? La respuesta es no en ambos casos.

Los estudiantes A, B, C y D respondieron correctamente a 19, 15, 10 y 6 preguntas respectivamente. La diferencia de puntuación entre los estudiantes A y B, y los estudiantes C y D, es la misma, es decir, 4 puntos. Parecería que la diferencia en conocimiento es la misma en ambos casos, pero esto no es correcto. La razón es que el estudiante A respondió correctamente a 4 preguntas de nivel de 2º de bachillerato, mientras que el estudiante B no lo hizo. En realidad, la diferencia en conocimiento entre estos dos estudiantes es mucho mayor que en el caso de la segunda pareja de estudiantes. No se ha tenido en cuenta la dificultad de las preguntas en esta prueba y es precisamente en el Modelo Rasch donde se tiene en cuenta.

El Modelo Rasch aborda este problema al tener en cuenta la dificultad de las preguntas y la capacidad de los estudiantes en la materia que se está intentando medir.

2.2.2. Modelo de Escala de Clasificación

El modelo original de Rasch que se acaba de presentar es para ítems de respuesta dicotómica, pero los ítems del cuestionario de identidad profesional en optometría se miden en una escala ordenada de respuestas categóricas, la escala Likert de 5 puntos.

Existen numerosas versiones del Modelo de Rasch Politémico. Entre las más importantes para respuestas ordinales se encuentran el Modelo de Crédito Parcial — Partial Credit Model — y el Modelo de Escala de Clasificación — Rating Scale Model —.

En el caso del Modelo de Crédito Parcial la dificultad β_i que antes se estimaba para cada ítem ahora se estima para cada categoría de cada ítem β_{ik} donde $i = 1 \dots p$ — p es el número de ítems — y $k = 1 \dots m_i$ — $m_i + 1$ es el número de categorías del ítem i —. Este modelo está preparado para analizar ítems con escalas de respuesta propias de cada ítem.

En cambio, el Modelo de Escala de Clasificación asume que la escala de respuesta es igual para todos los ítems. Estima la dificultad β_i como en el Modelo Rasch y añade m parámetros comunes para todos los ítems consiguiendo así agregar la dificultad de cada categoría a la de cada ítem.

Como los ítems del cuestionario que se pretende analizar se miden todos en la misma escala Likert de 5 puntos, parece razonable asumir que esta escala se comporta de forma similar en todos los ítems y por tanto se debe ajustar un Modelo de Escala de Clasificación. Es importante destacar que, una vez realizado el ajuste, se verificará la validez de este supuesto.

Como ya se ha mencionado anteriormente, el Modelo de Escala de Clasificación es una extensión del Modelo Rasch y por tanto comparte los supuestos y características fundamentales del modelo original. A continuación, se estudiarán las diferencias del modelo que propuso Andrich en 1978.

Formulación del modelo:

$$\log \left(\frac{P_{isj}}{P_{is(j-1)} + P_{isj}} \right) = \theta_s - \beta_i - \tau_j \quad (7)$$

$$i = 1 \dots p \quad s = 1 \dots n \quad j = 1 \dots m$$

Donde τ_j es el umbral o límite entre la categoría j y la categoría $j-1$, P_{isj} es la probabilidad de que la respuesta de la persona s al ítem i sea la categoría j , $m+1$ es el número de categorías y los demás parámetros son los mismos que en el Modelo Rasch original.

Teniendo en cuenta que $\sum_{j=0}^m P_{isj}$ se puede obtener una formulación equivalente del modelo en función de las probabilidades:

$$P_{isj} = \frac{\exp\{\sum_{k=0}^j (\theta_s - \beta_i - \tau_k)\}}{\sum_{l=0}^m \exp\{\sum_{k=0}^l (\theta_s - \beta_i - \tau_k)\}} \quad (8)$$

$$i = 1 \dots p \quad s = 1 \dots n \quad j = 0 \dots m$$

$$\text{con } \tau_0 = 0 \quad \text{y} \quad \exp\{\sum_{k=0}^0 (\theta_s - \beta_i - \tau_0)\} = 1$$

Generalmente se define el concepto de “pasos” entre categorías como la diferencia en dificultad requerida para pasar de una categoría de respuesta a la siguiente. Los umbrales son los puntos donde la probabilidad de elegir las dos categorías coincide, así cuando pasa el umbral entre dos categorías se da el paso y si no lo supera no da el paso.

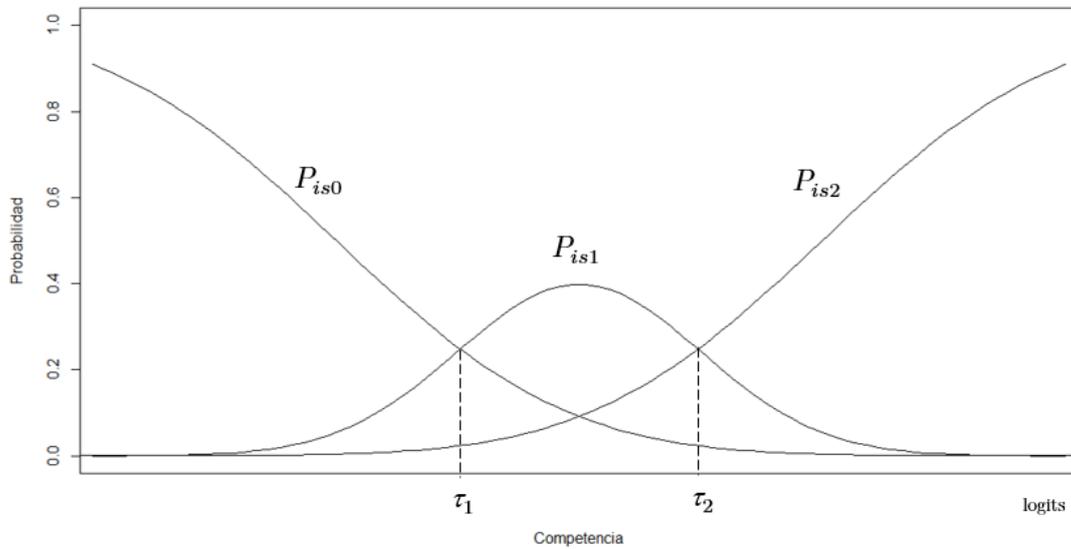


Figura 2. Curvas de probabilidad de las categorías de una escala Likert de 3 puntos

Se ilustraran estas nuevas ideas con el siguiente ejemplo.

- 1.- Calcula:
- a) $3 + 5$
 - b) $(3 + 5)^2$
 - c) $\int (3 + 5)^2 x dx$

Se tiene el problema de la izquierda en un examen de matemáticas. Consta de 3 apartados y cada uno puntúa 1 si la respuesta es correcta, pero tiene una peculiaridad, si el sujeto no contesta bien al apartado anterior no puede contestar al siguiente. Así, si una persona no contesta bien el apartado a) no puede contestar al b) y así con los 3 apartados. Además, los apartados están ordenados en aumento por su nivel de dificultad. Desde el punto de vista del Modelo de Escala de Clasificación se tiene un ítem con 4 categorías relacionadas con las posibles puntuaciones del problema (0,1,2,3) y 3 “pasos” desde una categoría a otra. El primer paso sería “pasar de tener un 0 a tener un 1 contestando bien en el apartado a)”, el segundo paso sería “pasar de tener un 1 a tener un 2 contestando bien en el apartado b)” y el tercer paso sería “pasar de tener un 2 a tener un 3 contestando bien en el apartado c)”.

A continuación, se representa la escala que se obtendría al ajustar el modelo para este ejemplo junto a la estimación β de la dificultad del ítem, la estimación τ_i de los umbrales entre categorías y las propias categorías – calificaciones – en rojo.

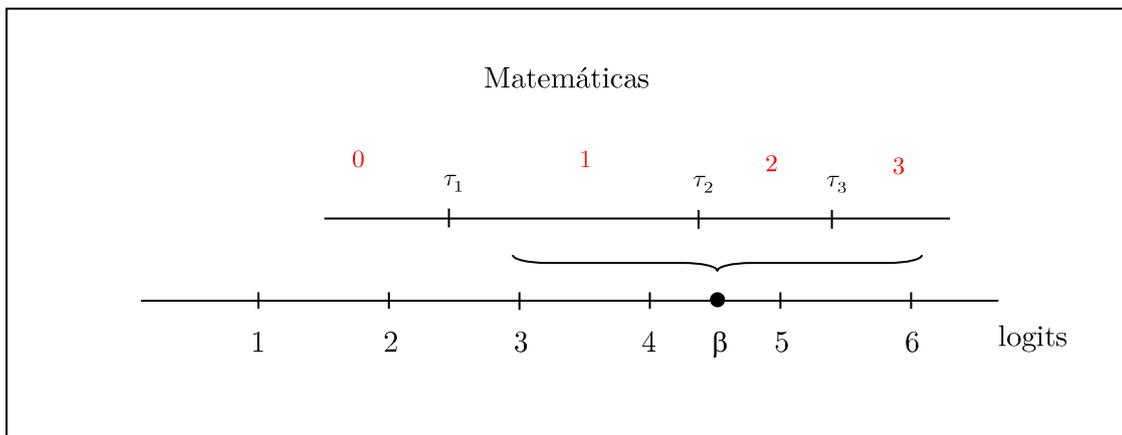


Figura 3. Representación de los umbrales τ_i de un ítem en el ejemplo del examen de matemáticas

Así, un alumno con una competencia en matemáticas de 2 logits obtendría una calificación de 0, ya que no ha superado el umbral τ_1 y por tanto no tiene el conocimiento

necesario para acertar el apartado a). Por otro lado, un alumno con una competencia en matemáticas de 5 logits obtendría un 2 ya que ha superado los umbrales τ_1 y τ_2 y por tanto ha conseguido responder correctamente al apartado a) y b) pero no ha superado τ_3 y no ha podido responder correctamente al c).

Ahora que ya se ha establecido el funcionamiento de estos nuevos parámetros se abordará desde el punto de vista de nuestro cuestionario.

Aunque no sea tan evidente, las categorías de la escala Likert se comportan como las categorías del problema anterior. La dificultad de un ítem, β , será el nivel de exigencia del ítem. Un ítem con un alto nivel de exigencia — un ítem difícil — será aquel que requiera un grado de identificación con la profesión de optometría muy alto para estar de acuerdo. En cambio, un ítem con un nivel de exigencia bajo será aquel que no requiera un grado de identificación tan alto para estar de acuerdo con la afirmación.

En este contexto, las categorías están ordenadas en la escala Likert de menor acuerdo a mayor y por eso, dar un paso significaría escoger una categoría más a la derecha de la escala, es decir, estar más de acuerdo con el ítem — y por tanto más identificado —. Por ejemplo, dar el primer paso sería escoger la categoría “en desacuerdo” en vez de “totalmente en desacuerdo”. Y no dar el paso, por tanto, sería escoger “totalmente en desacuerdo” en vez de “en desacuerdo”. De igual manera que en el ejemplo del problema matemático, no se puede dar un paso si no se ha dado el anterior. Así, para que un individuo haya escogido la opción “de acuerdo” ha tenido que dar los primeros tres pasos, pero no ha dado el último.

Seguidamente se muestra la representación de un ítem y las categorías de acuerdo en la escala de identidad profesional.

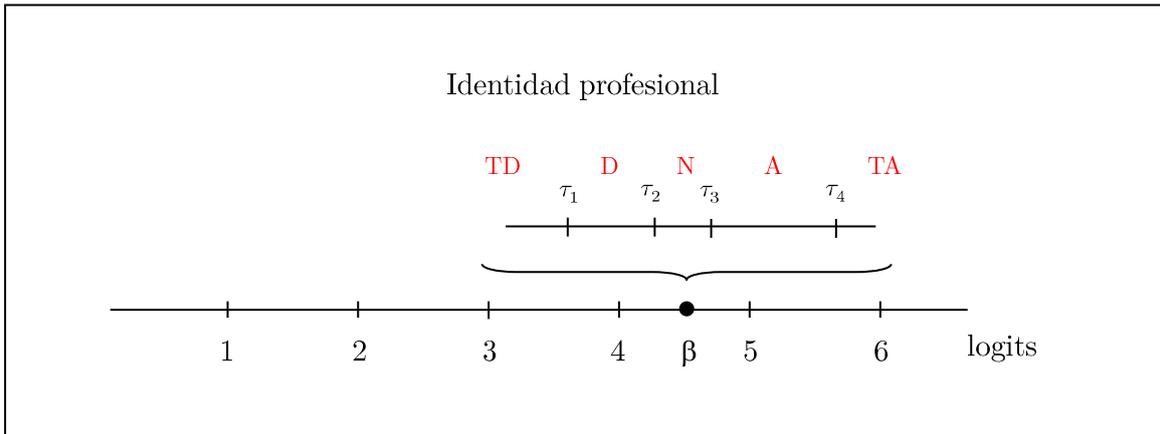


Figura 4. Representación de los umbrales τ_i de un ítem en el cuestionario de identidad profesional en optometría

Como en el Modelo de Escala de Clasificación se asume que los pasos no varían de un ítem a otro, solo se estimarán 4 umbrales comunes para todos los ítems con toda la muestra.

3. Caso práctico: ajuste del Modelo de Escala de Clasificación al cuestionario de identidad profesional en optometría

En este capítulo se va a realizar el ajuste y la validación del cuestionario de identidad profesional en optometría paso a paso con la versión gratuita del programa estadístico Winsteps llamada Ministeps. Las dos versiones del programa están diseñadas para ajustar distintos modelos de la Teoría de Respuesta al Ítem como el Modelo Rasch, el Modelo de Escala de Clasificación o el Modelo de Crédito Parcial. Además, proporcionan tablas y gráficos con las estimaciones de los parámetros del modelo y medidas para evaluar el ajuste, el cumplimiento de los supuestos y la calidad del cuestionario. La única diferencia entre las dos versiones del programa es que Ministeps está restringido a 25 ítems y 75 personas encuestadas.

Se codificaron las categorías de la Figura 5 como 1 = "totalmente en desacuerdo", 2 = "en desacuerdo", 3 = "neutral", 4 = "de acuerdo" y 5 = "totalmente de acuerdo".

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
-----------------------------	------------------	---------	---------------	--------------------------

Figura 5. Categorías de respuesta con la escala Likert de 5 puntos

Para el ajuste se eliminaron los sujetos que no habían contestado a todos los ítems, quedando la muestra piloto con 71 observaciones. Una vez realizada la validación del cuestionario no es necesario eliminar las observaciones incompletas, ya que solo se necesitan las respuestas a unos pocos ítems para estimar el grado de identificación de una persona. Incluso se pueden predecir las respuestas a los ítems que no se contesten. Pero para validar el cuestionario sí que es necesario excluir estos individuos.

Paso 1: Polaridad de los ítems

Lo primero que se ha de comprobar es que todos los ítems tienen la misma polaridad, esto es, que cuanto más se identifique un individuo con la profesión de optometría más de acuerdo estará con la afirmación del ítem. Se sospecha que dos ítems del cuestionario de identidad están formulados con la polaridad invertida a todos los demás.

Los ítems son:

2. Pongo excusas por pertenecer a esta profesión.
24. Siento que pongo las necesidades del negocio antes que las decisiones clínicas.

Un ejemplo de un ítem sin la polaridad invertida sería el siguiente.

1. Estoy orgulloso de pertenecer al colectivo de profesionales de la salud visual.

Se puede intuir que alguien con un alto grado de identificación elegiría la categoría “totalmente de acuerdo” para responder al ítem 1, mientras que para los ítems 2 y 24 elegiría la categoría “totalmente en desacuerdo”. Para comprobar si esta hipótesis es cierta se debe ajustar el Modelo de Escala de Clasificación a los datos originales y estudiar la correlación de cada ítem con la variable latente identidad.

Una correlación biserial puntual negativa indica ítems con polaridad invertida (Linacre, 2002). Se trata de una correlación de Pearson entre las observaciones del ítem y las puntuaciones de las personas — obtenidas sumando las respuestas en formato numérico del cuestionario —. En la se puede observar las correlaciones para cada ítem. Como se había previsto, el ítem 2 y el ítem 24 tienen correlaciones negativas.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTBISERL-CORR.	-EX-EXP.	EXACT OBS%	MATCH EXP%	ITEM
1	313	71	-1.13	.20	1.06	.39	.94	-.25	.63	.35	59.2	56.2	Q2
2	143	71	2.57	.15	2.26	5.56	2.65	6.65	-.42	.38	33.8	47.6	Q3
3	289	71	-.35	.17	.83	-.92	.82	-.96	.71	.40	46.5	54.2	Q4
4	305	71	-.84	.18	.79	-1.13	.75	-1.45	.63	.37	60.6	55.6	Q5
5	310	71	-1.02	.19	.85	-.76	.81	-1.05	.54	.36	60.6	56.2	Q6
6	330	71	-1.91	.24	.97	-.11	.86	-.63	.42	.30	74.6	69.2	Q7
7	292	71	-.44	.17	.79	-1.18	.73	-1.57	.70	.39	60.6	54.3	Q9
8	277	71	-.04	.16	.64	-2.28	.63	-2.29	.73	.41	59.2	52.7	Q10
9	273	71	.06	.15	.61	-2.47	.64	-2.30	.65	.42	56.3	52.0	Q11
10	219	71	1.14	.13	.74	-1.89	.77	-1.58	.46	.43	45.1	40.1	Q12
11	264	71	.27	.15	.99	-.03	.94	-.27	.32	.42	54.9	49.6	Q13
12	238	71	.79	.14	1.24	1.47	1.27	1.59	.32	.43	33.8	41.0	Q14
13	279	71	-.09	.16	1.26	1.39	1.19	1.04	.33	.41	63.4	53.3	Q15
14	279	71	-.09	.16	.71	-1.74	.69	-1.87	.55	.41	63.4	53.3	Q16
15	269	71	.15	.15	1.20	1.14	1.10	.61	.47	.42	54.9	51.3	Q17
16	297	71	-.59	.18	.86	-.70	.86	-.72	.13	.38	52.1	54.0	Q23
17	313	71	-1.13	.20	1.01	.13	1.08	.48	.25	.35	60.6	56.2	Q24
18	140	71	2.64	.15	1.33	1.81	1.78	3.67	-.08	.38	42.3	47.5	Q25
MEAN	268.3	71.0	.00	.17	1.01	-.07	1.03	-.05			54.5	52.5	
P.SD	51.9	.0	1.16	.02	.37	1.86	.48	2.19			10.3	6.2	

Tabla 1. Correlación biserial puntual de cada ítem

No se trata de un grave problema. Simplemente se deben codificar al revés los ítems 2 y 24 e invertir la afirmación para el ajuste del Modelo de Escala de Clasificación. Si un sujeto había respondido al ítem 2 eligiendo la categoría “en desacuerdo”, ahora se deberá cambiar por la categoría “de acuerdo” y los ítems pasarán a ser los siguientes:

2. No pongo excusas por pertenecer a esta profesión.
24. Siento que pongo las decisiones clínicas antes que las necesidades del negocio.

Ahora para todos los ítems un alto grado de acuerdo con la afirmación significa un alto grado de identificación con la profesión de optometría.

Puede surgir la pregunta de por qué no se cambian estos ítems en el cuestionario como se está haciendo para el análisis. La respuesta es porque los ítems con polaridad invertida ayudan a que el encuestado no pierda la atención.

Paso 2: Ajuste de los ítems

Se considera que un ítem está mal ajustado cuando las respuestas de los encuestados a ese ítem no se comportan como se esperan en el Modelo de Escala de Clasificación. Esto puede ocurrir por muchas razones, por ejemplo, porque no esté redactado adecuadamente y por tanto los encuestados no estén entendiendo bien la afirmación o porque el ítem no esté realmente relacionado con el constructo que se desea medir.

Existen 4 estadísticos para identificar si un ítem se ajusta bien al modelo o no: *infit* MNSQ, *infit* ZSTD, *outfit* MNSQ y *outfit* ZSTD. A continuación, se expondrán las diferencias entre los cuatro estadísticos.

Los estadísticos *infit* son más sensibles a las respuestas inesperadas por personas con un grado de identificación cercano al nivel de exigencia del ítem, mientras que los estadísticos *outfit* son más sensibles a respuestas inesperadas por personas con un grado de identificación alejado del nivel del ítem y, por tanto, detecta los casos atípicos.

MNSQ se refiere al *infit* u *outfit* cuadrático medio y ZSTD es el estadístico MNSQ estandarizado.

Un valor del estadístico MNSQ dentro del rango 0.5-1.5 indica un buen ajuste del ítem (Linacre, 2012). También se puede comprobar la calidad del ajuste con el estadístico ZSTD. Si está entre los valores -2 y 2 el ítem se ajusta correctamente al modelo (Bond & Fox, 2007).

En la se observa que el único ítem que presenta valores fuera del rango es el ítem 24. Esto quiere decir que las respuestas observadas en el ítem 24 son demasiado impredecibles. A continuación, se examinarán estas respuestas para tomar una decisión sobre qué hacer con este ítem.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	ITEM
1	313	71	-.88	.20	1.01	.13	.89	-.53	.61	.44	65.7	58.6	Q1
2	283	71	.13	.17	1.28	1.48	1.29	1.54	.51	.52	44.3	55.1	Q2
3	289	71	-.04	.17	.83	-.94	.84	-.89	.65	.51	50.0	55.7	Q3
4	305	71	-.57	.19	.86	-.71	.83	-.94	.58	.47	60.0	56.9	Q4
5	310	71	-.76	.20	.87	-.68	.81	-1.02	.55	.45	60.0	58.0	Q5
6	330	71	-1.70	.24	.97	-.09	.88	-.44	.44	.37	74.3	69.4	Q6
7	292	71	-.14	.18	.81	-1.05	.74	-1.54	.66	.50	64.3	55.9	Q8
8	277	71	.30	.16	.68	-1.93	.73	-1.64	.67	.53	64.3	53.8	Q9
9	273	71	.40	.16	.61	-2.48	.64	-2.26	.69	.54	61.4	52.4	Q10
10	219	71	1.59	.14	.91	-.53	.98	-.10	.54	.60	44.3	41.6	Q11
11	264	71	.63	.16	1.22	1.26	1.25	1.38	.37	.55	54.3	50.2	Q12
12	238	71	1.20	.14	1.30	1.78	1.29	1.68	.49	.58	35.7	43.6	Q13
13	279	71	.24	.17	1.34	1.76	1.28	1.49	.45	.53	64.3	54.6	Q14
14	279	71	.24	.17	.73	-1.55	.74	-1.56	.58	.53	61.4	54.6	Q15
15	269	71	.50	.16	1.31	1.66	1.25	1.37	.51	.54	45.7	51.3	Q16
16	297	71	-.30	.18	.90	-.46	.95	-.20	.29	.49	45.7	56.3	Q22
17	313	71	-.88	.20	1.05	.33	1.10	.58	.35	.44	57.1	58.6	Q23
18	286	71	.04	.17	1.47	2.29	1.81	3.66	.19	.51	45.7	55.7	Q24
MEAN	284.2	71.0	.00	.18	1.01	.02	1.02	.03			55.5	54.6	
P.SD	26.2	.0	.76	.02	.25	1.37	.29	1.50			9.9	5.8	

Tabla 2. Estadísticos de diagnóstico del ajuste de los ítems

La Figura 6 aporta información sobre las respuestas inesperadas en cada ítem. Las respuestas esperadas, en las que la categoría que ha elegido el sujeto ha coincidido con la categoría que el modelo predice para ese sujeto, se representan con un punto, mientras que las respuestas inesperadas se representan con la categoría observada.

Para ver cuáles eran las respuestas esperadas de estos individuos se dispone de una ilustración en la Figura 7. En ella se puede determinar la respuesta esperada a cada ítem dependiendo del grado de identificación estimado del individuo.

```

MOST MISFITTING RESPONSE STRINGS
      |PERSON
      |54463137455534431115313 423 6 56621
OUTMNSQ |24253400896526093754117959627712830 ITEM
      high-----
1.81 A|3.3.3.....2.....5.. 18 Q24
1.28 B|.....3.....2...1...1.....5... 13 Q14
1.25 C|.....22...2.....1.... 15 Q16
1.29 D|.....22.....5.....4.. 12 Q13
1.29 E|.....323.....2222..... 2 Q2
1.25 F|...3....2...2.2.....4. 11 Q12
1.10 G|...4.....2.....3..... 17 Q23
.89 H|...4.....3..3.....2.. 1 Q1
.98 I|.....2.....3 10 Q11
.88 i|.....4.....33..... 6 Q6
.95 h|.4..... 16 Q22
.81 g|.....3.....3..... 5 Q5
.83 f|.....3.3..... 4 Q4
.74 d|.....2..... 7 Q8
.74 c|.....2..... 14 Q15
.73 b|...3..... 8 Q9
      |-----low
      |54463137455534431115313942326756621
      |24253400896526093754117 596 7 12830

```

Figura 6. Respuestas inesperadas de las 26 personas con peor ajuste

```

-5 -4 -3 -2 -1 0 1 2 3 4 5 6 NUM ITEM
|-----|
1 | 1 : 2 : 3 : 4 : 5 5 10 Q11
|
1 | 1 : 2 : 3 : 4 : 5 5 12 Q13
|
1 | 1 : 2 : 3 : 4 : 5 5 11 Q12
1 | 1 : 2 : 3 : 4 : 5 5 15 Q16
1 | 1 : 2 : 3 : 4 : 5 5 9 Q10
1 | 1 : 2 : 3 : 4 : 5 5 8 Q9
1 | 1 : 2 : 3 : 4 : 5 5 13 Q14
1 | 1 : 2 : 3 : 4 : 5 5 14 Q15
1 | 1 : 2 : 3 : 4 : 5 5 2 Q2
1 | 1 : 2 : 3 : 4 : 5 5 18 Q24
1 | 1 : 2 : 3 : 4 : 5 5 3 Q3
1 | 1 : 2 : 3 : 4 : 5 5 7 Q8
1 | 1 : 2 : 3 : 4 : 5 5 16 Q22
|
1 | 1 : 2 : 3 : 4 : 5 5 4 Q4
|
1 | 1 : 2 : 3 : 4 : 5 5 5 Q5
1 | 1 : 2 : 3 : 4 : 5 5 1 Q1
1 | 1 : 2 : 3 : 4 : 5 5 17 Q23
|
1 | : 2 : 3 : 4 : 5 5 6 Q6
|-----|
-5 -4 -3 -2 -1 0 1 2 3 4 5 6 NUM ITEM

```

Figura 7. Respuestas esperadas según la medida de la persona

Es importante tener presente que por cómo está codificado el ítem 24 tras el análisis de su polaridad, la afirmación que se va a estudiar es “Siento que pongo las decisiones clínicas antes que las necesidades del negocio”.

El individuo 29 ha respondido “en desacuerdo”, pero según su grado de identificación medido, 1.06 logits, se esperaba que respondiera “de acuerdo”. Se podría considerar que no leyó detenidamente la afirmación y como el ítem tiene la polaridad invertida, pensó que estaba contestando lo contrario a como respondió. Si las demás respuestas inesperadas son similares a este caso, la solución sería no invertir la polaridad del ítem en el cuestionario, ya que más que ayudar a que los encuestados no pierdan la atención está provocando que lo interpreten de manera contraria y algunas respuestas sean incorrectas.

El siguiente individuo, el 68, que tiene un grado de identificación de -0.16 logits ha respondido la categoría “totalmente de acuerdo” pero se esperaba que escogiera “neutral”. Los individuos 33, 42 y 53, con grados de identificación 3.10, 3.32 y 4.23 respectivamente, han elegido “neutral”, pero según el Modelo de Escala de Clasificación ajustado deberían haber elegido “totalmente de acuerdo”.

Estas respuestas sugieren que el problema no es solo la polaridad del ítem. Revisando todos los ítems en conjunto se ha considerado que el ítem 24 genera confusión por su similitud con el ítem previo. La única diferencia entre los dos ítems es la palabra “puedo” y la colocación de los términos “decisiones clínicas” y “necesidades del negocio”, que son claves en los dos ítems. A continuación, se presentan las afirmaciones de las que se está hablando.

23. Siento que puedo poner las decisiones clínicas antes que las necesidades del negocio.

24. Siento que pongo las necesidades del negocio antes que las decisiones clínicas.

Por estas razones, se decide eliminar el ítem 24.

Paso 3: Estructura de las categorías

El siguiente paso consiste en identificar la estructura de categorías óptima para el cuestionario de identidad profesional. Se utilizó una escala Likert de 5 puntos para que los encuestados respondieran su nivel de acuerdo con cada ítem, pero realmente no se tiene la certeza de que sea la escala adecuada.

Para que las estimaciones de los umbrales sean estables debe haber un mínimo de 10 observaciones por categoría.

CATEGORY LABEL	SCORE	OBSERVED COUNT	OBSVD %	SAMPLE AVRG	INFINIT EXPECT	OUTFIT MNSQ	ANDRICH THRESHOLD	CATEGORY MEASURE
1	1	15	1	-.71	-.76	.99 1.09	NONE	(-3.33)
2	2	84	7	.12	-.03	1.13 1.16	-2.12	-1.44
3	3	171	14	.73	.75	1.03 .99	-.35	-.13
4	4	551	46	1.51	1.58	1.11 .99	-.02	1.39
5	5	386	32	2.75	2.67	.91 .93	2.48	(3.64)

Tabla 3. Resumen de la estructura de las categorías (escala Likert de 5 puntos)

Si las categorías son las óptimas, se cumplirán una serie de condiciones (Linacre, 2002):

- 1) Que la medida media de cada categoría, calculada como la media de la diferencia entre las medidas de las personas y las medidas de los ítems, vaya en aumento de forma monótona.

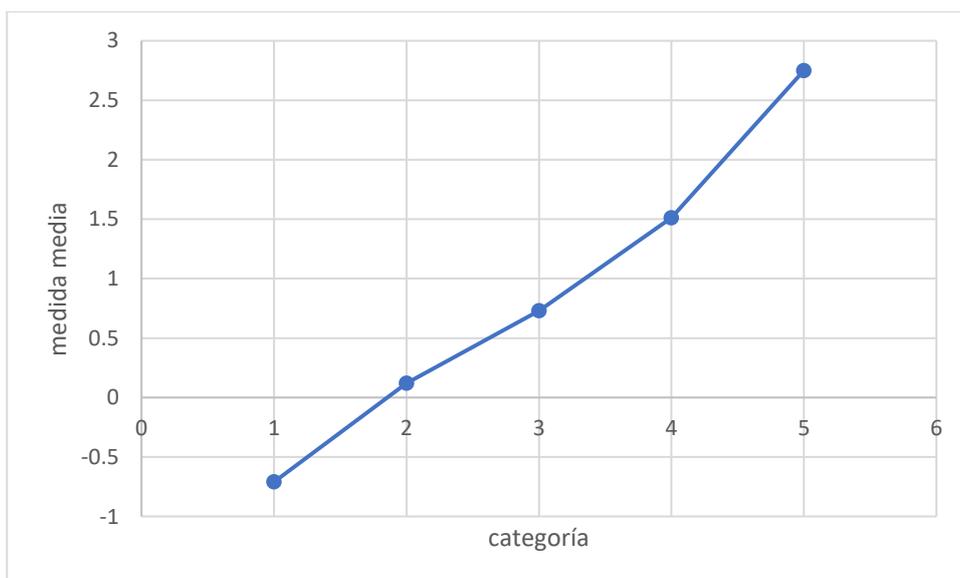


Figura 8. Medidas medias de cada categoría

Se cumple.

- 2) Que los umbrales aumenten también de forma monótona.

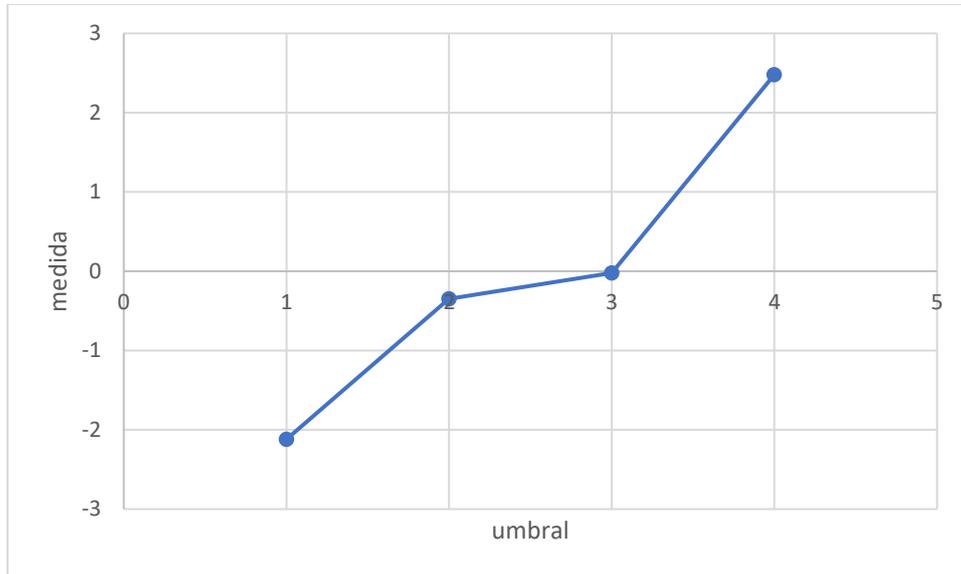


Figura 9. Medida estimada para cada umbral

No se cumple, el aumento del segundo umbral al tercero es mínimo comparado con los demás.

- 3) Que los estadísticos infit MNSQ y outfit MNSQ de cada categoría sean inferiores a 2.

Se cumple.

- 4) Que las distancias entre umbrales sean de al menos 1.4 logits para mostrar distinción entre categorías, pero no más de 5 logits porque entonces las categorías estarían demasiado separadas e indicaría la falta de una categoría intermedia.

No se cumple. Las distancias entre los umbrales son de 1.77, 0.33 y 2.5 logits.

La separación entre el segundo y tercer umbral no alcanza los 1.4 logits necesarios.

También se puede comprobar en la Figura 10.

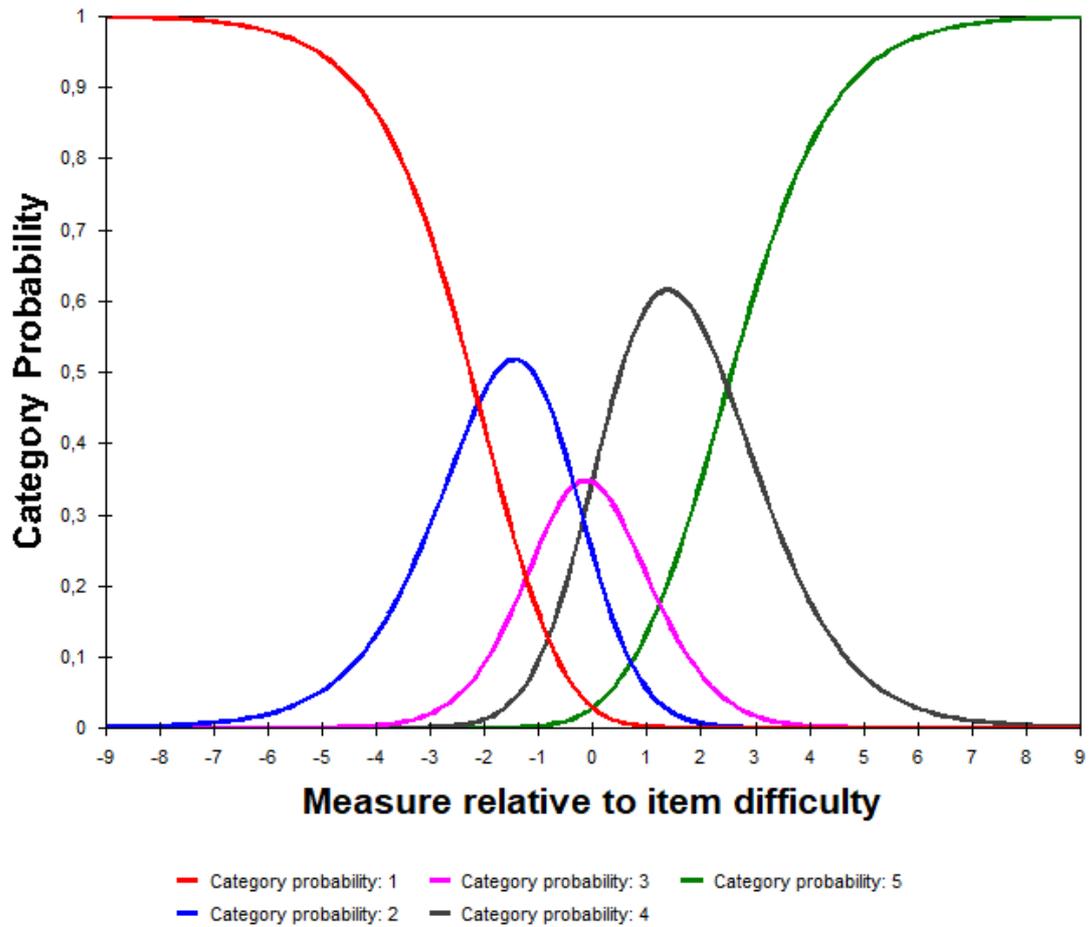


Figura 10. Curvas de probabilidad de las categorías

Tras el análisis de la estructura original de las categorías se decide eliminar la categoría central y repetir el análisis.

CATEGORY	OBSERVED	OBSVD	SAMPLE	INFINIT	OUTFIT	ANDRICH	CATEGORY			
LABEL	SCORE	COUNT	%	AVRGE	EXPECT	MNSQ	MNSQ	THRESHOLD	MEASURE	
1	1	15	1	-1.38	-1.51	1.03	1.17	NONE	(-3.66)	1
2	2	84	8	.14	.15	.99	.96	-2.43	-1.63	2
4	3	551	53	1.96	1.97	1.02	.95	-.80	1.24	4
5	4	386	37	3.77	3.75	1.00	.98	3.24	(4.35)	5
MISSING		171	14	1.18						

Tabla 4. Resumen de la estructura de las categorías (escala Likert de 4 puntos)

A continuación, se comprobará el cumplimiento de los criterios anteriormente citados con la nueva estructura de categorías.

1) Se cumple

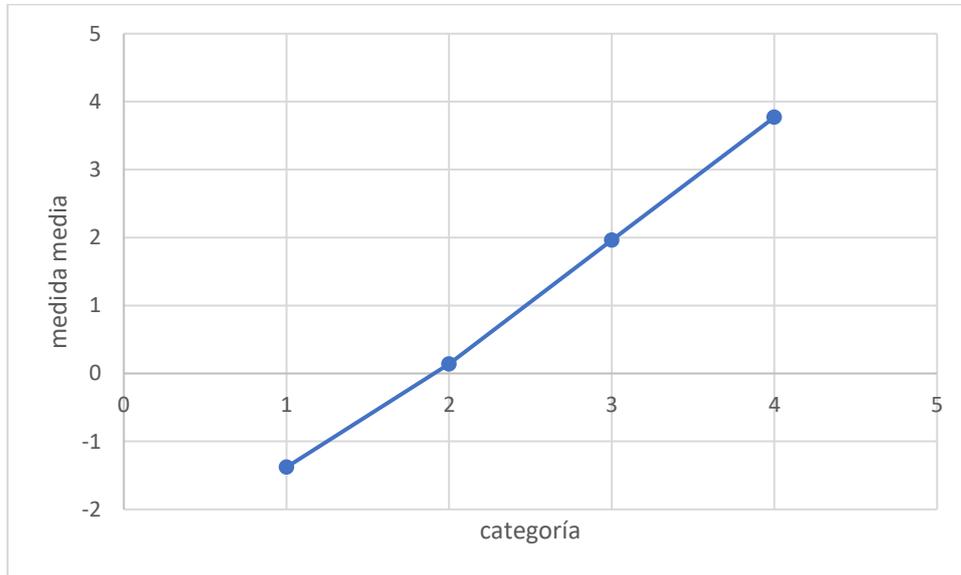


Figura 11. Medidas medias de cada categoría

2) Se cumple

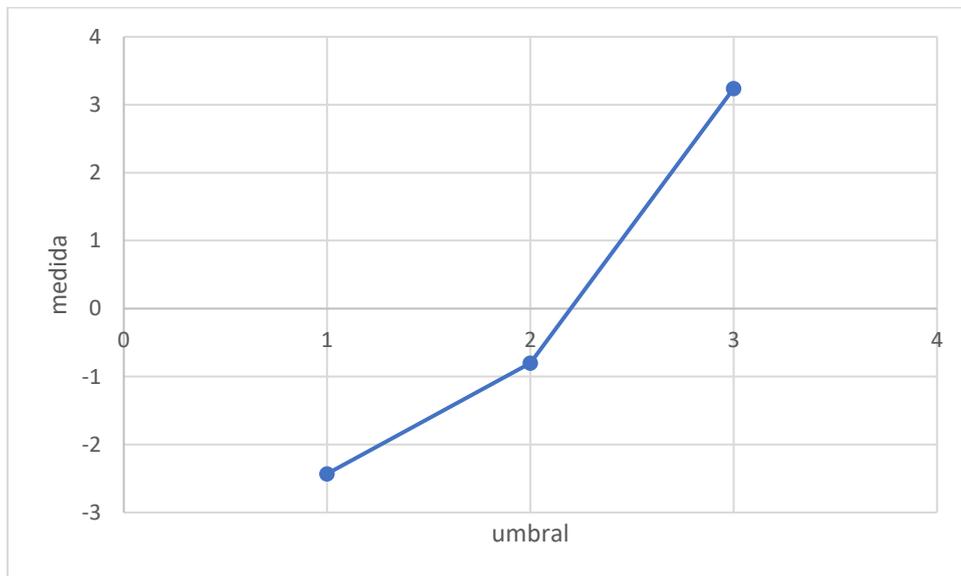


Figura 12. Medida estimada para cada umbral

3) Se cumple

4) Se cumple. Ahora existe una separación de 1.63 y 4.04 logits entre cada umbral.

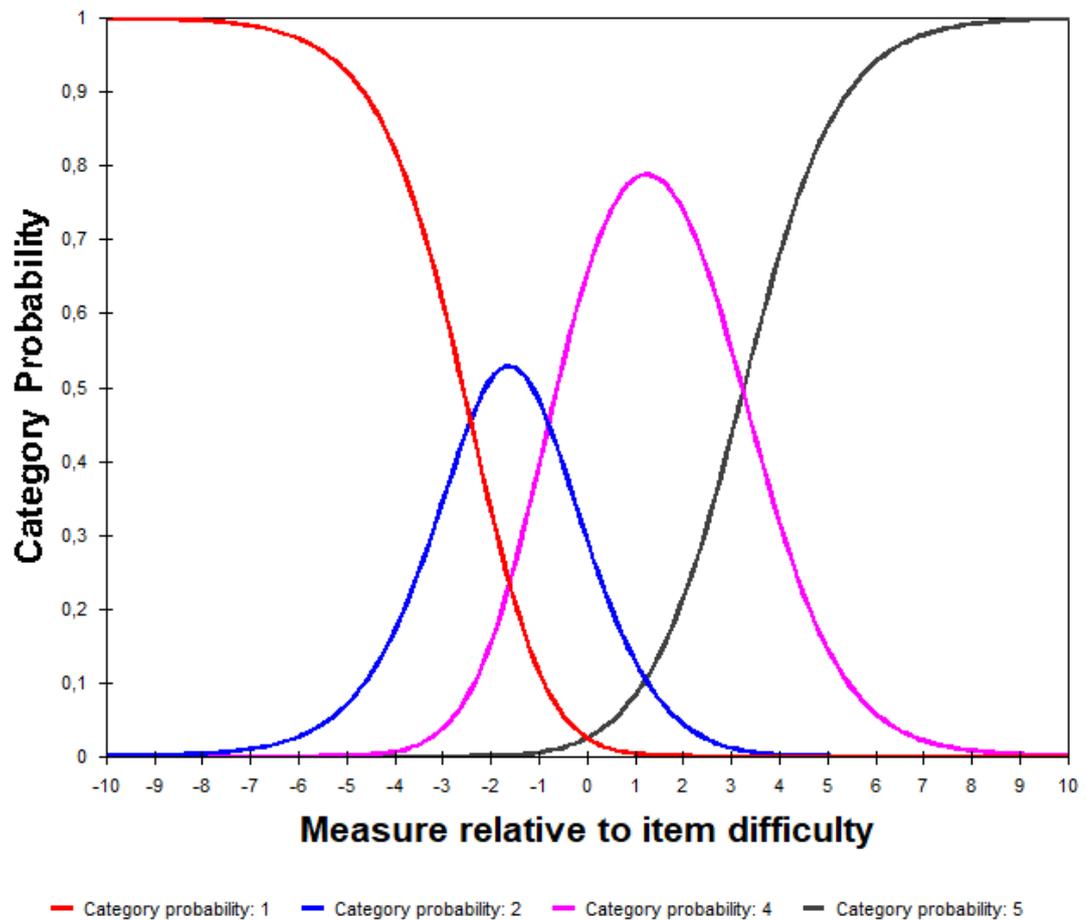


Figura 13. Curvas de probabilidad de las categorías

Se cumplen todos los requisitos, así que la categorización óptima para los ítems del cuestionario de identidad profesional en optometría será la siguiente:

Totalmente en desacuerdo	En desacuerdo	De acuerdo	Totalmente de acuerdo
-----------------------------	---------------	------------	--------------------------

Figura 14. Categorías de respuesta con la escala Likert de 4 puntos

Por último, se debe comprobar que después de esta modificación en la estructura de las categorías ningún ítem se ajusta mal. Siguiendo los criterios empleados en el Paso 2:

Ajuste de los ítems se concluye que los 17 ítems se ajustan bien al Modelo de Escala de Clasificación.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	ITEM
1	298	66	-1.37	.27	.85	-.97	.83	-.59	.63	.51	70.8	69.1	Q1
2	250	60	-.10	.26	1.30	1.56	1.26	1.22	.54	.58	64.4	68.4	Q2
3	253	59	-.33	.27	.95	-.23	.94	-.22	.65	.57	62.1	67.3	Q3
4	278	62	-1.01	.27	.77	-1.58	.70	-1.31	.60	.51	70.5	67.9	Q4
5	292	65	-1.10	.27	.77	-1.62	.69	-1.36	.61	.52	70.3	68.4	Q5
6	321	68	-2.20	.30	1.02	.16	.96	.04	.40	.43	79.1	76.9	Q6
7	277	66	-.10	.25	.78	-1.31	.72	-1.52	.71	.58	75.4	68.2	Q8
8	244	60	.28	.26	.59	-2.42	.59	-2.22	.77	.60	84.7	68.4	Q9
9	216	52	.26	.28	.55	-2.59	.52	-2.53	.76	.58	80.4	68.4	Q10
10	120	38	2.33	.28	1.18	.80	1.19	.82	.65	.66	43.2	61.4	Q11
11	228	59	1.05	.25	1.02	.18	.97	-.08	.50	.59	77.6	70.1	Q12
12	208	61	1.98	.24	1.43	1.95	1.50	2.11	.47	.65	46.7	66.5	Q13
13	252	62	.38	.26	1.46	2.13	1.42	1.82	.49	.60	68.9	70.0	Q14
14	255	63	.45	.25	.65	-1.99	.65	-1.84	.68	.61	77.4	69.5	Q15
15	245	63	.72	.25	1.24	1.16	1.22	1.04	.63	.61	66.1	70.6	Q16
16	276	64	-.26	.26	1.16	.96	1.22	1.07	.25	.57	54.0	67.9	Q22
17	304	68	-.99	.26	1.19	1.22	1.22	.97	.35	.50	59.7	68.2	Q23
MEAN	253.9	60.9	.00	.26	1.00	-.15	.98	-.15			67.7	68.7	
P.SD	45.0	6.9	1.13	.01	.28	1.52	.30	1.41			11.4	2.9	

Tabla 5. Estadísticos de diagnóstico del ajuste de los ítems

Paso 4: Ajuste de los sujetos

De igual forma que en el ajuste de los ítems, los valores de los estadísticos infit ZSTD y outfit ZSTD, calculados para cada persona encuestada, deben estar dentro del rango $[-2, 2]$ para indicar que la persona se ajusta al modelo (Linacre, 2012). Aun así, Linacre indica que valores menores que -2 son raramente preocupantes, así que estos valores no se examinarán. En la se pueden ver los estadísticos de diagnóstico de las dos personas que peor se ajustan. Se trata de los individuos 9 y 31. Se ajustó el mismo modelo eliminando estas dos personas y se compararon los resultados. Se realizó una regresión lineal con las medidas estimadas de los ítems y otra con las medidas estimadas de las personas, obteniendo coeficientes de regresión aproximados de 0 y 1 y una correlación de Pearson de 1. Se establece por lo tanto que los individuos 9 y 31 no influyen las estimaciones y no es necesario eliminarlos (Linacre, 2010).

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%
9	58	15	1.93	.51	2.79	3.23	3.02	3.43	A .54	.50	20.0	69.5
31	54	13	2.68	.56	2.46	2.94	2.57	3.00	B .00	.50	30.8	67.7

Tabla 6. Estadísticos de diagnóstico de las personas que peor se ajustan al Modelo de Escala de Clasificación

Paso 5: Unidimensionalidad

Para comprobar que se cumple el supuesto de unidimensionalidad se realiza un Análisis de Componentes Principales con los residuos estandarizados del ajuste. En la se puede ver como el modelo explica un 44.8% de la varianza de los datos. Se estudiará si el 55.2% de la varianza inexplicada por el modelo se puede atribuir a otra dimensión diferente a la identidad profesional. Si ocurriera esto se estaría ante un problema de multidimensionalidad, ya que el Modelo Rasch y el Modelo de Escala de Clasificación solo miden una variable latente y no múltiples como indicaría la multidimensionalidad.

El autovalor asociado a la primera componente – dimensión secundaria – es 2.9551. Un valor de hasta 3 no se considera una nueva dimensión, así que se cumple la unidimensionalidad (Linacre, 2012).

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units			
	Eigenvalue	Observed	Expected
Total raw variance in observations =	30.7821	100.0%	100.0%
Raw variance explained by measures =	13.7821	44.8%	45.1%
Raw variance explained by persons =	8.5593	27.8%	28.0%
Raw Variance explained by items =	5.2228	17.0%	17.1%
Raw unexplained variance (total) =	17.0000	55.2%	100.0%
Unexplned variance in 1st contrast =	2.9551	9.6%	17.4%
Unexplned variance in 2nd contrast =	2.1155	6.9%	12.4%
Unexplned variance in 3rd contrast =	1.7585	5.7%	10.3%
Unexplned variance in 4th contrast =	1.6676	5.4%	9.8%
Unexplned variance in 5th contrast =	1.4097	4.6%	8.3%

Tabla 7. Análisis en Componentes Principales de los residuos estandarizados

Paso 6: Separación y confiabilidad de los ítems y de las personas

El índice de separación de las personas se refiere a la capacidad del cuestionario de discriminar entre diferentes niveles de encuestados según la identidad. Como se puede ver en la , el cuestionario de identidad profesional en optometría tiene un índice de separación de 2.15, es decir, las respuestas a los ítems del cuestionario permiten diferenciar 2 grupos de personas: individuos que se identifican mucho con la profesión de optometría e individuos que se identifican muy poco con la profesión.

El coeficiente de confiabilidad, asociado al índice de separación, es una medida de la capacidad de reproducir la ubicación relativa de las medidas. Así, una alta confiabilidad de las personas significa que existe una alta probabilidad de que las medidas estimadas de las personas que se identifican mucho con la profesión realmente tengan medidas más altas que las personas que se identifican poco con la profesión (Linacre, 2012).

Si se desea una confiabilidad alta de las personas se necesita una muestra de individuos con un amplio rango de grados de identificación y un instrumento con muchos ítems.

El índice de separación de los ítems se refiere a la capacidad de discriminar entre diferentes niveles de exigencia en los ítems. Para el cuestionario de identidad profesional en optometría se tiene un índice de separación de 3.93, así que se los ítems del cuestionario consiguen diferenciar aproximadamente 4 grupos de ítems dependiendo de su nivel de exigencia.

El coeficiente de confiabilidad de los ítems es de 0.94. Si se desea una confiabilidad alta de los ítems se necesita un cuestionario con un amplio rango de dificultad de los ítems — nivel de exigencia — y una muestra grande de encuestados.

Se considera que un cuestionario discrimina bien si el índice de separación de las personas es superior a 2, el coeficiente de confiabilidad de las personas es superior a 0.80, el índice de separación de los ítems es superior a 3 y el coeficiente de confiabilidad de los ítems es superior a 0.90 (Linacre, 2012).

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	60.5	14.6	2.24	.55	.99	-.04	.97	-.09
SEM	1.5	.2	.17	.01	.06	.15	.06	.15
P.SD	12.5	2.0	1.43	.10	.49	1.22	.53	1.23
S.SD	12.6	2.0	1.44	.10	.49	1.22	.54	1.24
MAX.	84.0	17.0	6.56	1.08	2.79	3.23	3.02	3.43
MIN.	27.0	9.0	-1.89	.42	.19	-2.55	.19	-2.67
REAL RMSE	.60	TRUE SD	1.30	SEPARATION	2.15	PERSON	RELIABILITY	.82
MODEL RMSE	.56	TRUE SD	1.32	SEPARATION	2.37	PERSON	RELIABILITY	.85
S.E. OF PERSON MEAN = .17								

Tabla 8. Separabilidad y confiabilidad de las personas

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	253.9	60.9	.00	.26	1.00	-.15	.98	-.15
SEM	11.2	1.7	.28	.00	.07	.38	.07	.35
P.SD	45.0	6.9	1.13	.01	.28	1.52	.30	1.41
S.SD	46.3	7.1	1.16	.01	.29	1.57	.31	1.45
MAX.	321.0	68.0	2.33	.30	1.46	2.13	1.50	2.11
MIN.	120.0	38.0	-2.20	.24	.55	-2.59	.52	-2.53
REAL RMSE	.28	TRUE SD	1.09	SEPARATION	3.93	ITEM	RELIABILITY	.94
MODEL RMSE	.26	TRUE SD	1.10	SEPARATION	4.15	ITEM	RELIABILITY	.95
S.E. OF ITEM MEAN = .28								

Tabla 9. Separación y confiabilidad de los ítems

Paso 7: Independencia local de los ítems

Otro de los supuestos que deben cumplir los datos para que el Modelo de Escala de Clasificación se valide es la independencia local de los ítems. Si los residuales de dos ítems tienen una correlación positiva alta (>0.7) indica dependencia entre los ítems (Linacre, 2012).

CORREL- ATION	ENTRY NUMBER ITE	ENTRY NUMBER ITE
.53	4 Q4	5 Q5
.38	16 Q22	17 Q23
.38	11 Q12	12 Q13
.30	1 Q1	3 Q3
.30	6 Q6	7 Q8
.29	1 Q1	5 Q5
-.44	7 Q8	16 Q22
-.38	5 Q5	16 Q22
-.36	3 Q3	12 Q13
-.35	4 Q4	17 Q23
-.35	3 Q3	13 Q14
-.32	6 Q6	12 Q13
-.32	10 Q11	14 Q15
-.30	7 Q8	11 Q12
-.29	6 Q6	16 Q22
-.28	1 Q1	15 Q16
-.28	11 Q12	14 Q15
-.28	7 Q8	12 Q13
-.27	1 Q1	13 Q14
-.27	4 Q4	12 Q13

Tabla 10. Correlaciones residuales estandarizadas más fuertes

Como se puede observar en la , las correlaciones entre pares de ítems no son mayores de 0.7, por lo que no existen problemas de independencia local.

Paso 8: Funcionamiento diferencial del ítem

El funcionamiento diferencial del ítem indica que un grupo de encuestados está obteniendo un nivel significativamente más alto de identificación que el otro grupo en un ítem. Se quiere contrastar si algún ítem del cuestionario de identidad profesional en optometría tiene un funcionamiento diferencial, ya que constituiría un problema para la validez del cuestionario y del Modelo de Escala de Clasificación.

Se contrastó si existían diferencias en los ítems por el género, la edad, si tenían algún familiar trabajando en el campo de la salud visual o no y por el lugar donde trabajaban. No se encontraron diferencias significativas en ninguno de los ítems.

Paso 9: Objetividad específica

En este último paso se confirmará una de las propiedades del Modelo de Escala de Clasificación: la objetividad específica. Esta propiedad consiste en que las medidas no dependen de las condiciones específicas con las que se obtuvieron. En otras palabras, la diferencia entre dos ítems no debe depender de las personas específicas que se utilicen para cuantificarla (Prieto & Delgado, 2003).

Para contrastar empíricamente esta propiedad se dividirá la muestra aleatoriamente en dos grupos, se estimará para cada uno las medidas de los ítems y se realizará una regresión lineal simple con las medidas de los dos grupos (Prieto & Delgado, 2003). Si se cumple la objetividad específica, deberán salir unos valores cercanos a 1 para la correlación lineal de Pearson y de la pendiente de la recta y un valor cercano a 0 para la ordenada en el origen. Con el cuestionario de identidad profesional en optometría se ha obtenido una correlación de Pearson de 0.9464946, una pendiente de 0.809081 y una ordenada en el origen de 0.001176471, por lo que se verifica la objetividad específica.

4. Conclusiones

En este trabajo fin de grado se abordaron las diferencias entre los dos marcos teóricos que existen para validar cuestionarios: la Teoría Clásica del Test (TCT) y la Teoría de Respuesta al Ítem (TRI). Se decidió ajustar el Modelo de Escala de Clasificación, perteneciente a la TRI, un modelo basado en el Modelo de Rasch porque los ítems del cuestionario son de respuesta categórica. Se estudiaron las características de los dos modelos para comprender cuáles eran sus diferencias y por qué se debía ajustar el Modelo de Escala de Clasificación para validar el cuestionario de identidad profesional en optometría.

Tras el estudio teórico, se evaluó con la muestra piloto compuesta por 71 individuos el cumplimiento de los supuestos del Modelo de Escala de Clasificación: la unidimensionalidad y la independencia local. Además se identificaron los ítems y las personas que no se ajustaban al modelo y se estudió la estructura de las categorías de respuesta. También se comprobó que la confiabilidad del test era alta y que el índice de separación de personas era superior a 2.

Después de todo este proceso, se encontró que se debía eliminar la categoría neutral, quedando una escala Likert de 4 puntos para las respuestas y que se debía eliminar además el ítem 24: “Siento que pongo las decisiones clínicas antes que las necesidades del negocio” para obtener un cuestionario válido para medir la identidad profesional en optometría.

5. Bibliografía

- Bond, T. G., & Fox, C. M. (2007). Applying the rasch model: Fundamental measurement in the human sciences: Second edition. In *Applying the Rasch Model: Fundamental Measurement in the Human Sciences: Second Edition*.
<https://doi.org/10.4324/9781410614575>
- Boone, W. J. (2016). Rasch analysis for instrument development: Why,when,and how? *CBE Life Sciences Education*, 15(4). <https://doi.org/10.1187/cbe.16-04-0148>
- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1).
- Linacre, J. M. (2012). A User's Guide to WINSTEPS MINISTEP: Rasch-Model Computer Programs. In *Winsteps*.
- Muñiz, J., & Hambleton, R. K. (1992). Medio siglo de Teoría de Respuesta a los ítems. *Anuario de Psicología*, 52.
- Prieto, G., & Delgado, A. R. (2003). Análisis de un test mediante el modelo de Rasch. *Psicothema*, 15(1).
- Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, 15(2). <https://doi.org/10.2307/1412107>
- Wright, B. D., & Stone, M. H. (1979). *Best test design: Rasch measurement*. Mesa Press.
- Stanley, J.C. (1971). Reliability. In R. L. Thorndike, Educational measurement (2nd ed., pp. 356-442). Washington, DC: American Council on Education.
- Linacre, J. M. (2010). When to stop removing items and persons in Rasch analysis? *Rasch Measurement Transactions*, 23(4), 1.

Anexo I: cuestionario

The professional identity of a vocational profession in a commercial setting in the UK

Identity

1. I am proud to belong to the eye care profession

*I strongly
agree*

1

2

3

4

5

I strongly disagree

2. I find myself making excuses for belonging to this profession

*I strongly
agree*

1

2

3

4

5

I strongly disagree

3. I identify strongly with my professional group

*I strongly
agree*

1

2

3

4

5

I strongly disagree

4. Being a member of the primary eye care profession is important to me

*I strongly
agree*

1

2

3

4

5

I strongly disagree

5. I see myself as belonging to the eye care professional group

*I strongly
agree*

1

2

3

4

5

I strongly disagree

6. Being part of a team is essential in the provision of eye care services.

*I strongly
agree*

1

2

3

4

5

I strongly disagree

7. From my role within primary eye care, I attend to:

- a. Customers
- b. Clients
- c. Patients
- d. Other, please specify

Engagement

8. I do this job because it provides me with a strong sense of fulfillment

*I strongly
agree*

1

2

3

4

5

I strongly disagree

9. I chose this job because it allows me to reach my life goals (values)

I strongly agree

I strongly disagree

1 2 3 4 5

10. I am in this job because this job fulfils my career plans

I strongly agree

I strongly disagree

1 2 3 4 5

11. I do this job because of the high reputation it brings to me

I strongly agree

I strongly disagree

1 2 3 4 5

12. I am in this job because it affords me a certain standard of living

I strongly agree

I strongly disagree

1 2 3 4 5

13. I consider my salary is appropriate to my training and experience

I strongly agree

I strongly disagree

1 2 3 4 5

14. I enjoy a feeling of job security

I strongly agree

I strongly disagree

1 2 3 4 5

15. I feel my level of responsibility is appropriate to my scope of practice

I strongly agree

I strongly disagree

1 2 3 4 5

16. I have opportunities for career advancement

I strongly agree

I strongly disagree

1 2 3 4 5

Commitment (Only for qualified staff: All those registered with the GOC)

17. My current job/professional title is:

- a. Optometrist
- b. Dispensing Optician
- c. Contact Lenses dispensing optician
- d. Optical Assistant
- e. Practice manager
- f. Locum

18. Please state the number of years since your registration:

19. I am a member of the following organizations:

- a. The College of Optometrists
- b. The Association of Optometrists
- c. The British Association of Dispensing Opticians (ABDO)
- d. Local Optical Committee
- e. BCLA
- f. Other, please specify

20. I engage in continuous education / accredited additional training to:

- a. To update my knowledge and skills
- b. To complete the requirements within the profession
- c. To progress within the profession
- d. To enhance my professional network

21. When engaging in continuous education / accredited additional training, I tend to:

- a. Access online relevant professional journals
- b. Attend CET events organised by the company I work for
- c. Attend CET events within my local area
- d. Attend yearly events / conferences

22. I am comfortable managing clinical decisions against the needs of the business

I strongly agree

I strongly disagree

1

2

3

4

5

23. I feel I can put clinical decisions before business needs

I strongly agree

I strongly disagree

1

2

3

4

5

24. I feel I put business needs before clinical decisions

I strongly agree

I strongly disagree

1

2

3

4

5

Profile

25. What qualifications do you hold? Please select all that apply

- a. Higher Education Certificate
- b. Higher Education Diploma
- c. Bachelor Degree (BSc)
- d. Postgraduate Certificate
- e. Postgraduate Diploma
- f. Master of Optometry (MOptom)
- g. Masters Degree (MSc)
- h. Professional Certificate Medical Retina (Prof Cert Med Ret)
- i. Professional Certificate Glaucoma (Prof Cert Glau)
- j. Professional Higher Certificate Glaucoma
- k. Professional Diploma Glaucoma
- l. Doctor of Philosophy (PhD)
- m. Other

26. Aside from my professional/job title, I would describe myself (or when asked) as:

- a. An optical assistant
- b. An optician
- c. A clinician
- d. An eye care practitioner
- e. A healthcare professional
- f. An eye care specialist
- g. A hospital optometrist /
- h. An optometrist
- i. A dispensing optician
- j. An academic / scientist
- k. A manager
- l. A director/owner of a company
- m. A product/equipment representative
- n. Other.....

27. Please state the number of years since your registration/working in the optical sector:

28. I work ... (PLEASE SELECT ALL THAT APPLY)

- a. In an independent
 - b. in a multiple
 - c. in a hospital
 - d. as a locum
 - e. in academia
- i. If you have multiple employments, please specify where you spend most of your time...
- 1. In an independent
 - 2. in a multiple
 - 3. in a hospital
 - 4. in academia

29. I work...

- a. Full-time
- b. Part-time (1 workplace)
- c. Part-time (more than 1 workplace)
 - i. If you work part-time, how many days a week do you work?
 - 1. 1 day
 - 2. 2 days
 - 3. 3 days
 - 4. 4 days

30. Have you got any family members that also work in Primary Eye Care?

- a. Yes
- b. No

31. The postcode of my workplace is (main workplace)

32. I describe myself as:

- a. Male
- b. Female
- c. Prefer not to say
- d. Other

33. Age

The researchers are keen to explore in depth the key themes identified through this survey. If you wish to take part in a focus group and share your experiences, please provide your details below or contact us at: luisa.simo@port.ac.uk Thank you.

Thank you for taking part in this survey.
Your time and answers are greatly appreciated.

Anexo II: regresión lineal simple para verificar la objetividad específica

se leen los dos ficheros que contienen las medidas estimadas de los ítems con cada división de la muestra

```
load("D:/users/medidas_item_muestra1.Rda")
```

```
load("D:/users/medidas_item_muestra2.Rda")
```

```
datos<-data.frame(y=medidas_item1$MEASURE,x=medidas_item2$MEASURE)
```

```
lm<-lm(y~x,datos)
```

```
lm$coefficients[1] #aproximadamente 0
```

```
lm$coefficients[2] #aproximadamente 1
```

```
cor(datos$x,datos$y,method="pearson") #aproximadamente 1
```