

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE OPTICA Y OPTOMETRÍA**



**TESIS DOCTORAL**

**Adaptación Transcultural al español de la escala**  
**“Convergence Insufficiency Symptom Survey” (CISS)**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Carlos Pérez Garmendia**

Directores

**Beatriz Antona Peñalba**  
**Ana Rosa Barrio de Santos**  
**Mariano González Pérez**

Madrid

© Carlos Pérez Garmendia, 2024

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE ÓPTICA Y OPTOMETRÍA**



**TESIS DOCTORAL**

**Adaptación Transcultural al español de la escala “Convergence  
Insufficiency Symptom Survey” (CISS)**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Carlos Pérez Garmendia**

Directores

Dra. Beatriz Antona Peñalba  
Dra. Ana Rosa Barrio de Santos  
Dr. Mariano González Pérez

**Madrid, 2024**

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE ÓPTICA Y OPTOMETRÍA**



**TESIS DOCTORAL**

PROGRAMA DE DOCTORADO EN ÓPTICA, OPTOMETRÍA Y VISIÓN

**Adaptación Transcultural al español de la escala “Convergence  
Insufficiency Symptom Survey” (CISS)**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Carlos Pérez Garmendia**

Directores

Dra. Beatriz Antona Peñalba  
Dra. Ana Rosa Barrio de Santos  
Dr. Mariano González Pérez

**Madrid, 2024**

## AGRADECIMIENTOS Y DEDICATORIAS

### Agradecimientos

En primer lugar, quiero dar las gracias desde lo más profundo del corazón a mis directoras de tesis, Beatriz y Ana, y a mi director, Mariano. Os agradezco toda vuestra ayuda, vuestros fantásticos consejos, vuestra paciencia infinita, y sobre todo, el cariño y el apoyo que desde el primer momento he sentido. Sin vosotros, no hubiera sido posible llegar hasta aquí.

También quiero dar las gracias a “mis chicas”, a mi mujer, Noemí, y a mi hija María. Vosotras sois el centro de mi vida, mi ilusión, mi equipo. Sin vuestro cariño, no tendría nada, y tampoco tendría sentido nada de lo que hago. Sois mi proyecto de vida. Os quiero.

Tampoco puedo olvidarme de mis padres, Maribel y Carlos. Sois un referente para mí y puedo presumir orgulloso de haber contado siempre con el mejor ejemplo. Vosotros me habéis enseñado casi todo lo que sé y seguís demostrándome cada día lo mucho que me queréis. Vuestra generosidad no tiene límite y la ilusión que ponéis en cualquiera de mis proyectos tampoco. Muchas gracias. Os quiero.

Gracias también al resto de la familia. A los que estáis y a los que no estáis, a mis amigos y a todos los que me queréis. Soy feliz y también mucha de la culpa es vuestra.

### Dedicatorias

Desde pequeño, mis padres siempre me inculcaron el respeto a los demás, la humildad y el valor del trabajo. Cuatro generaciones de doctores en la familia. Ya mi bisabuelo, Nicolás, íntimo amigo de Unamuno, fue Catedrático de Derecho en la Universidad de Salamanca. Después mi abuelo, José, también fue Catedrático de Física y mi padre, Carlos, es también Doctor en Matemáticas. Todos ellos han sido para mí grandes ejemplos y de todos ellos he aprendido a tener siempre espíritu crítico, a querer avanzar y a buscar siempre caminos nuevos.

Si hay una persona a la que me gustaría especialmente dedicar este trabajo, esa es, mi abuelo José Garmendia Iraundegui. Siempre recordaré sus historias con Francisco Morán, Julio Palacios y Miguel Azpiroz, que, junto con él, eran los 4 mosqueteros, o por lo menos así es como se hacían llamar en el círculo de la “Física del Aire”. Sé que a él, es a la persona que más ilusión haría ver a su nieto defender una Tesis. Para ti es este trabajo. Estés donde estés, sé que te encantará. Te quiero.

# ÍNDICE GENERAL

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PARA LA OBTENCIÓN DEL TÍTULO DE DOCTOR.....	iii
AGRADECIMIENTOS Y DEDICATORIAS.....	iv
ÍNDICE GENERAL.....	v
LISTA DE ABREVIATURAS.....	ix
ÍNDICE DE FIGURAS.....	xi
RESUMEN.....	xvi
ABSTRACT.....	xviii
PRESENTACIÓN DE LA ESTRUCTURA DE LA TESIS.....	1
<b>1 INTRODUCCIÓN.....</b>	<b>3</b>
1.1 SISTEMA VERGENCIAL Y ACOMODATIVO.....	3
1.1.1 INTERACCIONES ENTRE ACOMODACIÓN Y VERGENCIA.....	4
1.1.2 PRUEBAS CLÍNICAS PARA DETECTAR DISFUNCIONES VERGENCIALES Y ACOMODATIVAS.....	7
1.1.3 DISFUNCIONES BINOCULARES ACOMODATIVAS Y VERGENCIALES NO ESTRÁBICAS.....	9
1.2 LA INSUFICIENCIA DE CONVERGENCIA Y LA IMPORTANCIA DE LOS INSTRUMENTOS BASADOS EN LAS RESPUESTAS DE LOS SUJETOS (INSTRUMENTOS PRO) PARA SU INVESTIGACIÓN.....	11
1.2.1 LA INSUFICIENCIA DE CONVERGENCIA.....	11
1.2.2 LA IMPORTANCIA DE LOS INSTRUMENTOS “PATIENT REPORTED OUTCOMES” (PRO).....	14
1.2.3 LOS INSTRUMENTOS PRO COMO HERRAMIENTA PARA EL DIAGNÓSTICO DE INSUFICIENCIA DE CONVERGENCIA. LA ESCALA CISS “CONVERGENCE INSUFFICIENCY SYMPTOM SURVEY”.....	15
1.2.4 CRITERIOS DE CALIDAD PARA VALORAR UN INSTRUMENTO PRO.....	16
1.2.5 EL MODELO PROPUESTO POR GEORG RASCH. CONSIDERACIONES GENERALES SOBRE EL ANÁLISIS DE RASCH.....	19
1.2.6 IMPORTANCIA DE LA VALIDACIÓN DE LOS INSTRUMENTOS PRO PARA SU EMPLEO EN EL DIAGNÓSTICO CLÍNICO.....	26
<b>2 JUSTIFICACIÓN DEL ESTUDIO.....</b>	<b>37</b>
<b>3 HIPÓTESIS y OBJETIVOS.....</b>	<b>39</b>
3.1 HIPÓTESIS.....	39
3.1.1 HIPÓTESIS PRINCIPAL.....	39
3.1.2 HIPÓTESIS SECUNDARIAS.....	39
3.2 OBJETIVOS.....	40

3.2.1	OBJETIVO PRINCIPAL .....	40
3.2.2	OBJETIVOS SECUNDARIOS.....	40
3.2.3	OTROS OBJETIVOS ADICIONALES .....	41
<b>4</b>	<b>MÉTODOS</b> .....	<b>43</b>
4.1	ADAPTACIÓN TRANSCULTURAL DE LA ESCALA CISSv15 AL ESPAÑOL.....	44
4.1.1	BLOQUE I – TRADUCCIONES DIRECTAS .....	45
4.1.2	BLOQUE II – TRADUCCIONES INVERSAS.....	47
4.1.3	BLOQUE III – REUNIÓN DE EXPERTOS .....	48
4.1.4	PRE-TEST.....	49
4.2	VALIDACIÓN DE LA ESCALA CISSve.....	50
4.2.1	CONSIDERACIONES GENERALES.....	50
4.2.2	FASE CUALITATIVA. VALIDACIÓN DE CONTENIDO .....	51
4.2.3	FASE CUANTITATIVA. VALIDACIÓN DEL CONSTRUCTO, FIABILIDAD Y VALIDEZ CONVERGENTE Y DIVERGENTE. REPETIBILIDAD.....	53
4.3	COMPARACIÓN ENTRE LAS PROPIEDADES DE CISSv15 Y CISSve .....	64
<b>5</b>	<b>RESULTADOS</b> .....	<b>65</b>
5.1	ADAPTACIÓN TRANSCULTURAL DE LA ESCALA CISSv15 .....	65
5.1.1	BLOQUE I – TRADUCCIONES DIRECTAS .....	65
5.1.2	BLOQUE II – TRADUCCIONES INVERSAS.....	72
5.1.3	BLOQUE III – REUNIÓN DE EXPERTOS .....	75
5.1.4	PRE-TEST.....	79
5.2	VALIDACIÓN DE LA ESCALA CISSve.....	88
5.2.1	CONSIDERACIONES GENERALES.....	88
5.2.2	FASE CUALITATIVA. VALIDACIÓN DE CONTENIDO .....	88
5.2.3	FASE CUANTITATIVA. VALIDACIÓN DEL CONSTRUCTO CON ANALISIS DE RASCH, VALIDEZ CONVERGENTE Y DIVERGENTE Y REPETIBILIDAD .....	90
5.3	COMPARACIÓN ENTRE LAS PROPIEDADES DE CISSv15 Y CISSve .....	143
<b>6</b>	<b>DISCUSIÓN</b> .....	<b>147</b>
6.1	ADAPTACIÓN TRANSCULTURAL .....	147
6.2	PROPIEDADES PSICOMÉTRICAS DE LA ESCALA CISSve.....	148
6.3	LIMITACIONES DEL ESTUDIO .....	154
6.4	PROPUESTA PARA POSIBLES NUEVAS LÍNEAS DE INVESTIGACIÓN.....	155
<b>7</b>	<b>CONCLUSIONES</b> .....	<b>159</b>
7.1	CONCLUSIONES DERIVADAS DE LA ADAPTACIÓN TRANSCULTURAL DE LA ESCALA CISSv15 .....	159
7.2	CONCLUSIONES DERIVADAS DE LA APLICACIÓN DE LA ESCALA CISSve A POBLACIÓN DIAGNOSTICADA CLÍNICAMENTE.....	160

<b>8</b>	<b>APORTACIONES ADICIONALES DERIVADAS DE LA TESIS</b>	<b>163</b>
8.1	REDUCCIÓN DE ITEMS PARA OBTENCIÓN DE ESCALA CISSve ABREVIADA PARA “SCREENINGS”	163
8.1.1	OBJETIVOS	163
8.1.2	JUSTIFICACIÓN	163
8.1.3	HIPÓTESIS	163
8.1.4	MÉTODOS	164
8.1.5	RESULTADOS	164
8.1.6	DISCUSIÓN	187
8.1.7	CONCLUSIONES	188
8.2	CREACIÓN DE ESCALAS PARA MEDIDA DE SÍNTOMAS ASOCIADOS A INSUFICIENCIA DE CONVERGENCIA POR GRUPOS DE EDAD	188
8.2.1	OBJETIVOS	188
8.2.2	JUSTIFICACIÓN	188
8.2.3	HIPÓTESIS	189
8.2.4	MÉTODOS	189
8.2.5	RESULTADOS	189
8.2.6	DISCUSIÓN	191
8.2.7	CONCLUSIONES	191
8.3	CONSTRUCCIÓN DE UN BANCO DE ÍTEMS	191
8.3.1	OBJETIVOS	192
8.3.2	JUSTIFICACIÓN	192
8.3.3	HIPÓTESIS	192
8.3.4	MÉTODOS	192
8.3.5	RESULTADOS	192
8.3.6	DISCUSIÓN	195
8.3.7	CONCLUSIONES	196
8.4	UTILIZACIÓN DE METODOS PREDICTIVOS BASADOS EN ALGORITMOS DE CLASIFICACIÓN DE MACHINE LEARNING PARA REALIZAR CUESTIONARIOS ADAPTATIVOS.	197
8.4.1	OBJETIVOS	197
8.4.2	JUSTIFICACIÓN	197
8.4.3	HIPÓTESIS	197
8.4.4	MÉTODOS	198
8.4.5	RESULTADOS	198
8.4.6	DISCUSIÓN	204
8.4.7	CONCLUSIONES	204
<b>9</b>	<b>COMUNICACIONES CIENTÍFICAS DERIVADAS DE LA TESIS</b>	<b>209</b>

9.1	PUBLICACIONES EN REVISTAS INTERNACIONALES.....	209
9.2	COMUNICACIONES EN CONGRESOS NACIONALES E INTERNACIONALES .....	209
10	REFERENCIAS.....	211
11	ANEXOS .....	227
11.1	ANEXO I – GUÍA DE KHADKA ADAPTADA .....	227
11.2	ANEXO II – INFORME FAVORABLE DEL COMITÉ ÉTICO DE INVESTIGACIÓN CLÍNICA DEL HOSPITAL CLÍNICO SAN CARLOS .....	228
11.3	ANEXO III – MODELOS DE CONSENTIMIENTOS INFORMADO PARA EL TRATAMIENTO DE LOS DATOS .....	229
11.3.1	CONSENTIMIENTO INFORMADO DIRIGIDO A PADRES O TUTORES .....	229
11.3.2	ASENTIMIENTO INFORMADO DIRIGIDO A ADULTOS JÓVENES PARTICIPANTES 230	
11.4	ANEXO IV – TABLAS Y GRÁFICOS AUXILIARES. DESCRIPTIVOS DE LA ESCALA CISSve	231
11.4.1	DESCRIPTIVOS DE LA VARIABLE PUNTUACIÓN EN LA ESCALA CISSve.....	231
11.4.2	DESCRIPTIVOS DE LOS ITEMS DE LA ESCALA CISSve .....	232
11.4.3	COMPORTAMIENTO DE LAS CATEGORÍAS DE LA ESCALA CISSve.....	233
11.4.4	DISCRIMINACIÓN DE LOS ITEMS DE la ESCALA CISSve.....	237
11.4.5	DIFICULTAD DE LOS ÍTEMS DE LA ESCALA CISSve .....	239
11.4.6	INFIT Y OUTFIT DE LOS ITEMS DE LA ESCALA CISSve.....	240
11.4.7	MEDIDAS ESPERADAS vs MEDIDAS OBSERVADAS EN LA ESCALA CISSve .....	241
11.5	ANEXO V – ESCALA CISSv15.....	242
11.5.1	PERMISO DE UTILIZACIÓN DE CISSv15 .....	242
11.5.2	ESCALA CISSv15 .....	243
11.6	ANEXO VI – ESCALA CISSve CON ÍTEMS ADICIONALES.....	244
11.7	ANEXO VII – ESCALA CVSS17 .....	245
11.8	ANEXO VIII – ESCALA WEMWBS.....	248
11.8.1	PERMISO DE UTILIZACIÓN DE WEMWBS (versión española).....	248
11.8.2	ESCALA WEMWBS .....	249
11.9	ANEXO IX – PUBLICACIONES EN REVISTAS INTERNACIONALES .....	250

## LISTA DE ABREVIATURAS

AOA	Asociación Americana de Optometría
ARx	Autorefractometría
AV	Agudeza Visual
BN	Base Nasal
BT	Base temporal
$\beta_i$	Dificultad del ítem i
CCI	Coefficiente de Correlación Intraclase
CEIC	Comité Ético de Investigación Clínica
COR	Coefficiente de Repetibilidad (Coefficient of Repeatability)
CIRS	Convergence Insufficiency and Reading Study
CISS	Convergence Insufficiency Symptom Survey
CISSv15	Convergence Insufficiency Symptom Survey – v15
CISSve	Convergence Insufficiency Symptom Survey – version Española
CISSve-R2	Convergence Insufficiency Symptom Survey – R2 (2 ítems)
CVSS17	Computer-Vision Symptom Scale 17
D	Dioptrías
DIF	Funcionamiento Diferencial de los Ítems (Differential Item Functioning)
DIP	Distancia interpupilar
EA	Exceso acomodativo
EC	Exceso de convergencia
EMG	Electromiografía del Músculo Orbicular
FCF	Frecuencia Crítica de Fusión
IC	Intervalo de Confianza
IA	Insuficiencia acomodativa
ID	Insuficiencia de divergencia
ICC	Curva Característica del ítem
INE	Instituto Nacional de Estadística

IRT	Teoría de Respuesta al Ítem (Item Response Theory)
Nott	Valor en Dioptrías de la Retinoscopía Nott
OD	Ojo Derecho
OI	Ojo Izquierdo
$\Theta_s$	Capacidad de la persona s que responde al ítem
PCA	Análisis de los componentes principales (Principal Component Analysis)
PVD	Pantallas de Visualización de Datos
PPA	Punto Próximo de Acomodación
PPC	Punto Próximo de Convergencia
PRO	Resultados percibidos por el paciente (Patient Reported Outcomes)
RA	Respuesta Acomodativa
Rx	Retinoscopía estática
Rxh	Refracción habitual
RSM	Modelo de Escala de Clasificación (Rating Scale Model)
SOV	Síntomas Oculares y/o Visuales
TCT	Teoría Clásica de Test
TIC	Tecnologías de la Información y las Comunicaciones
TV	Terapia Visual
UGT	Unión General de Trabajadores
VF	Vergencia fusional
VFP	Vergencia fusional positiva
VFN	Vergencia fusional negativa

## ÍNDICE DE FIGURAS

Figura 1.1 - Acomodación Ocular .....	4
Figura 1.2 - Modelo de Hung et al. de los sistemas vergencial y acomodativo .....	6
Figura 1.3 – Elementos característicos de los sistemas vergencial y acomodativo e interrelaciones funcionales entre ambos.....	7
Figura 1.4 – Síntomas asociados a la IC.....	13
Figura 1.5 – Tipos de variables .....	18
Figura 1.6 – Probabilidad de acierto en función de la dificultad del ítem y la habilidad del sujeto .....	24
Figura 1.7 – Proceso de validación de un instrumento documental.....	27
Figura 4.1 – Proceso de adaptación y validación de la escala CISSv15 .....	43
Figura 4.2 – Proceso general de adaptación de la escala CISSv15 .....	44
Figura 4.3– Proceso detallado de adaptación de la escala CISSv15.....	45
Figura 4.4 – BLOQUE I – Traducciones directas .....	46
Figura 4.5 – BLOQUE II – Traducciones inversas .....	47
Figura 4.6 – BLOQUE III – Reunión de expertos .....	48
Figura 4.7 – Fase de pre-test .....	49
Figura 4.8 – Fases en la validación de un instrumento documental.....	50
Figura 4.9 – Cálculo de la muestra necesaria para estudio de validez convergente y divergente de la escala CISSve – Granmo <sup>146</sup> (IMIM).....	55
Figura 4.10 – Diferencia entre Análisis de las componentes principales y Análisis factorial.....	58
Figura 4.11 – Estimación del tamaño del DIF en logits según Mantel-Haenszel <sup>107</sup> (Winsteps) .	60
Figura 4.12 – Criterio de valoración sobre la existencia de diferencia de dificultad entre grupos .....	61
Figura 5.1 – Escala CISSv15 .....	66
Figura 5.2 – Traducción T1 .....	67
Figura 5.3 – Traducción T2 .....	68
Figura 5.4 – Preparativos de la sesión de consenso para obtención de T12 .....	70
Figura 5.5 – Traducción directa consensuada T12 .....	71
Figura 5.6 – Traducciones inversas – BT1 y BT2.....	72
Figura 5.7 – Diferencias entre BT1 y BT2 .....	74
Figura 5.8 – Traducción inversa consensuada (BT12) .....	75
Figura 5.9 – Diferencias entre BT1 y BT2 .....	76
Figura 5.10 – Comparación entre ítems de BT1 Y BT2 realizada por el equipo de investigación	76
Figura 5.11 – Revisión conjunta de todas las traducciones por parte del panel de expertos ....	77
Figura 5.12 – Revisión de la Equivalencia de las traducciones llevada a cabo por el panel de expertos y categorización de los ítems de BT12 según su correspondencia con la versión original.....	78
Figura 5.13 – Descriptivos de la versión preliminar de CISSve.....	80
Figura 5.14 – Descripción de la muestra que participa en pre-test de CISSve y su puntuación .	81
Figura 5.15 – Mapa de Wright del pre-test de la escala CISSve .....	82
Figura 5.16 – Correspondencia entre distancia ITEM-SUJETO en escala logit y probabilidad de éxito para el ítem de un determinado sujeto .....	83
Figura 5.17 – Ajuste al modelo de Rasch de las puntuaciones de los sujetos .....	84

Figura 5.18 – Resumen general de ajuste de ítems y sujetos en la versión preliminar de CISSve .....	85
Figura 5.19 – Distribución de las categorías de respuesta de la escala CISSve.....	85
Figura 5.20 – Distribución gráfica de las categorías de respuesta de la escala CISSve .....	86
Figura 5.21 – Ajuste de cada ítem al modelo de Rasch.....	87
Figura 5.22 – Índice de discriminación del ítem 2 en la escala preliminar CISSve .....	88
Figura 5.23 – Descripción (Winsteps) de la muestra que contestó la escala CISSve .....	91
Figura 5.24 – Descripción (SPSS) de la muestra que contestó la escala CISSve .....	92
Figura 5.25 – Resumen de la estructura de categorías de respuesta de la escala CISSve .....	93
Figura 5.26 – Categorías de respuesta por ítem de la escala CISSve .....	93
Figura 5.27 – Curva de probabilidad del ítem 1 de la escala CISSve .....	94
Figura 5.28 – Estadísticas de ajuste al modelo de los Items .....	95
Figura 5.29 – Análisis de las componentes principales (PCA) de la escala CISSve .....	96
Figura 5.30 – Ítems de la escala CISSve que constituyen la segunda característica encontrada	96
Figura 5.31 – Coeficiente de desatenuación entre el primer y segundo contraste .....	96
Figura 5.32 – Interdependencia entre ítems (análisis PCA con Winsteps) .....	98
Figura 5.33 – Person Separation Index (Winsteps) .....	99
Figura 5.34 – Person Separation Index and Reliability – Extreme and non-extreme (Winsteps)	99
Figura 5.35 -Person Separation Index and Reliability - Non-extreme (winsteps) .....	99
Figura 5.36 – Niveles de síntomas de la escala CISSve estadísticamente diferentes con independencia de la muestra de acuerdo al método de Wright .....	101
Figura 5.37 – Correspondencia de nivel de síntomas en función de la puntuación obtenida en la escala CISSve .....	102
Figura 5.38 – Distribución de la muestra en cuanto a sus niveles de síntomas.....	103
Figura 5.39 – Distribución porcentual de la muestra en los diferentes niveles de síntomas ...	103
Figura 5.40 – Targeting de la escala CISSve para la muestra estudiada .....	104
Figura 5.41 – Targeting de la escala CISSve para la muestra estudiada, descartando los-sujetos que peor se adaptan al modelo (n=430, M=15.5, DT=9,9) .....	104
Figura 5.42– Histograma que representa la distribución de las personas de la muestra analizada frente a los ítems de la escala CISSve .....	105
Figura 5.43 – Mapa de Wright de la escala CISSve - “M” (media) y “S” (Desviación Típica).....	105
Figura 5.44 – Probabilidad de Éxito de un sujeto sobre un ítem en función de la distancia en logits de ese sujeto a ese ítem <sup>107</sup> .....	106
Figura 5.45 – Clasificación de los ítems de la escala CISSve según su nivel de dificultad .....	107
Figura 5.46– Expresión gráfica de los ítems de la escala CISSve según su nivel de dificultad ..	108
Figura 5.47 – Análisis del DIF en cuanto al sexo (Winsteps) .....	109
Figura 5.48 – Representación gráfica del Análisis del DIF en cuanto al sexo (Winsteps) .....	110
Figura 5.49 - Análisis del DIF en cuanto a la edad (Winsteps) .....	112
Figura 5.50 – Representación gráfica del Análisis del DIF en cuanto a la edad (Winsteps).....	113
Figura 5.51 – Análisis del DIF en cuanto al sexo mediante el método de Mantel (Winsteps) ..	115
Figura 5.52 – Análisis del DIF en cuanto a la Edad mediante el método de Mantel (Winsteps) .....	116
Figura 5.53 – Pruebas no paramétricas que confirman las diferencias entre las puntuaciones de los sujetos según el rango de edad al que pertenezcan (9-17 o 18-30 años) .....	117
Figura 5.54 – Prueba KMO y prueba de esfericidad de Bartlett que confirman la viabilidad para realizar un análisis factorial.....	118
Figura 5.55 – Varianza explicada por los 3 factores encontrados en el análisis factorial .....	119
Figura 5.56 – Factores identificados a través del análisis factorial.....	120

Figura 5.57– Diagrama de sedimentación.....	121
Figura 5.58 – Representación gráfica del peso de cada factor .....	121
Figura 5.59 – Contribución de cada factor al total de síntomas percibidos (por grupos de edad) .....	122
Figura 5.60 – Representación gráfica de la contribución de cada factor al total de síntomas percibidos (por grupos de edad) .....	122
Figura 5.61 – Prueba de normalidad de K-S sobre las puntuaciones de las escalas CISSve, CVSS17 y WEMWBS.....	123
Figura 5.62– Test de “Friedman” para la comparación de las puntuaciones de las escalas CISSve, CVSS17 y WEMWBS .....	123
Figura 5.63 – Prueba “U de Mann-Whitney” para las muestras independientes de las escalas CISSve, CVSS17 y WEMWBS .....	124
Figura 5.64 – Pruebas de correlación “rho de Spearman” y “tau de Kendall” entre las puntuaciones de las escalas CISSve, CVSS17 y WEMWBS.....	125
Figura 5.65 – Diagrama de dispersión con recta de regresión entre las puntuaciones de las escalas CISSve y CVSS17 .....	126
Figura 5.66 – Diagrama de dispersión con recta de regresión entre las puntuaciones de las escalas CISSve y WEMWBS.....	126
Figura 5.67 – Diagrama de Bland-Altman sobre la repetibilidad de la escala CISSve .....	127
Figura 5.68– Sesión 1 vs Sesión 2 sobre la misma muestra - escala CISSve.....	128
Figura 5.69 – Comparación entre la Sesión 1 y La Sesión 2 - escala CISSve.....	128
Figura 5.70 – Coeficiente de correlación intraclase - escala CISSve .....	129
Figura 5.71 – Pacientes diagnosticados clínicamente con IC.....	130
Figura 5.72 – Estadísticos descriptivos de la muestra (sintomática/general).....	131
Figura 5.73 – Prueba de Levene y Prueba t para muestras independientes .....	131
Figura 5.74 - Estadísticos Descriptivos CISSve – Factores Población clínica vs general .....	132
Figura 5.75 - Prueba de Levene y T de Student - Clínica vs General .....	132
Figura 5.76 -Estadísticos Descriptivos CISSve agrupados por Sexo.....	133
Figura 5.77 - Pruebas de Levene y t de student – Comparativa en cuanto al sexo .....	133
Figura 5.78 - Estadísticos Descriptivos CISSve agrupados por Grupos de Edad.....	134
Figura 5.79 - Pruebas de Levene y t de student – Comparativa en cuanto a la edad .....	134
Figura 5.80 - Comparativa de Ítems - Medida Análisis de Rasch (Clínica vs General) .....	135
Figura 5.81 – Pacientes diagnosticados clínicamente de IC – PRE y POST terapia .....	135
Figura 5.82 - Estadísticos Descriptivos CISSve – Antes de Terapia.....	136
Figura 5.83 - Estadísticos Descriptivos CISSve – Después de Terapia .....	136
Figura 5.84 –Comparación estadísticos descriptivos Antes de Terapia vs Después de Terapia	137
Figura 5.85 – Comparación estadísticos descriptivos de cada factor, antes y después de terapia .....	137
Figura 5.86 - Prueba de Levene y T de Student factores .....	138
Figura 5.87 – Diagrama de caja y bigote – Variación de sintomatología pre vs post terapia...	139
Figura 5.88 – Prueba de Levene y T de Student - variación de síntomas en cuanto al sexo....	140
Figura 5.89 - Descriptivos factores para los 2 grupos de edad – niños (9-17) y adultos (18-30) .....	140
Figura 5.90 - Prueba de Levene y T de Student - variación de síntomas en cuanto a la edad.	140
Figura 5.91 - Análisis cluster – Respuestas pre y post terapia .....	141
Figura 5.92 – Tipología de síntomas por grupos de edad (9-17, 18-30, >40 años).....	142
Figura 5.93 - Comparación entre las propiedades psicométricas de la escala CISSve y la escala CISSv15 .....	144

Figura 5.94 – Comparación entre los ítems de la escala CISSve y la escala CISSv15 .....	145
Figura 5.95 – Nivel de dificultad de los ítems de la escala CISSve vs la escala CISSv15 .....	145
Figura 8.1 – Correlación biserial puntual de los ítems de la escala CISSve .....	165
Figura 8.2 – Infit y Outfit de los ítems de la escala CISSve .....	166
Figura 8.3 – Dificultad (susceptibilidad a la sintomatología) de los ítems de la escala CISSve	167
Figura 8.4 – Respuestas nulas en los ítems de la escala CISSve .....	168
Figura 8.5 – Porcentaje de ítems de la escala CISSve ubicados en la categoría más alta .....	169
Figura 8.6 – Desviación estándar de los ítems de la escala CISSve .....	170
Figura 8.7 – Asimetría y Curtosis de los ítems de la escala CISSve .....	171
Figura 8.8 – “Targeting” y “Fiabilidad” de la Escala CISSve frente a la escala CISSve-R5.....	172
Figura 8.9 – Análisis de los ítems de la escala CISSve-R5 .....	173
Figura 8.10 – Dimensionalidad de la escala CISSve-R5 .....	174
Figura 8.11 – Factores de la escala CISSve (izquierda) y CISSve-R5 (derecha).....	175
Figura 8.12 - “Targeting” y “Fiabilidad” de la Escala CISSve-R5 (arriba) frente a la escala CISSve-R4 (abajo) .....	176
Figura 8.13 – Análisis de los ítems de la escala CISSve-R4 .....	177
Figura 8.14 - Dimensionalidad de la escala CISSve-R4 .....	178
Figura 8.15 - “Targeting” y “Fiabilidad” de la Escala CISSve-R4 (arriba) frente a la escala CISSve-R3 (abajo) .....	179
Figura 8.16 – Análisis de los ítems de la escala CISSve-R3 .....	180
Figura 8.17 – Dimensionalidad de la escala CISSve-R3 .....	181
Figura 8.18 - “Targeting” y “Fiabilidad” de la Escala CISSve-R3 (arriba) frente a la escala CISSve-R2 (abajo) .....	181
Figura 8.19 – Análisis de los ítems de la escala CISSve-R2 .....	182
Figura 8.20 – Explicación sobre el “Targeting” de la Escala CISSve-R2 frente a la escala CISSve .....	183
Figura 8.21 – Análisis de los ítems de la escala CISSve-R2 en la escala de CISSve .....	184
Figura 8.22 – Invariabilidad de distancia entre los ítems CISSve-R2 en la escala de CISSve.....	185
Figura 8.23 – Dimensionalidad de la escala CISSve-R2 .....	185
Figura 8.24 – Correspondencia entre los puntos obtenidos en la escala CISSve-R2 y su valor en logits una vez se ha aplicado el modelo de Rasch con Winsteps.....	186
Figura 8.25 – Sujetos sintomáticos según la escala CISSve (abajo) y CISSve-R2 (centro) y comparación diferencial entre ambos resultados (arriba) .....	186
Figura 8.26 – Tabla cruzada que relaciona sintomáticos y no sintomáticos según las escalas CISSve y CISSve-R2.....	187
Figura 8.27 Niveles de síntomas de la escala CISSve-R2 estadísticamente diferentes con independencia de la muestra de acuerdo al método de Wright .....	187
Figura 8.28 – Ítems que constituyen los modelos reducidos de escala parametrizados para cada grupo de edad – niños (9-17) – adultos jóvenes (18-30) .....	190
Figura 8.29 – Fiabilidad de CISSve y de las escalas reducidas para cada grupo de edad – niños (9-17) – adultos jóvenes (18-30) .....	190
Figura 8.30 – Características de los 8 nuevos ítems introducidos – comparativa con las características de los 15 ítems pertenecientes a la escala CISSve .....	193
Figura 8.31 – Mapa de Wright que posiciona en la misma escala los 23 ítems (15 de la escala CISSve + 8 adicionales) y los sujetos de la muestra utilizada en este trabajo .....	194
Figura 8.32 – Análisis factorial sobre los 23 ítems (15 de la escala CISSve + 8 adicionales).....	195

Figura 8.33 – Esquema del proceso para la predicción de sujetos “sintomáticos” y “no sintomáticos” en las escalas CISSve y CISSve-R2, siguiendo un algoritmos de clasificación basado en árboles binarios .....	198
Figura 8.34 – Árbol de decisión binario que clasifica a los sujetos como “sintomáticos” o no sintomáticos de acuerdo a los criterios definidos, tomando en cuenta sus respuestas a los ítems de la escala CISSve.....	199
Figura 8.35 - Rendimiento del Árbol de decisión binario aplicado a CISSve.....	200
Figura 8.36 - Árbol de decisión binario que clasifica a los sujetos como “sintomáticos” o no sintomáticos de acuerdo a los criterios definidos para CISSve-R2 (puntuación $\geq$ 4), tomando en cuenta sus respuestas a los ítems de la escala CISSve-R2.....	200
Figura 8.37 - Rendimiento del Árbol de decisión binario aplicado a CISSve-R2 .....	201
Figura 8.38 - Árbol de decisión binario que clasifica a los sujetos como “sintomáticos” o no sintomáticos de acuerdo a los criterios definidos para CISSve, tomando en cuenta sus respuestas a los ítems de la escala CISSve-R2.....	201
Figura 8.39 - Rendimiento del Árbol de decisión binario aplicado a CISSve-R2 con los criterios de clasificación como sintomático definidos para CISSve.....	202
Figura 8.40 - Rendimiento del Árbol de decisión binario construido con las variables que aglutinan los 3 factores componentes de CISSve (fatiga visual, problemas de visión binocular y capacidad de concentración) .....	202
Figura 8.41 - Árbol de decisión binario que clasifica a los sujetos como “sintomáticos” o “no sintomáticos” en la escala CISSve, de acuerdo a la tipología de síntomas reportados .....	203

## RESUMEN

**INTRODUCCIÓN:** La insuficiencia de convergencia es la disfunción de la visión binocular más ampliamente estudiada. Se reportan síntomas asociados a ella con diversas etiologías. En los adultos, los más usuales son, fatiga visual, dolores de cabeza y visión doble, y en niños, diplopia y visión borrosa. El uso de instrumentos que recojan las respuestas percibidas por los pacientes (PRO) facilita la recopilación de datos de forma estandarizada para realizar inferencias sobre las poblaciones de interés. Un instrumento PRO de calidad debe tener dos características fundamentales: Evaluar un único constructo y hacerlo sobre una escala de intervalo. Hasta este momento el cuestionario Convergence Insufficiency Symptom Survey (CISSv15), es el instrumento más extendido para medir insuficiencia de convergencia; no existiendo ningún instrumento validado en lengua española para cubrir ese propósito. El análisis de Rasch, englobado dentro de la teoría de respuesta al ítem (IRT), ofrece un modelo matemático capaz de convertir las observaciones registradas mediante cuestionarios en medidas lineales y además, permite verificar si esa transformación presenta la calidad necesaria.

**OBJETIVO:** Realizar una adaptación transcultural de la escala CISSv15 que permita utilizar el instrumento en poblaciones hispanohablantes, comprobando con el análisis de Rasch que sus propiedades psicométricas son equivalentes a las de la versión original. Evaluar su validez y repetibilidad.

**MÉTODOS:** Se realizó la adaptación del instrumento siguiendo las recomendaciones de Gjersing, Caplehorn y Clausen. Tras la validación del contenido, una muestra (n=458, M=16.02, DT=5.68 años) de sujetos completaron el cuestionario CISSv15 y se evaluaron sus respuestas utilizando el modelo de Rasch y métodos estadísticos convencionales con el objeto de describir sus propiedades psicométricas. Validez y fiabilidad fueron analizadas mediante la evaluación del ajuste de los ítems de la escala al modelo de Rasch: el “análisis de las componentes principales” (PCA) y análisis factorial, el índice de separación para los sujetos y la confiabilidad del instrumento, el “funcionamiento diferencial de los ítems” (DIF) y el ajuste entre la dificultad de los ítems y la habilidad de los sujetos. Para el estudio de la validez convergente y divergente de la versión española de la escala CISSve se comprobó la correlación entre las puntuaciones obtenidas al administrar este instrumento y las obtenidas por los mismos sujetos con las escalas CVSS17 y WEMWBS respectivamente. La fiabilidad test-retest se estimó a partir del cálculo del coeficiente de repetibilidad (COR) y del coeficiente de correlación intraclass (ICC) con un intervalo de confianza al 95%. Por último, se administró la herramienta original CISSv15 a una población anglosajona de 218 sujetos y características demográficas similares para poder comparar los resultados obtenidos por ambas escalas

**RESULTADOS:** La escala CISSve, ha demostrado poseer una buena repetibilidad con un COR=9.14 y un ICC=0.874 Intervalo de confianza al 95% (0,840; 0,902), consistencia interna (infit y outfit dentro del intervalo [-0.5,1.5]), y propiedades psicométricas semejantes a las de la versión original (el autovalor del primer contraste del análisis PCA era 2.2, la separación para los sujetos era 2.22 y no había DIF en cuanto al sexo). Se han definido 6 estratos de sintomatología diferentes y se ha encontrado que el instrumento, pese a formar un constructo unidimensional, presenta 3 atributos que se corresponden con las diferentes etiologías de los síntomas (fatiga visual, disfunción visual, falta de concentración en actividades de cerca). Se comprobó analizando una muestra de población clínica (n=49, M=19,80, DT=9.13) que el

atributo asociado a la falta de concentración no resultaba ser tan relevante para el diagnóstico, al no observarse diferencias significativas entre la puntuación reportada en los ítems referidos a ese factor por la población diagnosticada clínicamente y la general, ni tampoco en una muestra de pacientes diagnosticados clínicamente que se sometieron a terapia visual (n=24, M=17.29, DT=6.74 años), antes y después de terapia. El estudio de la validez concurrente reveló una correlación significativa con la escala destinada a evaluar síntomas visuales relacionados con el uso de videoterminals en el trabajo CVSS17 ( $\rho=0.311$ ,  $p=0.02$ ) También se ha comprobado que la edad es un factor que influye en la percepción de los síntomas asociados a la IC.

**CONCLUSIONES:** Se ha creado la escala CISSve el primer instrumento PRO en lengua española destinado a medir síntomas asociados a IC. La escala CISSve proporciona medidas válidas y fiables y posee propiedades psicométricas similares a la escala original CISSv15 tal y como se ha podido comprobar a través del análisis de Rasch. La escala CISSve a pesar de medir un único constructo (la IC), está compuesto por varios factores que guardan correlación, un primer factor asociado con los síntomas relacionados con la fatiga visual tras actividades prolongadas relacionadas con la lectura o el trabajo de cerca. Un segundo factor asociado con síntomas asociados a la disfunción de la visión binocular y por último un tercer factor relacionado con los problemas de concentración para la lectura o actividades relacionadas con el trabajo de cerca, resultando este último menos significativo para el diagnóstico. Los diferentes grupos de edad, niños, adultos jóvenes y presbíteros, perciben distinta proporción de síntomas asociada a cada factor. A medida que aumenta la edad, crece la percepción de síntomas asociados a la fatiga visual tras actividades prolongadas relacionadas con la lectura o el trabajo de cerca. Las puntuaciones obtenidas en la escala por adultos jóvenes (18-30), son en promedio 5 puntos superiores a las obtenidas por niños (9-17). La escala muestra seis niveles de sintomatología estadísticamente diferenciables, lo que simplifica la interpretación clínica la puntuación alcanzada en el CISSve, posibilitando el discernir el grado de severidad de los síntomas en muestras sintomáticas. Los resultados de los análisis sobre validez convergente con el CVSS17 y validez divergente con el WEMWBS contribuyen a reforzar que esta nueva escala proporciona medidas válidas y mide adecuadamente el constructo que pretende medir. Las medidas obtenidas con la escala CISSve, presentan una buena repetibilidad ya que en este estudio el coeficiente de repetibilidad (COR) fue 9.14 y el coeficiente de correlación intraclase (ICC) fue 0.874.

## ABSTRACT

**INTRODUCTION:** Convergence insufficiency is the most widely studied dysfunction of binocular vision. Symptoms associated with it are reported with various etiologies. In adults, the most common symptoms are visual fatigue, headaches, and double vision, while children may experience diplopia and blurry vision. The use of Patient-Reported Outcome (PRO) instruments facilitates the standardized collection of data to make inferences about populations of interest. A quality PRO instrument should have two fundamental characteristics: assessing a single construct and doing so on an interval scale. Up until now, the Convergence Insufficiency Symptom Survey (CISSv15) has been the most widely used instrument to measure convergence insufficiency.. However, there is no validated instrument in the Spanish language to fulfill this purpose. Rasch analysis, encompassed within item response theory (IRT), offers a mathematical model capable of converting questionnaire responses into linear measures and allows to verify the required quality of this transformation.

**OBJECTIVE:** To perform a cross-cultural adaptation of the CISSv15 scale that enables the use of the instrument in Spanish-speaking populations, verifying through Rasch analysis that its psychometric properties are equivalent to those of the original version. To evaluate its validity and reliability.

**METHODS:** The adaptation of the instrument was conducted following the recommendations of Gjersing, Caplehorn, and Clausen. After content validation, a sample (n=458, M=16.02, SD=5.68 years) of subjects completed the CISSv15 questionnaire, and their responses were evaluated using Rasch modeling and conventional statistical methods to describe its psychometric properties. Validity and reliability were analyzed through the evaluation of item fit to the Rasch model, principal component analysis (PCA), factor analysis, person separation index (PSI), instrument reliability, differential item functioning (DIF), and the match between item difficulty and subject ability. To study the convergent and divergent validity of the CISSv15 scale, correlations were examined between scores obtained from this instrument and scores obtained from the CVSS17 and WEMWBS scales, respectively. Test-retest reliability was estimated by calculating the coefficient of repeatability (COR) and the intra-class correlation coefficient (ICC) with a 95% confidence interval. Finally, the original CISSv15 tool was administered to an Anglo-Saxon population of 218 subjects with similar demographic characteristics to compare the results obtained by both scales.

**RESULTS:** The CISSve scale has demonstrated good test-retest reliability with a COR=9.14 and an ICC=0.874, 95% confidence interval (0.840, 0.902). It also exhibits internal consistency (infit and outfit within the range [-0.5, 1.5]) and psychometric properties similar to the original version (the eigenvalue of the first contrast in the PCA was 2.2, subject separation index (PSI) was 2.22, and no DIF was observed regarding gender). Six different strata of symptomatology have been defined, and the instrument, despite measuring a unidimensional construct (convergence insufficiency), has three attributes corresponding to different etiologies of the symptoms: visual fatigue, visual dysfunction, and lack of concentration during near activities. Analyzing a clinical population sample (n=49, M=19.80, SD=9.13), it was found that the attribute associated with lack of concentration was not as relevant for diagnosis, as no significant differences were observed between the scores reported for items related to that factor by clinically diagnosed individuals and the general population, nor in a sample of clinically diagnosed patients undergoing vision therapy (n=24, M=17.29, SD=6.74 years), before

and after therapy. Concurrent validity analysis showed a significant correlation between the CISSve scale and the CVSS17 scale measuring visual symptoms associated with the use of video terminals at work ( $p=0.311$ ,  $p=0.02$ ). It was also found that age is a factor that influences the perception of symptoms associated with convergence insufficiency.

**CONCLUSIONS:** The CISSve scale has been developed as the first Spanish language PRO instrument designed to measure symptoms associated with CI. The CISSve scale provides valid and reliable measurements and exhibits psychometric properties similar to the original CISSv15 scale, as verified through Rasch analysis. Although the CISSve scale measures a single construct (CI), it is comprised of several factors that are correlated. The first factor is associated with symptoms related to visual fatigue after prolonged activities involving reading or close-up work. The second factor is associated with symptoms related to binocular vision dysfunction, and finally, a third factor is related to concentration problems when reading or carrying out close-up work, this last factor being less significant for diagnosis. Different age groups, including children, young adults, and presbyopes, perceive different proportions of symptoms associated with each factor. As age increases, there is a growing perception of symptoms associated with visual fatigue after prolonged activities related to reading or close-up work. The scores obtained on the scale by young adults (18-30) are, on average, 5 points higher than those obtained by children (9-17). The scale presents six statistically distinct symptomatology levels, which facilitates the clinical interpretation of the score obtained on the CISSve, enabling the discernment of the severity of symptoms in symptomatic samples. The results of the convergent validity analyses with CVSS17 and divergent validity with the WEMWBS contribute to reinforcing that this new scale provides valid measurements and accurately measures the construct it intends to measure. The measurements obtained with the CISSve scale show good repeatability, as in this study the repeatability coefficient (COR) was 9.14 and the intraclass correlation coefficient (ICC) was 0.874.



## PRESENTACIÓN DE LA ESTRUCTURA DE LA TESIS

La estructura del documento, sigue el formato clásico de las publicaciones científicas, es decir: Introducción, Métodos, Resultados y Discusión<sup>1,2</sup>.

La **introducción** expone los fundamentos teóricos sobre los que se sustenta el presente estudio. Se revisa cómo funciona el sistema vergencial y acomodativo y qué disfunciones de visión binocular de tipo acomodativo y vergencial son las más comunes. Se contextualiza en qué consiste la insuficiencia de convergencia y cuáles son métodos que existen para su detección. Se fundamenta el importante papel de los instrumentos basados en las respuestas de los pacientes ("PRO instruments") y se introducen los utilizados hasta ahora para investigar los síntomas oculares y/o visuales (SOV) asociados a la insuficiencia de convergencia. Se presenta el marco teórico sobre el que se sustenta nuestro trabajo, perfilando las recomendaciones más recientes sobre la construcción de instrumentos PRO<sup>3</sup> y se explica en qué consiste la adaptación transcultural de escalas, cuál es la metodología utilizada y cómo se validan este tipo de herramientas <sup>4</sup>. Se introduce el análisis de Rasch, justificando las razones por las que se utiliza tanto en la fase de adaptación como para su posterior validación<sup>5-7</sup>

La sección de **métodos** detalla los procedimientos empleados para la adaptación transcultural de la escala CISS (Convergence Insufficiency Symptom Survey) al español en sus diferentes fases. Se describe la población sobre la que se realiza el estudio y cómo se seleccionan las muestras. Se detalla la forma en la que se recogen los datos y el proceso que se sigue para analizar el comportamiento de la escala, valorar su funcionamiento y las características psicométricas que posee. Se explican los fundamentos de análisis de la validez de la escala (grado en el que escala mide con precisión lo que pretende evaluar) y el procedimiento empleado para el análisis de la repetibilidad del nuevo instrumento.

Los **resultados** presentan una descripción de las propiedades psicométricas de la escala y los ítems que la componen. Se realiza una valoración sistemática y objetiva de las mismas, así como del análisis de la validez y del estudio de la repetibilidad. Otra aportación que aparece en esta sección es la comparación estructurada entre la versión adaptada al español y la fuente original en inglés, describiendo objetivamente el comportamiento que se observa y las similitudes y diferencias que existen entre cada una de las dos versiones.

La **discusión** analiza cómo los resultados obtenidos permiten concluir que la nueva escala reúne los requisitos para considerarla válida y fiable. En este apartado subrayamos las principales aportaciones del trabajo y las limitaciones del estudio.

Tras la discusión se sintetizan en una nueva sección las principales **conclusiones** de esta Tesis y se proponen las líneas de investigación que podrían dar continuidad a este trabajo.

Al final del documento se encuentran varios **anexos** con el material empleado para realizar la investigación, tales como algunos de los documentos auxiliares utilizados para la adaptación, minutas de reuniones, cuestionarios usados en las diferentes fases, etc.

A lo largo del escrito, podemos encontrar referenciada nuestra escala como "cuestionario". Según J. Ventura-León <sup>8</sup> y J. Supo <sup>9</sup> formalmente, un cuestionario es un tipo de instrumento documental que se diferencia de una escala en algunos matices. El primero consta de un conjunto de ítems con categorías de respuesta del tipo "sí"/"no", "afirmativo"/"negativo", mientras que las escalas se construyen con categorías de respuesta organizadas de forma

## PRESENTACIÓN DE LA ESTRUCTURA DE LA TESIS

ordinal y gradada. Habría un tercer tipo de instrumentos documentales, los inventarios. Estos últimos presentan categorías de respuesta de tipo categórico ordenadas <sup>8</sup>.

Cuando se referencie el término “cuestionario”, no haremos distinción y nos referiremos al conjunto de ítems que responden los sujetos, con categoría de respuesta ordinal y gradada, es decir lo que formalmente se considera “escala” o “instrumento PRO” cuando se presenta junto con sus instrucciones y la tabla para asignar las respuestas.

# INTRODUCCIÓN



# 1 INTRODUCCIÓN

## 1.1 SISTEMA VERGENCIAL Y ACOMODATIVO

La visión binocular se define como la percepción simultánea con ambos ojos, y se manifiesta cuando la atención visual se centra en un determinado objeto. Esta visión se origina gracias a la perfecta coordinación motora de los ojos y a la fusión sensorial de las imágenes captadas por cada ojo en una dando lugar a una imagen única<sup>10</sup>.

Un alineamiento adecuado de los ojos es crucial para garantizar una correcta función sensorial, permitiendo la unificación de las imágenes captadas por ambos ojos. Los músculos extraoculares desempeñan un papel esencial en dicho alineamiento, ya que posibilitan la correcta orientación de las fóveas oculares al fijar el objeto de interés. Los movimientos oculares responsables de este ajuste se denominan movimientos vergenciales y pueden clasificarse en:

- **Vergencia Tónica:** Se refiere a la posición ocular en ausencia de disparidad, borrosidad de la imagen o estímulos próximos<sup>11</sup>.
- **Vergencia Proximal:** Surge debido al reconocimiento de la cercanía de un objeto cercano<sup>12</sup>.
- **Vergencia Acomodativa:** Inducida por cambios en la acomodación provocados por estímulos que generan borrosidad<sup>10</sup>.
- **Vergencia Fusional:** Evalúa la aptitud del sistema visual para mantener la fusión mientras se cambia la demanda de vergencia, manteniendo constante el estímulo de acomodación<sup>11</sup>.

En particular, las vergencias fusionales previenen la diplopía, evitando la proyección de objetos en áreas retinianas no correspondientes. En una visión normal, al centrar la vista en un objeto, los movimientos de vergencia actúan para prevenir la visión doble. No obstante, cuando esta fusión se interrumpe, los ejes visuales adoptan una posición denominada forias o heteroforia<sup>13</sup>. Las heteroforias se subdividen según la dirección de la desviación de los ejes visuales: ortoforia (sin desviación), exoforia (desviación hacia temporal), endoforia (desviación hacia nasal) e hiperforia (desviación vertical). Para mantener una adecuada visión binocular, es imperativo utilizar vergencias fusionales para neutralizar las heteroforias y, consecuentemente, evitar la diplopía.

Por otro lado, la acomodación se define como la habilidad del ojo para modificar su capacidad refractiva con el propósito de lograr una imagen retiniana clara de objetos situados a distintas distancias<sup>12</sup>. Esta facultad no solo implica ver con claridad objetos cercanos, sino que representa un cambio dinámico y activo en la potencia refractiva del ojo<sup>14</sup>. Durante la acomodación, el músculo ciliar se contrae, aumentando la potencia dióptrica del cristalino al aumentar su convexidad<sup>15</sup>. Además de la modificación dióptrica, el impulso de acomodación implica convergencia ocular y miosis. Estos tres mecanismos conforman la triada de la visión de cerca.

La acomodación puede categorizarse en:

## INTRODUCCIÓN

- **Acomodación Refleja:** Ajuste autónomo del estado refractivo para mantener una imagen nítida en la retina desencadenada por la presencia de imágenes borrosas<sup>16,17</sup>.
- **Acomodación de Vergencia:** Acomodación inducida por el movimiento de vergencia<sup>11</sup>.
- **Acomodación Proximal:** Generada por la percepción real o aparente de la proximidad de un objeto<sup>17</sup>.
- **Acomodación Tónica:** Acomodación en ausencia de estímulos, determinada por el equilibrio entre las inervaciones simpática y parasimpática<sup>18,19</sup>.

La Figura 1.1 muestra esquemáticamente cómo el sistema visual es capaz de adaptarse a los cambios en la fijación a diferentes distancias.

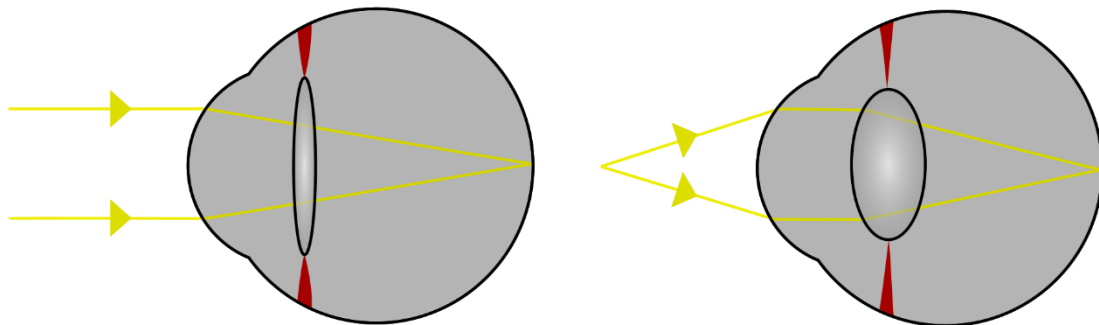


FIGURA 1.1 - ACOMODACIÓN OCULAR

### 1.1.1 INTERACCIONES ENTRE ACOMODACIÓN Y VERGENCIA

Para tener imagen nítida y simple de un objeto, es esencial la correcta coordinación entre los sistemas acomodativo y vergencial. Cualquier anomalía en un sistema puede impactar en el otro significativamente.

Esta relación se ilustra en el modelo propuesto por Hung et al.<sup>17,20</sup>, mostrado en la Figura 1.2. Ambos sistemas, acomodativo y vergencial, son semejantes. Pueden operar en un circuito cerrado o abierto. El circuito cerrado se refiere a la visión binocular, mientras que un circuito abierto se logra eliminando estímulos, como en la oscuridad o cubriendo un ojo.

En el sistema acomodativo, en condiciones de visión binocular normal (circuito cerrado), la borrosidad activa un cambio en la respuesta acomodativa. Se calcula un error inicial restando la respuesta acomodativa existente del estímulo acomodativo existente antes del cambio de borrosidad. Si este error supera un umbral, indicará que se necesita una respuesta acomodativa para reducir la borrosidad. Dos elementos entran en juego: el controlador de acomodación que da una respuesta inicial rápida y el subsistema adaptativo que se encarga después, adaptándose lentamente para mantener el enfoque.

La proximidad de un objeto desencadena la acomodación proximal, especialmente vital en condiciones de circuito abierto.

En este punto, los sistemas vergencial y acomodativo intercambian señales. Desde el circuito acomodativo, se envía una señal de vergencia que representa la relación  $AC/A^{21}$ . El sistema vergencial también envía una señal acomodativa relacionada con la relación  $CA/C^{22}$ .

La acomodación tónica que corresponde a la inervación neurológica básica, opera principalmente cuando no hay un estímulo acomodativo adecuado y el circuito está abierto. Es más influyente en la visión distante, ya que el controlador de acomodación está inactivo.

Todas las señales neurales se combinan para producir la señal nerviosa adecuada, garantizando que el músculo ciliar actúe en consecuencia. Este sistema se autoajusta constantemente, siempre permitiendo una ligera borrosidad, conocida como el retraso acomodativo<sup>17</sup>.

El circuito vergencial funciona de manera similar al acomodativo, pero el umbral está marcado por el área fusional de Panum. Aquí, el disparador principal es la disparidad retiniana. El controlador de vergencia responde rápidamente y, después de algunos segundos, la vergencia lenta mantiene esta respuesta. Dependiendo de la efectividad del componente adaptativo, esta vergencia puede durar minutos, horas o incluso días. La vergencia proximal juega un papel importante en la visión disociada pero menos en la visión binocular. Este sistema también intercambia información con el sistema acomodativo<sup>17</sup>.

La convergencia tónica influye en la respuesta vergencial principalmente en condiciones de circuito abierto, pero su papel es prácticamente despreciable con circuito cerrado.

La señal nerviosa aquí asegura que los músculos extraoculares actúen para producir la respuesta vergencial. Al igual que el sistema acomodativo, se autoajusta, dejando siempre un ligero error de vergencia conocido como disparidad de fijación<sup>23</sup>.

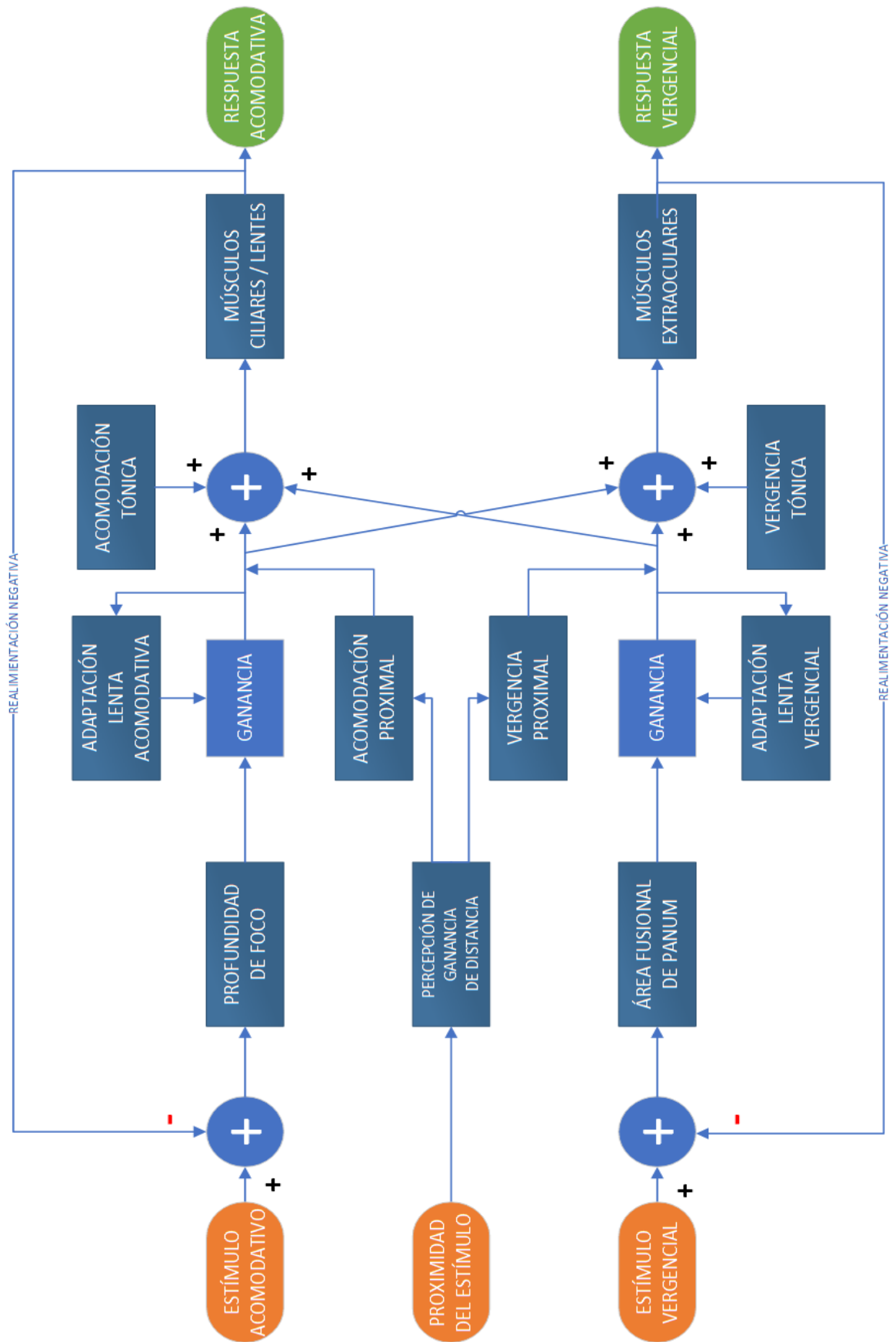


FIGURA 1.2 - MODELO DE HUNG ET AL. DE LOS SISTEMAS VERGENCIAL Y ACOMODATIVO

### 1.1.2 PRUEBAS CLÍNICAS PARA DETECTAR DISFUNCIONES VERGENCIALES Y ACOMODATIVAS

La funcionalidad sinérgica de los sistemas vergencial y acomodativo es crucial para garantizar una visión binocular armónica. Estos sistemas, aunque tienen características distintas, están inevitablemente interconectados, manifestando interacciones que pueden repercutir en la visión binocular. El objetivo de este apartado es introducir cuáles son las principales pruebas clínicas que evalúan estos sistemas, subrayando la relevancia de su diagnóstico en la práctica clínica.

En la Figura 1.3 podemos ver las diferentes características de cada sistema y cuáles son las que relacionan a ambos



FIGURA 1.3 – ELEMENTOS CARACTERÍSTICOS DE LOS SISTEMAS VERGENCIAL Y ACOMODATIVO E INTERRELACIONES FUNCIONALES ENTRE AMBOS

Comenzando con la **Amplitud de Acomodación (AA)**, es la medida dióptrica que se extiende desde el punto más cercano hasta el más lejano de acomodación del ojo. Se puede evaluar mediante técnicas como el acercamiento, alejamiento y el uso de lentes negativas, las cuales proporcionan una evaluación precisa del rango acomodativo<sup>13,24</sup>.

Otra variable de la acomodación es la **Flexibilidad Acomodativa**. Ésta se refiere a la destreza del sistema acomodativo para alternar rápidamente su nivel de enfoque. Las pruebas clínicas, tales como la alternancia de dos estímulos situados respectivamente en lejos y en visión próxima o de lentes de diferente potencia, permiten cuantificar la habilidad del sistema

## INTRODUCCIÓN

acomodativo para realizar cambios rápidos de acomodación y comprobar la resistencia a la fatiga en un tiempo determinado <sup>11,24</sup>

Para obtener el perfil de comportamiento de la respuesta del sistema ante estímulos en visión próxima se mide la **Respuesta Acomodativa (RA)**, que se examina mediante pruebas subjetivas como los Cilindros Cruzados Fusionados (CCF) u objetivas (retinoscopías dinámicas), como el método de estimación monocular (retinoscopia MEM) y la retinoscopia Nott <sup>11</sup>. Estas pruebas miden el plano de enfoque del sujeto con respecto al estímulo acomodativo, permitiendo identificar posibles discrepancias.

La **Acomodación Relativa (AR)** nos revela cuánta acomodación puede ser estimulada (Acomodación Relativa Positiva, ARP) o relajada (Acomodación Relativa Negativa, ARN) manteniendo una convergencia constante. Para su determinación se utilizan lentes negativas (para la ARP) y lentes positivas (para la ARN) dispuestas binocularmente, e indirectamente permiten estudiar parámetros pertenecientes al sistema vergencial <sup>11,24-26</sup>.

En cuanto a las **Desviaciones Oculares (DO)**, es necesario medir la dirección y magnitud de dichas desviaciones para comprender la orientación de los ejes visuales. Herramientas como el Cover Test <sup>24,27</sup> para evaluar de forma objetiva la desviación latente o manifiesta y el test de Wesson para evaluar la disparidad de fijación son algunos ejemplos de las numerosas pruebas clínicas disponibles <sup>28</sup>.

Un aspecto esencial de la visión binocular es la capacidad de mantener una imagen única a pesar de las desviaciones oculares leves que pueda tener el paciente, es decir las **Reservas Fusionales (RF)** disponibles. La Vergencia Fusional Positiva (VFP) y la Vergencia Fusional Negativa (VFN), se miden mediante la introducción de prismas base temporal (BT) y base nasal (BN) respectivamente y evaluando la capacidad de un individuo para compensar la demanda vergencial inducida. También es importante evaluar el Punto Próximo de Convergencia que evalúa la máxima capacidad para converger de un sujeto, es decir, el punto de rotura de la fusión y el posterior recobro de la visión binocular <sup>11,24,26</sup>.

La **Flexibilidad de Vergencias (FV)** se refiere a la capacidad de alternar rápidamente entre distintos niveles de vergencia. Esta puede ser valorada mediante una demanda prismática BN, BT, alterna durante un intervalo de tiempo.

Para comprender las interrelaciones entre acomodación y vergencia, los cocientes **AC/A** y **CA/C** se erigen como herramientas esenciales <sup>11,13,16</sup>. El método de la heteroforia y el método del gradiente son los más utilizados clínicamente y permiten evaluar el cociente AC/A.

Es importante también evaluar la **Fusión** y el tamaño del escotoma de supresión en caso de existir <sup>11</sup>. El test de Worth, el estudio de la fusión con filtro rojo y la prueba de fusión con prisma de  $4\Delta$  BT son pruebas que se utilizan habitualmente en la práctica clínica.

Por último, la **Estereopsis** es esencial en la visión binocular, derivada de la disparidad de imágenes percibidas por la separación horizontal de los ojos, posibilita la percepción de profundidad relativa <sup>11</sup>. Una estereoagudeza óptima requiere buena agudeza visual en ambos ojos, alineación ocular adecuada y experiencia previa de binocularidad <sup>29,30</sup>. Por lo tanto, las pruebas de estereoagudeza se emplean en la práctica clínica como herramientas de detección para identificar disfunciones binoculares, como la heterotropía y la ambliopía <sup>17,31,32</sup>.

Existen diversas pruebas clínicas para evaluar la estereoagudeza, categorizadas en: estereogramas locales o test de contornos, que introducen disparidad horizontal en los límites

de imágenes; estereogramas de puntos al azar o test de estereopsis global, basados en el desplazamiento horizontal de elementos aleatorios; y test de profundidad real, donde la profundidad se logra mediante separación real, eliminando la necesidad de filtros disociantes. Aunque el test de Titmus ha sido tradicionalmente el más utilizado, la introducción de estereogramas de puntos al azar ha llevado al desarrollo de nuevas pruebas, como Random-dot-E y TNO<sup>27,33</sup>

### 1.1.3 DISFUNCIONES BINOCULARES ACOMODATIVAS Y VERGENCIALES NO ESTRÁBICAS

Las disfunciones acomodativas y binoculares no estrábicas son trastornos del sistema óculo-visual que perturban la binocularidad y a la eficiencia de dicho sistema, particularmente en actividades en cerca. Estas disfunciones emergen cuando el sistema visual no logra proporcionar una respuesta acomodativa o vergencial adecuada. Como resultado del esfuerzo continuo al enfrentar tareas en visión próxima, los sistemas visuales afectados por estas alteraciones pueden experimentar una disminución en su eficacia, lo que podría obstaculizar el desempeño normal de actividades <sup>11</sup>.

#### 1.1.3.1 DISFUNCIONES ACOMODATIVAS

Las disfunciones acomodativas se refieren a una inapropiada respuesta del sistema visual frente a demandas visuales específicas, resultando en una deficiencia en el enfoque. Estos trastornos, frecuentemente observados en la práctica optométrica<sup>34</sup>, se manifiestan con síntomas como visión borrosa, dificultades en la lectura, astenopia y dolores de cabeza, especialmente hacia el final del día<sup>11,23</sup>. En cuanto a la prevalencia existe controversia en cuanto a su cuantificación<sup>11,35,36</sup>, debido a las variaciones en los criterios diagnósticos empleados por diversos autores.

Dentro de estas disfunciones, se destacan tres tipos principales según la clasificación de Donders:

- **Insuficiencia de Acomodación (IA):** Esta anomalía sensorial se caracteriza por la incapacidad de enfocar o mantener el enfoque no atribuible a escleriosis del cristalino u otras patologías oculares<sup>11,37</sup>. La IA es originada funcionalmente y sus síntomas se asocian primordialmente a la visión cercana. Se pueden observar problemas en pruebas con lentes negativas, reflejando dificultades para estimular la acomodación<sup>38-40</sup>.
- **Exceso Acomodativo (EA):** Esta disfunción implica problemas al relajar la acomodación. De origen funcional, se manifiesta con síntomas como astenopia, visión borrosa intermitente, y fatiga visual, especialmente durante tareas de cerca<sup>38,41</sup>. En la exploración clínica, se observan dificultades con lentes positivas y una serie de variaciones en indicadores específicos de acomodación. Es importante distinguir el EA del espasmo acomodativo, una condición relacionada pero menos común<sup>11,42,43</sup>.
- **Inflexibilidad Acomodativa:** Se refiere a dificultades en adaptarse a cambios de demanda acomodativa<sup>38,41</sup>. Los pacientes muestran latencia y velocidades anómalas en

## INTRODUCCIÓN

su respuesta acomodativa. Los síntomas pueden presentarse tanto en visión lejana como próxima e incluyen astenopia, problemas de lectura y dificultades en cambios de enfoque. Las pruebas clínicas revelan problemas tanto para relajar como para estimular la acomodación<sup>11</sup>.

A pesar de la variedad de estudios, la prevalencia de estas disfunciones varía ampliamente, dependiendo de las poblaciones estudiadas y los criterios diagnósticos adoptados. Estas disfunciones subrayan la importancia de una evaluación optométrica rigurosa y específica para proporcionar diagnósticos precisos y soluciones adecuadas<sup>35</sup>.

### 1.1.3.2 DISFUNCIONES VERGENCIALES

Las disfunciones binoculares vergenciales no estrábicas se han consolidado en la literatura oftalmológica como un conjunto de anomalías que surgen debido a la incapacidad del sistema visual para mantener un alineamiento bifoveal óptimo o para fusionar las imágenes de ambos ojos de manera adecuada<sup>44</sup>. Es frecuente que estas disfunciones se manifiesten de forma latente, pasando inadvertidas<sup>45</sup> hasta que se presentan circunstancias que intensifican la demanda visual, como podría ser un incremento sostenido en actividades que requieren de visión cercana. Las estadísticas sobre la prevalencia de estas afecciones varían<sup>46</sup>, y se estima que oscilan entre un 0.1% y un 33%, dependiendo de la disfunción específica y de los criterios tomados en cada estudio<sup>13,35,38,41,47</sup>.

Para organizar estas disfunciones, la clasificación más comúnmente aceptada, es la propuesta por Wick<sup>11</sup> que destaca por su aplicación práctica y su respaldo teórico. Esta clasificación, que se basa en una revisión modificada del esquema original propuesto por Duane<sup>48</sup> modificada, ordena las anomalías en función de dos parámetros centrales: la vergencia tónica y el cociente AC/A.

Tanto las disfunciones binoculares horizontales como las verticales son anomalías que afectan la capacidad del sistema visual para mantener una alineación adecuada de los ojos y una visión binocular eficiente. Estas disfunciones se categorizan según el cociente AC/A, que representa la relación entre acomodación y convergencia.

#### **Disfunciones con cociente AC/A bajo:**

- **Insuficiencia de Convergencia (IC):** Esta anomalía se manifiesta cuando hay una incapacidad para converger adecuadamente en visión cercana<sup>49</sup>. Los síntomas asociados incluyen visión borrosa, diplopia, incomodidad visual, dolor de cabeza y fatiga visual, especialmente durante tareas prolongadas de lectura<sup>37,38,41,42,47,50-61</sup>.
- **Insuficiencia de Divergencia (ID):** Se caracteriza por una endoforia más pronunciada en visión lejana que en visión cercana<sup>11,44</sup>. Los síntomas comunes son diplopia y visión borrosa a distancia<sup>24,41,62</sup>.

#### **Disfunciones con cociente AC/A alto:**

- **Exceso de Convergencia (EC):** Se presenta con una endoforia en visión cercana más pronunciada que en visión lejana<sup>63</sup>. Los síntomas pueden variar desde visión borrosa y diplopia, hasta dolor de cabeza, especialmente relacionados con tareas en visión cercana<sup>35,38,41,47</sup>.

- **Exceso de Divergencia (ED):** Esta disfunción se caracteriza por una exodesviación más pronunciada en visión lejana que en visión cercana<sup>13</sup>. Los síntomas suelen estar asociados a la visión lejana, como la diplopia intermitente<sup>47,64</sup>.

#### Disfunciones con cociente AC/A normal:

- **Endoforia Básica:** Se manifiesta con una endoforia de magnitud similar en visión lejana y cercana<sup>11</sup>. Los síntomas pueden incluir astenopia, dolor de cabeza y fatiga visual<sup>41,47</sup>.
- **Exoforia Básica:** Esta anomalía se caracteriza por una exoforia de magnitud similar en ambas distancias<sup>26</sup>. Los síntomas comunes son astenopia y problemas relacionados con la lectura y otras tareas de cerca<sup>41,47</sup>.
- **Disfunción de Vergencia Fusional:** Aunque los pacientes presentan ortoforia o forias leves<sup>11</sup>, tienen vergencias fusionales reducidas, lo que puede llevar a síntomas como dolor de cabeza y visión borrosa<sup>41,47</sup>.

**Disfunciones Verticales:** Estas desviaciones se refieren a desalineaciones del ojo en el eje vertical<sup>34,65</sup>. Las hiperforias son desviaciones hacia arriba, mientras que las hipoforias son desviaciones hacia abajo. Los síntomas más comunes son diplopia, dolores de cuello y astenopia. También pueden presentarse mareos, visión borrosa y dolor de cabeza<sup>11,41,47</sup>.

## 1.2 LA INSUFICIENCIA DE CONVERGENCIA Y LA IMPORTANCIA DE LOS INSTRUMENTOS BASADOS EN LAS RESPUESTAS DE LOS SUJETOS (INSTRUMENTOS PRO) PARA SU INVESTIGACIÓN

### 1.2.1 LA INSUFICIENCIA DE CONVERGENCIA

En el ámbito clínico, es frecuente identificar pacientes que presentan insuficiencia de convergencia (IC), una anomalía de la visión binocular introducida en el apartado anterior, que se caracteriza por la incapacidad para conseguir o para ser capaz de sostener una convergencia óptima sin esfuerzo significativo<sup>11,49</sup>. Esta condición, generalmente benigna e idiopática, denota una discordancia entre las habilidades visuales del individuo y las exigencias asociadas a visión próxima<sup>66,67</sup>.

En cuanto a su prevalencia, la literatura científica muestra valores que oscilan entre un 0.3% y un 33%<sup>38,41,47,52,56-61,68-71</sup>, dependiendo de las franjas de edad, de sí las poblaciones en las que se realizan los estudios son de tipo general o de tipo clínico y de los criterios empleados para catalogar a los sujetos como sintomáticos o no sintomáticos.

En la literatura científica, se documentan como manifestaciones clínicas asociadas a la IC, un PPC alejado con diplopía al menos a 5 cm<sup>72</sup> en niños de edad escolar<sup>59</sup> y en adultos a 5 cm y recuperación de 7 cm, una mayor exoforia de cerca que de lejos, AC/A bajo (inferior a 3:1) y reserva fusional positiva disminuida o insuficiente para la magnitud de la exoforia y exceso de acomodación en visión próxima frecuentemente asociado<sup>11,49,73,74</sup>.

## INTRODUCCIÓN

Los últimos estudios del grupo de investigación “Convergence Insufficiency Treatment Trial” (CITT), consideran sujetos sintomáticos los que presentan exoforia en la visión próxima ( $\geq 4\Delta$  mayor que en visión lejana), insuficiente amplitud de fusión ( $\leq 15\Delta$  en la prueba de vergencia fusional positiva VFP a una distancia de 40 cm ó o si falla el criterio de Sheard) y dificultad en las pruebas de punto próximo de convergencia (rotura  $\geq 6$  cm)<sup>74</sup>. Tal y como describen Hussaindeen et al., la mejor prueba de cribado para identificar la IC es el PPC <sup>75</sup>.

Diversas etiologías han sido propuestas en relación a la IC<sup>50</sup> en la literatura científica. Sin embargo, se postula que su origen es primordialmente central, lo que conlleva a una insuficiencia en la relación acomodación-convergencia (AC/A). Esto se traduce en una disrupción en la interacción entre la acomodación y la convergencia o en una convergencia subóptima<sup>37,42,49,50,67</sup>. De acuerdo con Pickwell y Evans<sup>76</sup>, otras causas potenciales incluyen insuficiencia acomodativa, factores anatómicos tales como una distancia interpupilar incrementada, así como otras etiologías menos prevalentes como ambliopías o forias verticales descompensadas.

Entre los diagnósticos diferenciales de la IC se pueden incluir la pseudo-insuficiencia de convergencia, la exoforia básica, la parálisis de convergencia y el exceso de divergencia <sup>46,77</sup>. Varios autores califican la IC como primaria, cuando se da como consecuencia de cambios en la demanda visual de cerca, problemas de salud, ansiedad o falta de sueño y como secundaria, en el caso de exoforia, exotropía u otro tipo de problemas como paresias o parálisis del recto medio, lesiones cerebrales o traumas <sup>76,78</sup>. En cuanto a la clasificación de la IC asociada con amplitudes de acomodación bajas, existen ciertas controversias en la literatura científica, encontrándose autores que la clasifican como primaria <sup>79,80</sup>, mientras otros la clasifican como secundaria<sup>76</sup>. Para muchos autores, una amplitud de acomodación reducida en los tests clínicos se considera insuficiencia acomodativa como disfunción visual primaria y la IC se identifica como secundaria (pseudo-insuficiencia de convergencia).

Los síntomas asociados a la IC presentan gran diversidad. Típicamente emergen cuando el individuo se encuentra en contextos de lectura de duración prolongada. Cabe reseñar que la intensidad de estos síntomas tiende a aumentar durante la adolescencia y persiste hasta la segunda década de vida. Es frecuente que tales manifestaciones se acentúen con el uso de dispositivos electrónicos, como ordenadores, o en situaciones que impliquen una alta demanda visual<sup>50-52</sup>. Tal y como se refleja en la Figura 1.4, los síntomas comunes de la IC incluyen: astenopia, tensión ocular, cefalea, visión borrosa, diplopía intermitente, somnolencia, dificultad para la concentración y pérdida de comprensión después de cortos periodos de lectura o realización de actividades de cerca <sup>11,49,74</sup>. Otros síntomas asociados a la IC son quemazón, lagrimeo y diplopía constante <sup>13,73</sup>. En ciertos pacientes, se observa una reducción en la percepción de profundidad, o incluso experimentan malestar general asociado al movimiento o durante su desplazamiento <sup>41,44,50,57</sup> y también son numerosos los autores que sugieren una relación entre la IC, el déficit de atención y desórdenes de hiperactividad <sup>53-55</sup>. Todos estos síntomas están relacionados con la visión próxima (VP) y muchos de ellos empeoran al final del día, es decir, van asociados a la fatiga visual <sup>13,30,42</sup>.

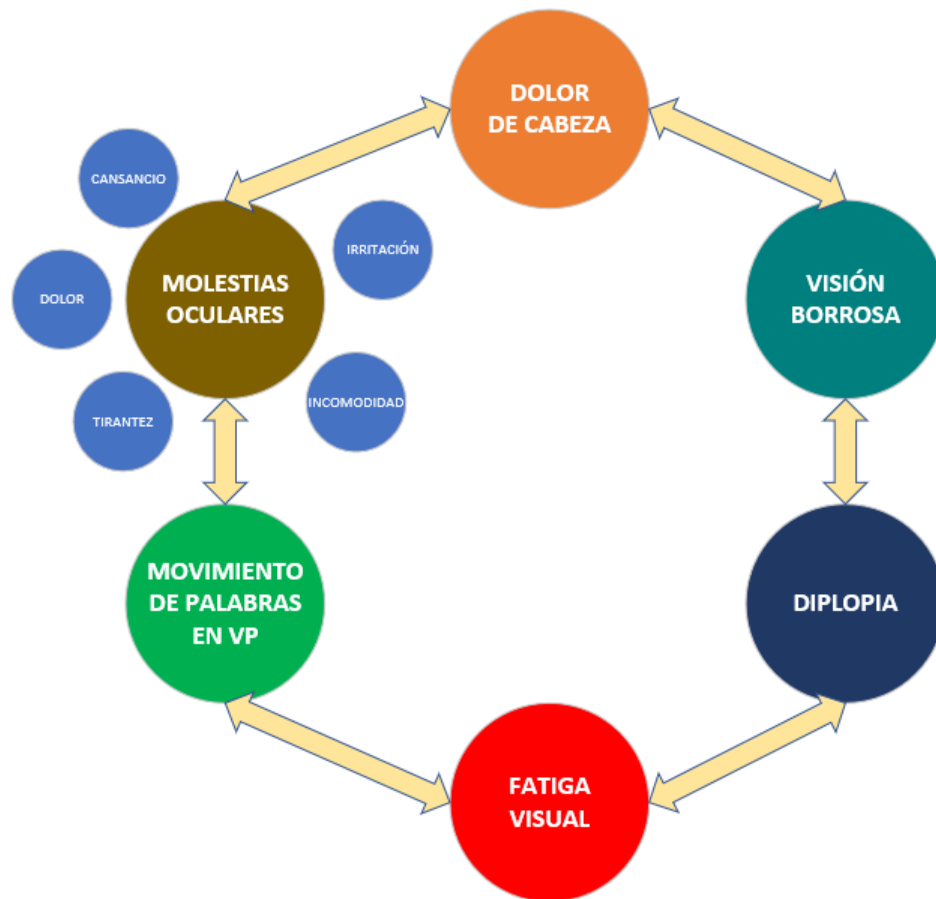


FIGURA 1.4 – SÍNTOMAS ASOCIADOS A LA IC

También es frecuente encontrar pacientes que reportan ser asintomáticos, debido a que existe supresión. Por ello, es necesario investigar si el paciente evita las actividades en visión próxima o se tapa un ojo mientras lee <sup>11</sup>. La IC interfiere con la lectura y el trabajo de cerca realizado en la escuela o el trabajo, impactando negativamente en el rendimiento y la calidad de vida de los pacientes <sup>74</sup>.

En cuanto al tratamiento de la IC, después de una correcta compensación óptica, la primera opción es la terapia visual (TV) con ejercicios de convergencia como por ejemplo acercamiento del lápiz o seguimiento de imágenes. Adicionalmente, existen otras opciones no quirúrgicas como las gafas de lectura con prisma de base interna, que ha mostrado ser útil en pacientes presbitas <sup>81,82</sup>. La cirugía es también una opción de tratamiento, pero rara vez se usa debido a su naturaleza invasiva y las posibles complicaciones que podría generar <sup>83,84</sup>. El CITT en su último estudio presenta resultados sobre la efectividad de la terapia visual en sujetos sintomáticos frente al placebo, midiendo PPC, VFP y las puntuaciones en la escala CISS pre y post terapia <sup>85</sup>.

### 1.2.2 LA IMPORTANCIA DE LOS INSTRUMENTOS “PATIENT REPORTED OUTCOMES” (PRO)

Las herramientas con las que cuenta el investigador para estudiar los síntomas asociados a la IC son, bien interrogar a los sujetos en estudio ayudándose de cuestionarios que pregunten sobre esos síntomas o revisar la historia clínica de los pacientes. Los cuestionarios empleados por los autores en ocasiones no se encuentran accesibles en las publicaciones, lo que obstaculiza su comparativa. Adicionalmente, en gran parte de los casos, no se detalla el proceso de elaboración de estos instrumentos, ni se proporciona información relativa a su validez y/o precisión. Lo que no es lo más adecuado como se justifica en los siguientes apartados. Comúnmente, son los autores quienes sugieren, fundamentándose en literatura previamente publicada sobre el tema, los potenciales síntomas, y son los sujetos quienes cuantifican el grado en que estos les afectan. Para recopilar la información en los ítems, suelen utilizar diferentes escalas de respuesta. Algunas dicotómicas (medición cualitativa con dos posibles resultados como sí/no, presente/ausente, etc.), otras ordinales (cualitativas o cuantitativas en las que el valor de respuesta refleja un orden entre las posibles respuestas) y también la escala visual analógica (*Visual Analogue Scale*), que representa un tipo de escala ordinal compuesta por una línea, ya sea horizontal o vertical, en la cual el paciente señala el punto que mejor representa su nivel de síntomas.<sup>86,87</sup>

La tendencia actual es utilizar escalas ordinales con un número impar de niveles, en general entre tres o cinco, manteniendo una opción neutra frente a las alternativas afirmativa o negativa en cuanto a la percepción del síntoma<sup>86</sup>. Existen multitud de alternativas, pero las más utilizadas son las escalas de tipo Likert<sup>88</sup>, en las que las opciones de respuesta de cada ítem están gradadas de forma ordinal, conteniendo varias alternativas, de carácter afirmativo, negativo y neutro en cuanto a la percepción del síntoma por parte del sujeto en análisis.

El uso de cuestionarios proporciona grandes beneficios, como son, la facilidad de uso, la escasa inversión de recursos necesaria para plantearlos, la rapidez, y la uniformidad en la captura de datos que permite su posterior análisis de forma adecuada<sup>89,90</sup>. Estos cuatro elementos son cruciales, no solamente en el ámbito de la investigación, sino también en otros contextos, tales como la práctica clínica optométrica, la prevención de riesgos laborales, la administración en el sector salud y la gestión empresarial.

En resumen, para investigar los síntomas y poder evaluar el éxito de los tratamientos de las anomalías de visión binocular no estrábicas y más en concreto la IC, es fundamental el uso de cuestionarios, dado que se requiere que los pacientes describan y/o cuantifiquen sus síntomas sin intervención interpretativa de un profesional. Al conjunto conformado por el cuestionario, acompañado de la información y documentación, que justifica su aplicación, se le conoce como instrumento sobre resultados percibidos por los pacientes (PRO - Patient Reported Outcomes)<sup>91</sup>.

### 1.2.3 LOS INSTRUMENTOS PRO COMO HERRAMIENTA PARA EL DIAGNÓSTICO DE INSUFICIENCIA DE CONVERGENCIA. LA ESCALA CISS “CONVERGENCE INSUFFICIENCY SYMPTOM SURVEY”

Aunque se han desarrollado y usado en el pasado diferentes escalas, hay instrumentos estandarizados diseñados para determinar los cambios en los síntomas asociados al tratamiento de insuficiencia de convergencia (IC) y otras anomalías de visión binocular no estrábicas.

En una revisión bibliográfica llevada a cabo por García-Muñoz et al <sup>13</sup> de los artículos publicados entre 1988 y 2012 sobre sintomatología asociada a las disfunciones acomodativas y binoculares no estrábicas, se encontró que de los 35 cuestionarios analizados sólo 3 habían sido validados para medir los síntomas asociados con incomodidad visual: el cuestionario Convergence Insufficiency Symptom Survey (CISSv15) <sup>59</sup> y su versión previa CIRS <sup>57</sup>, ambos diseñados para evaluar la insuficiencia de convergencia, y el cuestionario de Conlon <sup>92</sup> elaborado para anomalías visuales en general.

El cuestionario más ampliamente utilizado (en 21 de 56 estudios) fue el CISSv15, sobre el que se centra este trabajo. García-Muñoz et al <sup>14</sup> establecieron 34 categorías de síntomas y encontraron que 26 de ellas estaban relacionadas con la insuficiencia de convergencia, lo que explica la importancia dada en la literatura científica a esta disfunción visual respecto al resto. Los síntomas más frecuentemente mencionados fueron: dolor de cabeza, fatiga visual, visión borrosa, diplopía, fatiga visual y movimiento de las palabras en la visión de cerca.

El cuestionario CISS recoge las molestias más frecuentes relacionadas con tareas visuales en cerca y en numerosos estudios se ha determinado su validez, repetibilidad y eficacia en la identificación y control del efecto del tratamiento en pacientes con insuficiencia de convergencia (IC) <sup>57,59,60,74</sup>. Este cuestionario también ha contribuido a la cuantificación de síntomas de malestar visual en estudiantes <sup>40</sup> y en la identificación de individuos susceptibles de tener quejas visuales asociadas con incomodidad visual ligada a la visualización de pantallas y dispositivos 3D <sup>60</sup>.

El CISS fue desarrollado por el Convergence Insufficiency and Reading Study (CIRS). Inicialmente contenía 13 ítems y más tarde fue modificado para incluir otras actividades de visión próxima tales como videojuegos, lectura y otras actividades de ocio <sup>57,59,60,74</sup>, pasando a tener 15 ítems (CISSv15). En cada ítem el sujeto debe indicar la frecuencia de cada síntoma en una escala Likert <sup>88</sup> con puntuaciones entre 0 y 4, de acuerdo con la siguiente escala: nunca (0), poco frecuente (1), algunas veces (2), frecuentemente (3) y siempre (4). La puntuación final del cuestionario, cuyo valor estará comprendido entre 0 y 60, se calcula sumando los puntos correspondientes a cada ítem <sup>40</sup>. El cuestionario permite analizar, por un lado, si los síntomas están presentes y por otro, con qué frecuencia aparecen.

Esta última versión ha sido validada en niños de 9 a 18 años <sup>59</sup> y en adultos jóvenes de 19 a 30 años <sup>60</sup>. Se ha comprobado el valor diagnóstico del CISS tanto en niños como en adultos, ya que los pacientes con IC presentan una mayor puntuación en el cuestionario CISSv15 que sujetos con visión binocular normal. En esos estudios se establecieron como puntos de corte 16 puntos para niños (9-18 años) y 21 para adultos jóvenes (18-30 años) <sup>59,60</sup>, observándose una excelente discriminación, en el caso de los niños (sensibilidad 96%, especificidad 88%) <sup>59</sup>, y en el caso de los adultos jóvenes (sensibilidad 97,8%, especificidad 87%) <sup>60</sup>. El CISSv15, ha

## INTRODUCCIÓN

mostrado tener una alta consistencia interna y ser capaz de detectar cambios clínicos durante el tratamiento de la IC, considerando como el valor mínimo detectable (significativo y fuera del rango de variabilidad normal) un cambio de 10 puntos <sup>74</sup>.

Además de su enorme contribución para el diagnóstico de IC, se ha probado su valor diagnóstico en otras alteraciones binoculares no estrábicas con síntomas similares a la IC, tales como la insuficiencia de acomodación <sup>93</sup>.

### 1.2.4 CRITERIOS DE CALIDAD PARA VALORAR UN INSTRUMENTO PRO

En los últimos años se han publicado varias guías con directrices para la elaboración de nuevos instrumentos y/o para el análisis de los ya existentes<sup>9,91,94</sup>.

Por otra parte, el aumento del uso de cuestionarios en distintos países hace que sea frecuente la necesidad de utilizar aquellos que se han mostrado especialmente útiles y válidos, pero muchas veces están elaborados en un idioma diferente y no están validados para la población en la que se quieren emplear. En estos casos, es necesario hacer un proceso de traducción, adaptación transcultural y validación de los mismos. Este proceso no sólo implica la traducción y adaptación del cuestionario original al idioma de la población que se investiga, sino que también es necesario asegurarse de que las puntuaciones obtenidas con el cuestionario traducido sean equivalentes a las obtenidas con el original. Para facilitar esta tarea y garantizar que se hace correctamente, se cuenta con guías publicadas que dirigen el proceso para asegurar la equivalencia entre el cuestionario inicial validado y la versión traducida <sup>95</sup>.

Existe una iniciativa a nivel mundial llamada " Consensus-based Standards for the selection of health Measurement Instruments" (COSMIN) que tiene por objeto la estandarización de un sistema de puntuación para evaluar la calidad de los instrumentos PRO en base a nueve propiedades métricas <sup>96</sup>.

Un buen instrumento de medida debe tener dos características fundamentales:

- 1) Debe ser unidimensional, es decir, las medidas del instrumento deben evaluar un único constructo <sup>9,91,94</sup>.

La unidimensionalidad constituye una premisa fundamental tanto en la Teoría Clásica de los Test (TCT) como en la Teoría de Respuesta al Ítem (IRT). Esta premisa sostiene que, en un test específico, todos los ítems evalúan exclusivamente una única característica o constructo <sup>9,91,94</sup>.

Si bien se reconoce que no hay una escala que sea perfectamente unidimensional, se han ideado métodos para cuantificar la unidimensionalidad y establecer criterios para determinar si el grado de esta es adecuado para asumir que se está evaluando una única dimensión. Ejemplos de estas técnicas incluyen el Alpha de Cronbach y el análisis de Rasch. <sup>9,94,97</sup>

- 2) Debe conformar una variable de intervalo <sup>94</sup>

Las variables de intervalo son aquellas que cumplen las siguientes características:

- Su unidad de medida es cuantitativa
- Existe un origen, que está definido arbitrariamente
- Todas sus subdivisiones son iguales

Un ejemplo de variable de intervalo es, la temperatura. Su unidad de cuantificación es el grado, existe un origen que es arbitrario (es distinto en los grados centígrados que en los grados Kelvin) y sus subdivisiones son iguales. Este tipo de variables agregan información respecto a las nominales (que solo permiten nombrar y comparar utilizando la operación igualdad) y a las ordinales (que respecto a las anteriores introducen la posibilidad de ordenar, pero al no tener “un cero” no permiten operaciones aritméticas). En la Figura 1.5 se presenta un mapa sinóptico con los diferentes tipos de variable junto a su definición.

Al elaborar un cuestionario, cada ítem se presenta junto a una escala de respuesta. Dicha escala ofrece múltiples alternativas que denotan un orden y están vinculadas a una puntuación específica. Con ese orden, se pretende medir el constructo al que se refiere el ítem.

Las escalas de este tipo más extendidas en instrumentos documentales para la medida de síntomas, son las de tipo Likert <sup>88,98,99</sup> en las que se plantean diferentes puntos de referencia. La puntuación indicada en cada uno de estos puntos nos describe una variable ordinal, que permite inferir que alguien que seleccione la opción más positiva, ejemplo “mucho” experimenta un síntoma diferente a aquel que marque la más negativa “poco” y adicionalmente, que el primero experimenta un dolor más intenso que el segundo. No obstante, al estar ante una variable ordinal, se corre el riesgo de asumir que la diferencia entre “Mucho” y “Poco” es la misma que la que existe entre “Algo” y “Nada”, o pensar que entre “Mucho” y “Poco” hay el doble de dolor que entre “Algo” y “Poco”. Ese matiz, que es el que diferencia una variable ordinal de una variable de intervalo, impide la posibilidad de realizar con ellas operaciones matemáticas y estadísticas que proporcionen inferencias válidas.

Todos los datos que se recogen directamente en los cuestionarios son inicialmente de tipo nominal (es decir, solamente permiten comparar si las respuestas son iguales o distintas), u ordinales (es decir, solamente permiten comparar si una respuesta es mayor menor o igual que otra) <sup>100</sup>. Es aquí donde surge la necesidad de transformar las observaciones recogidas en variables de intervalo, que introducen la posibilidad de poder realizar operaciones aritméticas y cálculos estadísticos con las variables, de forma que los resultados obtenidos puedan considerarse válidos en términos cuantitativos.

En 1953, el matemático Georg Rasch desarrolló una solución para realizar la transformación de los datos capturados en variables nominales y ordinales a variables de intervalo solucionando así el problema inicial <sup>101</sup>. Desde entonces, esta metodología ha sido implementada en variadas disciplinas<sup>102</sup>, incluyendo las ciencias de la salud, la educación y la economía. La progresión tecnológica en las últimas décadas ha facilitado considerablemente su aplicación, elevando así su popularidad y propiciando el surgimiento de nuevas aplicaciones fundamentadas en el modelo de Rasch.

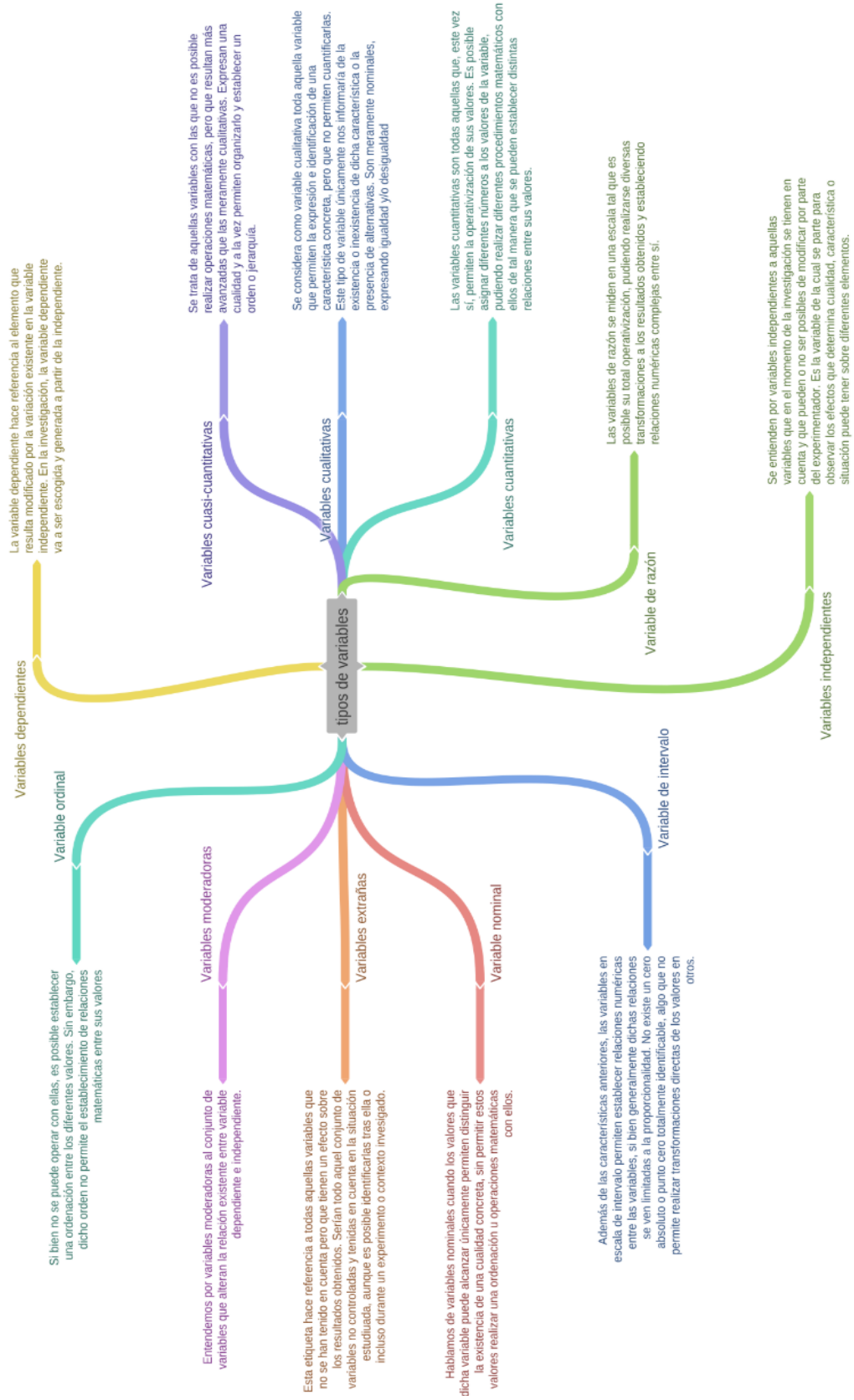


FIGURA 1.5 – TIPOS DE VARIABLES

### 1.2.5 EL MODELO PROPUESTO POR GEORG RASCH. CONSIDERACIONES GENERALES SOBRE EL ANÁLISIS DE RASCH

El modelo de Rasch forma parte de los modelos que agrupa la Teoría de Respuesta al Ítem (IRT), cuya idea fundamental es utilizar un modelo matemático para determinar la probabilidad de un sujeto de responder con éxito a un ítem, en función de su capacidad y de la dificultad de dicho ítem. Estos modelos ofrecen un procedimiento que convierte las observaciones registradas mediante cuestionarios (datos ordinales) en una medida de intervalo y además, permite verificar si esa transformación presenta la calidad necesaria <sup>103</sup>.

El modelo de Rasch establece una relación probabilística entre la dificultad inherente a un ítem específico y la aptitud del sujeto que contesta dicho ítem. Este modelo fue originalmente concebido para abordar ítems con una estructura de respuesta dicotómica, es decir, ítems que ofrecen únicamente dos alternativas de respuesta. Sin embargo, para extender su aplicabilidad a ítems que presentan múltiples opciones de respuesta, denominados ítems politómicos, se han derivado modelos adicionales. Entre estos modelos derivados se encuentran el Modelo de Crédito Parcial (Partial Credit Model, PCM) y el Modelo de Escala de Clasificación (Rating Scale Model, RSM) <sup>104</sup>.

El PCM opera bajo la premisa de que cada ítem define una escala de medida autónoma. En contraste, el RSM postula que las categorías de respuesta en todos los ítems de una escala particular, o en determinados agrupamientos de ítems dentro de dicha escala, son equidistantes en términos de su dificultad. Por lo tanto, en términos teóricos, el Modelo de Rasch para ítems con estructura de respuesta dicotómica puede ser conceptualizado como un caso particular del PCM.

El modelo inicial propuesto por Rasch, con una concepción dicotómica (Sí/No, Presente/Ausente, etc.), se fundamenta en los siguientes supuestos:

- 1) El atributo que se desea medir puede representarse en una única dimensión en la que se situarían conjuntamente las personas y los ítems
- 2) El nivel de la persona en el atributo y la dificultad del ítem determinan la probabilidad de que la respuesta sea correcta. Si el nivel de la persona se corresponde con el de la dificultad del ítem, así debe quedar representado en el modelo matemático elegido.

Rasch utilizó la función logística para modelar la relación. Llamando  $P$  a la probabilidad de una respuesta correcta a un ítem y por lo tanto,  $1-P$  a la probabilidad de una respuesta incorrecta a ese mismo ítem, y refiriéndonos al sujeto como  $s$  y al ítem como  $i$ , encontramos que el logaritmo neperiano de la razón de probabilidad  $p/(1-p)$ , (log-odd), es función de la diferencia entre la capacidad  $\theta$  del sujeto que responde,  $s$  y la dificultad  $\beta$  del ítem  $i$ .

$$\ln\left(\frac{P_{is}}{1 - P_{is}}\right) = (\theta_s - \beta_i)$$

#### ECUACIÓN 1

Así, cuando una persona responde a un ítem equivalente a su umbral de competencia, tendrá la misma probabilidad de una respuesta correcta y de una respuesta incorrecta:

## INTRODUCCIÓN

$$\ln\left(\frac{0,5}{0,5}\right) = (\theta_s - \beta_i) = 0$$

### ECUACIÓN 2

En este caso, particular, la dificultad del ítem es equivalente al nivel de competencia de la persona, es decir,  $\theta_s = \beta_i$ .

Este constructo matemático postula que si la competencia del individuo supera la dificultad del ítem ( $\theta_s - \beta_i > 0$ ), la probabilidad de emitir una respuesta correcta superará a la de generar una respuesta incorrecta<sup>4,105,106</sup>. Inversamente, si la competencia del individuo es inferior a la dificultad del ítem ( $\theta_s - \beta_i < 0$ ), la probabilidad de una respuesta correcta será menor que la de una respuesta incorrecta

La formulación de la Ecuación 3 Ecuación 3 se deriva de la modelización de la probabilidad de obtener una respuesta correcta, basada en la diferencia entre el atributo de competencia del individuo  $\theta_s$ , y el atributo de dificultad del ítem  $\beta_i$ . En la ecuación,  $e$  representa la base de los logaritmos neperianos.

$$P_{is} = \frac{e^{\theta_s - \beta_i}}{1 + e^{\theta_s - \beta_i}}$$

### ECUACIÓN 3

Los valores escalares de las personas y los ítems pueden expresarse en distintas métricas<sup>106</sup>. La más utilizada es la escala logit, que se corresponde de acuerdo a la Ecuación 1, con  $\ln\left(\frac{P_{is}}{1 - P_{is}}\right)$  es decir,  $\theta_s - \beta_i$ .

En esta escala la localización del “punto 0” es arbitraria, cumpliendo todos los requisitos de las variables de intervalo. Habitualmente se suele situar dicho “punto 0” de la escala en la dificultad media de los ítems. En ese caso particular es muy sencilla la interpretación de los parámetros de las personas. Significa que para los valores de  $\theta_s$  mayores que 0, las personas tienen una probabilidad superior a 0,50 de éxito en los ítems de dificultad media ( $\beta_i = 0$ ).

Aunque la escala logit puede adoptar valores entre más y menos infinito, la gran mayoría de los casos se sitúa en el rango  $\pm 5$ .

Otros usuarios del modelo prefieren, considerando los objetivos que tengan y la muestra que utilicen, situar el punto 0 en la habilidad media de las personas para evaluar el comportamiento de los ítems.

Asimismo, la familiaridad con la distribución normal de la escala logit ha llevado a multiplicar por la constante 1,7 el exponente de la Ecuación 2 para asimilar la escala logit a la distribución normal. Realizando esa transformación, la media y la desviación típica de la escala son similares a las de las conocidas puntuaciones típicas de la función Z (0 para la media y 1 para la desviación típica), consiguiendo por tanto que la casi totalidad de los casos quede incluida en el rango  $\pm 3$  (3 desviaciones típicas).

Cuando se administra un test, inicialmente se necesita estimar los parámetros (en la variable en cuestión) asociados tanto a los sujetos que se examinan  $\theta_s$ , como a los ítems que componen la prueba  $\beta_i$ . En ciertas circunstancias, uno de estos conjuntos de parámetros

puede ser conocido a priori, lo que simplifica el proceso de estimación. Un escenario común se presenta cuando los parámetros de los ítems ya han sido determinados a través de aplicaciones previas de la prueba, permitiendo así centrar la estimación exclusivamente en los parámetros de los sujetos mediante un enfoque denominado "estimación condicional". Por otro lado, en casos donde tanto los parámetros de los ítems como de los sujetos son desconocidos, se recurre a un proceso de "estimación conjunta".

La metodología predominante en este contexto es la conocida como de "máxima verosimilitud", que tiene como objetivo identificar los parámetros que maximizan la probabilidad de las respuestas observadas en los sujetos para cada ítem en particular. En el caso de la estimación condicional de los parámetros asociados a los sujetos, el método se asemeja a un proceso de búsqueda: dado un conjunto de parámetros de ítems ya conocidos, se calcula la probabilidad conjunta de las respuestas observadas a esos ítems para cada posible valor de  $\theta$ . Se asigna al sujeto el valor de  $\theta$  que maximiza esta probabilidad, y este valor es referido como el "estimador de máxima verosimilitud" ( $\theta'$ ).

Los procedimientos de cálculo necesarios para completar el análisis son largos y tediosos, por lo que es imprescindible recurrir a programas informáticos. Algunos de los más utilizados son, R (con sus paquetes eRm, ltm y difR), Jmetrik, Quest, ConQuest, RUMM, Winsteps, etc. <sup>107</sup>. En nuestro trabajo nos hemos apoyado principalmente en Winsteps, aunque hemos trabajado también con R, ConQuest y Jmetrik para algunos análisis.

Los estimadores de  $\theta$  son asintóticos e insesgados cuando los test son suficientemente largos. Su desviación típica, denominada "error típico de medida", es igual a:

$$SE(\theta') = \frac{1}{\sqrt{TI(\theta)}}$$

#### ECUACIÓN 4

El término  $TI(\theta)$  se designa como "función de información del test". Este constructo es particularmente relevante dado que el error típico de medida es inversamente proporcional a la cantidad de información proporcionada por el test. Por lo tanto, la función de información del test guarda una analogía conceptual con el principio de fiabilidad tal como se concibe en la Teoría Clásica del Test (TCT). La función de información del test se deriva de la suma acumulada de las funciones de información asociadas a cada uno de los ítems que conforman la escala completa.

La  $TI(\theta)$  representa un indicador que de forma sencilla sintetiza las contribuciones de cada ítem a la fiabilidad del test, y habilita por lo tanto la posibilidad de poder optimizar su diseño de cara a maximizar esa confiabilidad.

$$TI(\theta) = \sum I(\theta)$$

#### ECUACIÓN 5

Donde  $TI(\theta)$  es la "función de información del ítem":

$$I(\theta) = P_i(\theta)(1 - P_i(\theta))$$

#### ECUACIÓN 6

A partir de la Ecuación 6, se deducen dos implicaciones fundamentales. Por una parte, la información que un ítem particular proporciona no es constante, sino que varía a lo largo del continuo de la variable latente, y por otra, que el punto en el cual un ítem contribuye con la máxima cantidad de información se alcanza cuando la habilidad del sujeto  $\theta_s$  es idéntica al parámetro de dificultad del ítem ( $\theta_s = \beta_i$ )

### 1.2.5.1 VENTAJAS DEL MODELO DE RASCH

Las ventajas del modelo de Rasch respecto a la TCT y a otros modelos IRT han sido ampliamente difundidas <sup>6,106,108</sup>. Las más relevantes son las siguientes:

#### 1) Medición conjunta.

Significa que los parámetros de las personas y de los ítems se expresan en las mismas unidades y se localizan en la misma escala. Esta propiedad confiere al modelo de Rasch un carácter más realista que el de la TCT, ya que no es razonable mantener el supuesto de la invarianza de los ítems del modelo TCT (es obvio que no todos los ítems miden la misma cantidad del constructo) <sup>6</sup>. Esta particularidad posibilita el análisis de las interrelaciones entre individuos e ítems. Por lo tanto, la interpretación de las puntuaciones no se basa esencialmente en criterios grupales, sino en la detección de los ítems que el sujeto posee una elevada o reducida probabilidad de responder acertadamente. Tal atributo confiere al modelo de Rasch una notable profundidad diagnóstica.

y por ello hace que se ajuste perfectamente en cuestionarios relacionados con diagnósticos asociados a síntomas referidos por pacientes, propios de los instrumentos PRO <sup>94</sup>

#### 2) Objetividad específica.

Una medida sólo puede ser generalizada si su valor es invariante ante cambios en las condiciones específicas bajo las cuales se obtuvo. Desde este prisma, lo que se postula es que la diferencia entre dos sujetos en un atributo particular no depende de los ítems específicos empleados para su estimación y de forma análoga, la diferencia observada entre dos ítems en términos de su dificultad en un atributo particular tampoco depende de la muestra específica de individuos empleada para su cuantificación.

Esta premisa posibilita que tanto ítems como sujetos puedan ser remplazados por otros sin que ese hecho impacte a las inferencias obtenidas de la prueba. Fue el mismo Rasch el que denominó a esta propiedad “objetividad específica”

Matemáticamente, si dos individuos de niveles diferentes responden al mismo ítem, utilizando la Ecuación 1:

$$\ln\left(\frac{P_{i1}}{1 - P_{i1}}\right) = \theta_1 - \beta_i$$

$$\ln\left(\frac{P_{i2}}{1 - P_{i2}}\right) = \theta_2 - \beta_i$$

Por lo tanto, la diferencia entre ambas personas será igual a:

$$\ln\left(\frac{P_{i1}}{1 - P_{i1}}\right) - \ln\left(\frac{P_{i2}}{1 - P_{i2}}\right) = (\theta_1 - \beta_i) - (\theta_2 - \beta_i) = \theta_1 - \theta_2$$

Análogamente, si un mismo individuo responde a dos ítems con distinto grado de dificultad:

$$\ln\left(\frac{P_{1s}}{1 - P_{1s}}\right) = \theta_s - \beta_1$$

$$\ln\left(\frac{P_{2s}}{1 - P_{2s}}\right) = \theta_s - \beta_2$$

Y por lo tanto, la diferencia en dificultad entre ambos ítems será igual a:

$$\ln\left(\frac{P_{1s}}{1 - P_{1s}}\right) - \ln\left(\frac{P_{2s}}{1 - P_{2s}}\right) = (\theta_s - \beta_1) - (\theta_s - \beta_2) = \beta_1 - \beta_2$$

Por lo tanto, si los datos empíricos se alinean adecuadamente con el modelo descrito, se espera que las comparaciones entre individuos sean invariantes respecto a los ítems específicos que se utilicen para la evaluación. Además, las estimaciones de los parámetros asociados a estos ítems deberían permanecer inafectadas ante cualquier composición particular de la muestra empleada para su calibración.

Es importante contrastar este principio con la Teoría Clásica del Test (TCT), teoría en la que la puntuación obtenida por una persona está intrínsecamente ligada a los ítems específicos que se le administran y, la dificultad asociada a un ítem también puede cambiar dependiendo del grupo específico de individuos que se evalúen.

En esta propiedad de objetividad específica se fundamentan aplicaciones muy importantes como la equiparación de puntuaciones obtenidas con distintos test, la construcción de bancos de ítems de dificultades conocidas la realización de test adaptados a sujetos de nivel conocido o la construcción de cuestionarios adaptativos, por ejemplo.

### 3) Propiedades de intervalo

Es importante hacer notar que con la transformación que hace el modelo de Rasch la interpretación de las variaciones en la escala se mantiene constante a lo largo del atributo evaluado. Esto implica que, ante diferencias equivalentes entre un individuo y un ítem, se asocian probabilidades consistentes de obtener una respuesta acertada. Por esta razón, la escala logit exhibe características de intervalo, en contraposición a la TCT, en los que las puntuaciones están casi siempre modeladas por variables ordinales.

La variable de intervalo es especialmente importante, ya que constituye un prerrequisito esencial para poder aplicar las técnicas estadísticas paramétricas más comúnmente utilizadas en el ámbito de las ciencias sociales (como el análisis de varianza y la regresión, entre otros). Además, asegura la invarianza de las diferencias entre las puntuaciones a lo largo del continuo de la variable que se examina, habilitando de esta forma la posibilidad de realizar operaciones aritméticas y comparaciones de forma sencilla entre los parámetros de los ítems y de los sujetos. Especificidad del error típico de medida.

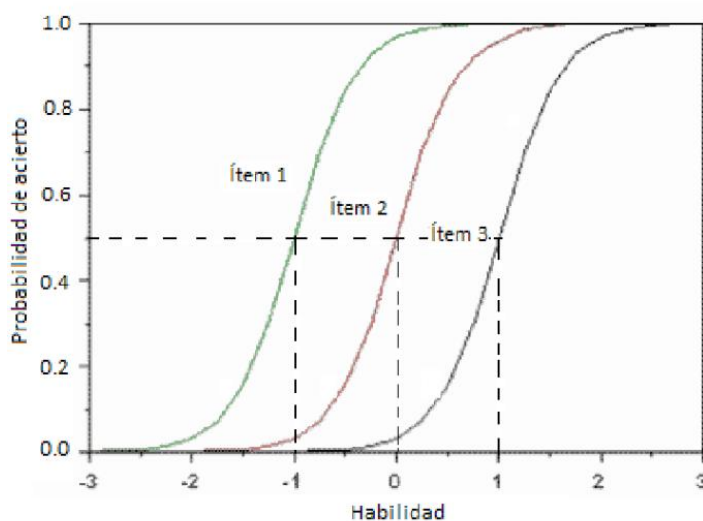
## INTRODUCCIÓN

La objetividad específica no garantiza que la exactitud en las estimaciones de los parámetros sea homogénea en diferentes conjuntos de ítems y personas<sup>105,106</sup>. Si los ítems resultan ser sencillos, los parámetros de los sujetos de nivel inferior se estimarán con mayor precisión. Por otro lado, si los individuos poseen un nivel elevado, los parámetros de los ítems más desafiantes se determinarán con mayor exactitud.

Bajo la Teoría Clásica de los Test (TCT), se presupone que los tests evalúan con idéntica fiabilidad en todas las áreas de la variable. El modelo de Rasch, en contraste, no adopta este postulado, que puede considerarse poco plausible, y en su lugar, permite cuantificar el volumen de información con el que se evalúa en cada segmento de la dimensión y escoger ítems que potencien la información en áreas del atributo previamente definidas. Esta última característica es de particular relevancia en tests orientados al criterio, donde es esencial maximizar la fiabilidad en las proximidades de los puntos de corte. Por ejemplo, un cuestionario que distingue entre pacientes que tienen o no tienen síntomas de determinada enfermedad

Para poder obtener todas las ventajas del modelo de Rasch, es necesario que los datos empíricos se ajusten al modelo.

De acuerdo con la Ecuación 2 la probabilidad de respuesta a un ítem depende sólo de los niveles de la persona y del ítem en el atributo medido como se puede ver gráficamente en la Figura 1.6.<sup>109</sup>



**FIGURA 1.6 – PROBABILIDAD DE ACIERTO EN FUNCIÓN DE LA DIFICULTAD DEL ÍTEM Y LA HABILIDAD DEL SUJETO**

La manifestación de respuestas atípicas, como el hecho de que individuos con competencias limitadas respondan acertadamente a ítems de alta dificultad, sugeriría que los parámetros asociados a sujetos e ítems son simplemente valores numéricos desprovistos de relevancia teórica.. Es por eso por lo que es importante analizar la coherencia en el ajuste al modelo. So se identificara una falta de ajuste, esta podría estar originada por una variedad de factores, como la multidimensionalidad, sesgo en los ítems, imprecisión en las alternativas de respuesta, respuestas aleatorias de los participantes o falta de motivación, cooperación o interés de los sujetos, errores de anotación de las respuestas por parte del examinador, copiado de las respuestas entre sujetos, etc<sup>110</sup>.

Los métodos analíticos existentes habilitan la posibilidad de identificar qué ítems y qué sujetos que desvían del modelo<sup>110 111</sup>. Las métricas más utilizadas para cuantificar la adecuación de los datos<sup>112</sup> se basan en el análisis de los residuos, es decir, en la cuantificación de las discrepancias entre las puntuaciones observadas y las que se esperaban a priori.

Conceptualmente, el residuo cuantifica la distancia entre el valor observado y el esperado, y por tanto su formulación es la siguiente:

$$y_{is} = (x_{is} - P_{is})$$

**ECUACIÓN 7**

Donde  $x_{is}$  representa la respuesta observada y  $P_{is}$  la probabilidad de que el sujeto  $s$  responda correctamente al ítem  $i$ .

Es común estandarizar los residuos dividiéndolos por su desviación estándar:

$$z_{is} = \frac{(x_{is} - P_{is})}{\sqrt{P_{is}(1 - P_{is})}}$$

**ECUACIÓN 8**

Para la evaluación cuantitativa de la conformidad con el modelo, se utiliza principalmente el estadístico "Infit", que se calcula como la media aritmética de los residuos cuadráticos, ponderada por su respectiva varianza ( $W_{is}$ ).

$$Infit = \frac{\sum z_{is}^2 W_{is}}{\sum W_{is}}$$

**ECUACIÓN 9**

El estadístico Infit para un ítem o un sujeto específico calcula realizando la media ponderada de los residuos cuadráticos, donde cada residuo cuadrático se pondera por su varianza asociada. La expectativa teórica para este estadístico es de 1, y según la convención establecida, los valores que superan 1,3 sugieren un desajuste cuando las muestras son inferiores a 500 sujetos. Cuando las muestras están comprendidas entre los 500 y 1000 sujetos, se endurece el criterio y en este caso valores que superen 1,2 se considerarían mal ajustados al modelo. Para muestras muy grandes (más de 1000 sujetos), se fijaría en 1,1 el umbral de aceptabilidad<sup>94,113-115</sup>.

### 1.2.6 IMPORTANCIA DE LA VALIDACIÓN DE LOS INSTRUMENTOS PRO PARA SU EMPLEO EN EL DIAGNÓSTICO CLÍNICO

De acuerdo con la definición de la Real Academia Española (RAE), un instrumento es un objeto fabricado “relativamente sencillo, con el que se puede realizar una actividad”<sup>116</sup>. Esta definición se aplica en varios ámbitos, encontrando los que se refieren a la música (instrumento musical), a la cirugía (instrumento quirúrgico) o a la ciencia (instrumentos de medida).

Los instrumentos de medida, pueden ser instrumentos mecánicos, como los utilizados en ciencias experimentales (balanza, termómetro...), que miden de forma objetiva variables cuantitativas (altura, presión arterial, temperatura...) o instrumentos documentales (cuestionario, escala, inventario), que miden magnitudes cualitativas con carácter subjetivo (calidad de vida, depresión, calidad de atención, inteligencia emocional, o como el que nos ocupa en este estudio, los síntomas de la insuficiencia de convergencia).

Todos los instrumentos necesitan certificar que son apropiados para lo que quieren medir y por lo tanto, tienen que ser validados. En el caso de los instrumentos mecánicos, las validaciones las realizan los fabricantes, mientras que en el caso de los instrumentos documentales la validación la realiza el investigador, como en el caso que nos ocupa. El objetivo es verificar sus propiedades psicométricas y garantizar que el instrumento es válido para el diagnóstico clínico.

Los pasos necesarios tanto para construir como para adaptar instrumentos documentales existentes que sean válidos y fiables representa un procedimiento complejo. Es importante garantizar que se cumplen unos estándares de calidad, y por esta circunstancia, el proceso utilizado, aunque pueda variar ligeramente de unos casos a otros, ha de tener siempre una serie de elementos comunes que repasaremos a continuación.

Cualquier investigación consta de una primera fase cualitativa a la que sigue otra cuantitativa soportada por métodos estadísticos y técnicas de análisis.

La fase cualitativa de la validación de instrumentos consiste en comprobar la validez del contenido de ese instrumento, es decir en asegurar que las preguntas y respuestas que contiene son las más adecuadas para conseguir medir la variable latente que se pretende.

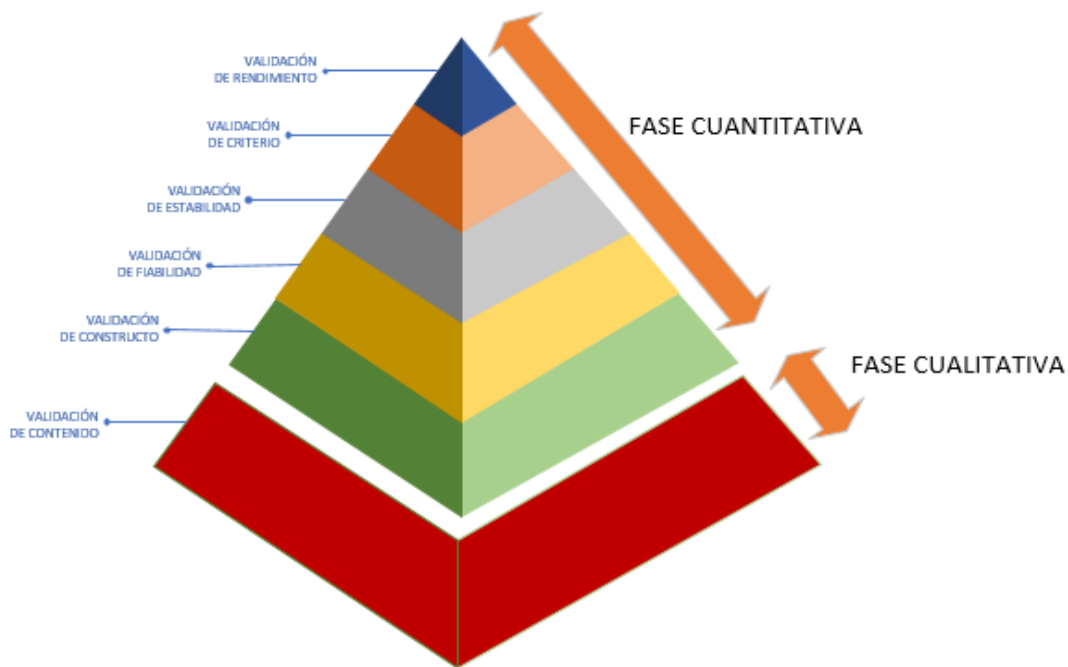
Dentro de la fase cuantitativa entrarían:

- La validación del constructo en la que se comprueba que el instrumento mide lo que debe medir y sólo eso, es decir, que es unidimensional.
- La validación de la fiabilidad del instrumento, es decir, su capacidad de discriminación.
- La validación de la estabilidad del instrumento, es decir, se examina si el instrumento arroja resultados consistentes al utilizarlo varias veces con la misma muestra, es decir, su repetibilidad y su reproducibilidad.
- La validación de criterio, esto es, si tomando diferentes grupos muestrales dentro de una misma población que fue observada por un mismo examinador, cambiando de examinador y volviendo a pasar el cuestionario a la misma población, los resultados persisten. También se pueden realizar pruebas de validación convergente (con un cuestionario que mida otra variable latente relacionada con la que mide el cuestionario en análisis), o pruebas de validez divergente, observando los resultados

frente a los de otro cuestionario que mide una variable sin relación, o con relación inversa con la variable latente que se pretende medir.

- La validación del rendimiento, es decir la comparación de los resultados del instrumento frente a los obtenidos en realidad mediante una exploración clínica, es decir observando los falsos positivos y falsos negativos y analizando el comportamiento del instrumento frente a la valoración clínica.

En la Figura 1.7 podemos encontrar un resumen gráfico de lo expuesto



**FIGURA 1.7 – PROCESO DE VALIDACIÓN DE UN INSTRUMENTO DOCUMENTAL**

En términos estadísticos, la validez de contenido se realiza a nivel exploratorio, de forma cualitativa. La validez del constructo se realiza a nivel descriptivo, la fiabilidad a nivel relacional. La estabilidad a nivel explicativo, la validez de criterio a nivel predictivo y la validez del rendimiento a nivel aplicativo.

# JUSTIFICACIÓN DEL ESTUDIO





## 2 JUSTIFICACIÓN DEL ESTUDIO

Es muy frecuente encontrar en los pacientes que acuden a los gabinetes de optometría, síntomas relacionados con disfunciones binoculares acomodativas y vergenciales no estrábicas y más concretamente con la IC <sup>41,47</sup>. Este tipo de anomalías presentan un espectro muy amplio de síntomas que son muchas veces difíciles de describir por las personas que los padecen y aún más de ser registrados adecuadamente por los profesionales de la salud visual <sup>11,49,74</sup>. De ahí la importancia de utilizar instrumentos que permitan realizar medidas de máxima calidad y servir de herramienta tanto para el análisis de la sintomatología como para el diagnóstico y el asesoramiento a los pacientes.

El desarrollo y la validación de los cuestionarios para evaluar sintomatologías clínicas se ha transformado en un área importante de la investigación. Sin embargo, de entre los cuestionarios de síntomas utilizados apenas existen instrumentos estandarizados y válidos ya que, con frecuencia, los cuestionarios, han sido confeccionados por los mismos autores, sin proporcionar detalles sobre su proceso de creación ni sobre su validez y/o consistencia. Se postula que un instrumento óptimo sobre resultados percibidos por los pacientes (PRO) debe ostentar dos características esenciales: evaluar un solo constructo (unidimensionalidad) y hacerlo a través de una escala de intervalo. La Teoría de Respuesta al Ítem (IRT), derivada del modelo de Rasch, propone un método para transformar las observaciones obtenidas mediante cuestionarios en una medida lineal y comprobar la calidad de dicha transformación<sup>117</sup>.

El aumento del uso de cuestionarios para valorar la calidad de vida <sup>118-120</sup>, cuantificar síntomas y evaluar actitudes y conductas en estudios de distintos países hace necesaria la traducción y adaptación cultural a otros idiomas de aquellos cuestionarios que se han mostrado especialmente útiles y válidos. Para facilitar esta tarea y garantizar que se hace correctamente se cuenta con guías publicadas que dirigen el proceso para asegurar la equivalencia entre el cuestionario inicial validado y la versión traducida<sup>115,121-124</sup>.

La elección de la escala debe ir vinculada a la idoneidad de su contenido teniendo en cuenta el contexto de su uso, ya que no existen herramientas únicas para la evaluación; se debe buscar, por lo tanto, aquellas que mejor se adapten al objetivo deseado, ya que cada instrumento tiene ventajas y desventajas respecto a otros <sup>3-5</sup>. El cuestionario que se traduzca y valide a otro idioma debe ser de aplicación práctica y tener la capacidad de generalización, para poder ser aplicado posteriormente a diferentes poblaciones, sin perder sus características básicas.

Aunque se han desarrollado y usado en el pasado diferentes escalas <sup>124,125</sup> no hay instrumentos estandarizados en español diseñados para determinar los cambios en los síntomas asociados al tratamiento de insuficiencia de convergencia (IC) y otras anomalías de visión binocular no estrábicas.

García-Muñoz et al. establecieron 34 categorías de síntomas y encontraron que 26 de ellas estaban relacionadas con la insuficiencia de convergencia, lo que explica la importancia dada en la literatura científica a esta disfunción visual respecto al resto. Los síntomas más frecuentemente mencionados fueron: dolor de cabeza, visión borrosa, diplopía, fatiga visual y movimiento de las palabras en la visión de cerca.

## JUSTIFICACIÓN DEL ESTUDIO

Con objeto de paliar la carencia existente hasta la fecha de un instrumento PRO en español capaz de medir los síntomas de IC con la calidad requerida, se pretende realizar la adaptación transcultural de la escala CISSv15 y someter a este conocido cuestionario por primera vez a las premisas actualmente recomendadas por el análisis de Rasch.

El cuestionario CISSv15 recoge las molestias más frecuentes relacionadas con tareas visuales en cerca y en numerosos estudios se ha determinado su validez, repetibilidad y eficacia en la identificación y control del efecto del tratamiento en pacientes con insuficiencia de convergencia (IC) <sup>57,59,60,74</sup>. Este cuestionario también ha contribuido a la cuantificación de síntomas de malestar visual en estudiantes <sup>40</sup> y en la identificación de individuos susceptibles de tener quejas visuales asociadas con incomodidad visual ligada a la visualización de pantallas y dispositivos 3D <sup>60</sup>.

El CISS fue desarrollado por el Convergence Insufficiency and Reading Study (CIRS). Inicialmente contenía 13 ítems y más tarde fue modificado para incluir otras actividades de visión próxima tales como videojuegos, lectura y otras actividades de ocio <sup>57,59,60,74</sup>, pasando a tener 15 ítems (CISSv15). En cada ítem el sujeto debe indicar la frecuencia de cada síntoma en una escala Likert <sup>88</sup> con puntuaciones entre 0 y 4, de acuerdo con la siguiente escala: nunca (0), poco frecuente (1), algunas veces (2), frecuentemente (3) y siempre (4). La puntuación final del cuestionario, cuyo valor estará comprendido entre 0 y 60, <sup>40</sup> se calcula sumando los puntos correspondientes a cada ítem. El cuestionario permite analizar, por un lado, si los síntomas están presentes y por otro, con qué frecuencia aparecen.

Esta última versión ha sido validada en niños de 9 a 18 años <sup>59</sup> y en adultos jóvenes de 19 a 30 años <sup>60</sup>. Se ha comprobado el valor diagnóstico del CISS tanto en niños como en adultos, ya que los pacientes con IC presentan una mayor puntuación en el cuestionario CISSv15 que sujetos con visión binocular normal <sup>59,60</sup>. Por otra parte, ha mostrado tener una alta consistencia interna y ser capaz de detectar cambios clínicos durante el tratamiento de la IC <sup>74</sup>.

Además de su enorme contribución para el diagnóstico de IC, se ha probado su valor diagnóstico en otras alteraciones binoculares no estrábicas con síntomas similares a la IC, tales como insuficiencia de acomodación<sup>93</sup>.

A través de la implementación de los procedimientos que, hasta la fecha, han demostrado ser más efectivos para la elaboración de escalas que evalúan síntomas, tal como el análisis de Rasch<sup>115,126</sup>, se espera poder aportar una nueva herramienta que supere las limitaciones de los instrumentos disponibles en español y validar de forma objetiva sus propiedades psicométricas de forma que se constate cuantitativamente su adecuación para la detección de síntomas asociados a IC y su correspondencia con la versión original del instrumento. Es la primera vez que se aplica el análisis de Rasch al conocido cuestionario CISSv15.

Esta nueva herramienta permitirá identificar a los sujetos más afectados por sintomatología asociada a IC y cuantificar de una forma más precisa del nivel de síntomas asociados con la alteración tras aplicar los tratamientos y terapias específicas, comparando los resultados en diferentes momentos del tiempo.

# HIPÓTESIS Y OBJETIVOS



## 3 HIPÓTESIS y OBJETIVOS

### 3.1 HIPÓTESIS

#### 3.1.1 HIPÓTESIS PRINCIPAL

La escala CISSv15 puede ser adaptada al español, de forma que un posterior análisis de sus propiedades psicométricas mediante la aplicación de las recomendaciones formuladas para la validación de instrumentos PRO<sup>3</sup> de calidad y el análisis de Rasch<sup>103,129</sup> confirme que el valor diagnóstico de la herramienta original se conserva íntegramente en la nueva escala.

De esta forma, quedará evidenciado que es posible crear una escala en español que proporcione una medida válida y fiable de los Síntomas Oculares y Visuales (SOV) asociados a la IC<sup>11,49,74</sup>

#### 3.1.2 HIPÓTESIS SECUNDARIAS

- 1) Gracias al soporte proporcionado por el análisis de Rasch<sup>103,129</sup> se podrá observar el comportamiento de cada ítem y su contribución a la medida en la nueva escala.
- 2) El análisis factorial de las respuestas aportadas a la nueva escala, posibilitará discernir los factores predominantes de los síntomas oculares visuales (SOV) relacionados con la IC.<sup>11,49,74,135</sup>
- 3) El análisis estadístico de las respuestas a la nueva escala, proporcionadas por la muestra empleada en el desarrollo de la misma, mostrará los niveles de normalidad de síntomas vinculados a la IC en el grupo poblacional de interés<sup>38-40</sup>.
- 4) El estudio de la correlación entre las puntuaciones de la nueva escala y las obtenidas por otros instrumentos PRO<sup>3</sup> ya validados, revelará la existencia de convergencia o divergencia entre las sintomatologías asociadas a las variables latentes a que se refieren esos constructos.
- 5) La nueva escala será capaz de identificar variaciones entre grupos que presenten diferentes grados de sintomatología<sup>103</sup>.

### 3.2 OBJETIVOS

#### 3.2.1 OBJETIVO PRINCIPAL

Realizar la adaptación transcultural al español y la validación de la escala “Convergence Insufficiency Symptom Survey”, CISSv15, consiguiendo así proporcionar a los especialistas de la salud visual un instrumento PRO capaz de medir la sintomatología visual y ocular asociada a la IC, la escala “Convergence Insufficiency Symptom Survey” versión española (CISSve). Esta nueva herramienta podrá ser utilizada tanto con fines diagnósticos como en seguimientos evolutivos de pacientes y en investigación.

#### 3.2.2 OBJETIVOS SECUNDARIOS

El análisis requerido para cumplir con el objetivo principal de la investigación también facilitará la consecución de diversos objetivos secundarios, los cuales tienen el potencial de enriquecer la calidad de la información científica existente respecto a este tema:

- 1) Validar la nueva escala mediante el estudio de sus propiedades psicométricas, utilizando para ello el análisis de Rasch <sup>103</sup> y las recomendaciones de Khadka et al. <sup>97, 115</sup>.
- 2) Comparar las propiedades psicométricas del CISSve con la versión original, CISSv15 <sup>59,60</sup>.
- 3) Analizar la correlación existente entre las puntuaciones obtenidas sobre la misma muestra por el CISSve y las obtenidas con otro instrumento validado que también evalúa síntomas asociados a VP, concretamente el CVSS17 <sup>51,127</sup>, que estudia síntomas oculares y visuales (SOV) asociados al uso de Pantallas de Visualización de Datos (PVD), determinando si existe o no convergencia en los resultados que pueda ratificar la validez de ambas escalas.
- 4) Analizar la correlación existente entre las puntuaciones obtenidas sobre la misma muestra por el CISSve y las obtenidas con otro instrumento validado que evalúa la percepción del bienestar emocional de esos pacientes <sup>94,128</sup>, concretamente el WEMWBS <sup>118,119</sup>, determinando si existe o no convergencia en los resultados que pueda establecer relación entre ambas variables.
- 5) Analizar la repetibilidad de la nueva escala a través del análisis de la repetibilidad test-retest <sup>123</sup>.
- 6) Identificar el tipo de síntomas de índole común asociados a la IC. La relación entre ellos y el peso de cada uno en el diagnóstico de esta disfunción de la visión binocular <sup>103,129</sup>.
- 7) Identificar, gracias al potencial que ofrece el análisis de Rasch <sup>103,129</sup> los ítems de la escala más valiosos a nivel diagnóstico en poblaciones generales y también en poblaciones sintomáticas.

- 8) Analizar los potenciales beneficios de la terapia visual en pacientes diagnosticados clínicamente de IC mediante el análisis de las puntuaciones obtenidas en la escala CISSve de una muestra piloto antes y después de recibir el tratamiento. Observar para qué grupo de síntomas la terapia resulta ser más eficaz.
- 9) Plantear el potencial valor diagnóstico de la escala CISSve comparando las puntuaciones obtenidas en la escala CISSve por los sujetos de una muestra piloto de pacientes sintomáticos diagnosticados clínicamente de IC frente a la obtenida por otros pacientes pertenecientes a una población general extraídos de la muestra utilizada para la validación de la escala CISSve.

### 3.2.3 OTROS OBJETIVOS ADICIONALES

Se abordarán objetivos adicionales surgidos como consecuencia de las investigaciones previas en el ámbito de la insuficiencia de convergencia y la adaptación transcultural del Convergence Insufficiency Symptom Survey (CISSv15) al español (CISSve). Estos objetivos se enfocan en la exploración de alternativas para mejorar la eficacia y adaptabilidad de la escala CISS en diferentes contextos y poblaciones.

- 1) Plantear, a través del conocimiento de las propiedades psicométricas <sup>103,129</sup> de la escala CISSv15 y CISSve, posibles alternativas de reducción de ítems para realizar “test rápidos” a través de cuestionarios más cortos.
- 2) Plantear, a través del conocimiento de las propiedades psicométricas <sup>103,129</sup> de la escala CISSv15 y CISSve, posibles escalas alternativas mejor adaptadas a los diferentes grupos de edad.
- 3) Plantear una posible metodología para ampliar el banco de ítems referidos a IC que constituyen CISSve, ofreciendo la posibilidad de un sistema basado en el análisis de Rasch para poder agregar o descartar algunos de ellos con la garantía de conservar o mejorar la información diagnóstica. Este objetivo es alcanzable en base a las propiedades del modelo planteado por Rasch que permite garantizar la calidad de la medida de cada ítem de forma independiente a los demás <sup>54,103</sup>.
- 4) Estudiar a través de los datos recogidos y modelos estadísticos predictivos de minería de datos, la posible creación futura de cuestionarios adaptativos para optimizar el tiempo de diagnóstico a través de la construcción de cuestionarios capaces de adaptar <sup>130</sup> el orden de las preguntas a la sintomatología percibida en el sujeto en los primeros ítems planteados, de forma que se consiga clasificarlos con un número menor de ítems y de forma más rápida <sup>131-134</sup>.



# MÉTODOS



## 4 MÉTODOS

Con el objeto de obtener la versión traducida y adaptada a español desde la versión original de CISSv15 de acuerdo a la metodología científica, tras realizar una exhaustiva búsqueda bibliográfica <sup>14,16-21,23,24</sup>, y evaluar las diferentes opciones propuestas por diferentes autores, optamos por utilizar el método propuesto por Gjersing, Caplehorn y Clausen <sup>122</sup> (muy similar al propuesto por Beaton, Guillemin y Bombardier <sup>124</sup>). Como se explicará en detalle más adelante, este proceso incluye varias etapas: la realización de dos traducciones directas (inglés- español) (T1 y T2) llevadas a cabo por traductores independientes, la versión consensuada de la traducción directa (T12), su posterior traducción inversa (español-inglés), también realizada independientemente (BT1 y BT2), la revisión realizada por parte de un panel de expertos para llegar por consenso a la versión prefinal del cuestionario y la aplicación de un pre-test a un grupo acotado de sujetos para obtener la versión final del cuestionario en español, la lengua destino, tal y como describen Gjersing, Caplehorn y Clausen <sup>122</sup> en sus recomendaciones para la adaptación transcultural de instrumentos de investigación. Una vez se tiene la versión final adaptada al español, se pasa a la fase de validación que también reúne varias etapas, tal y como se refleja en la Figura 4.1.

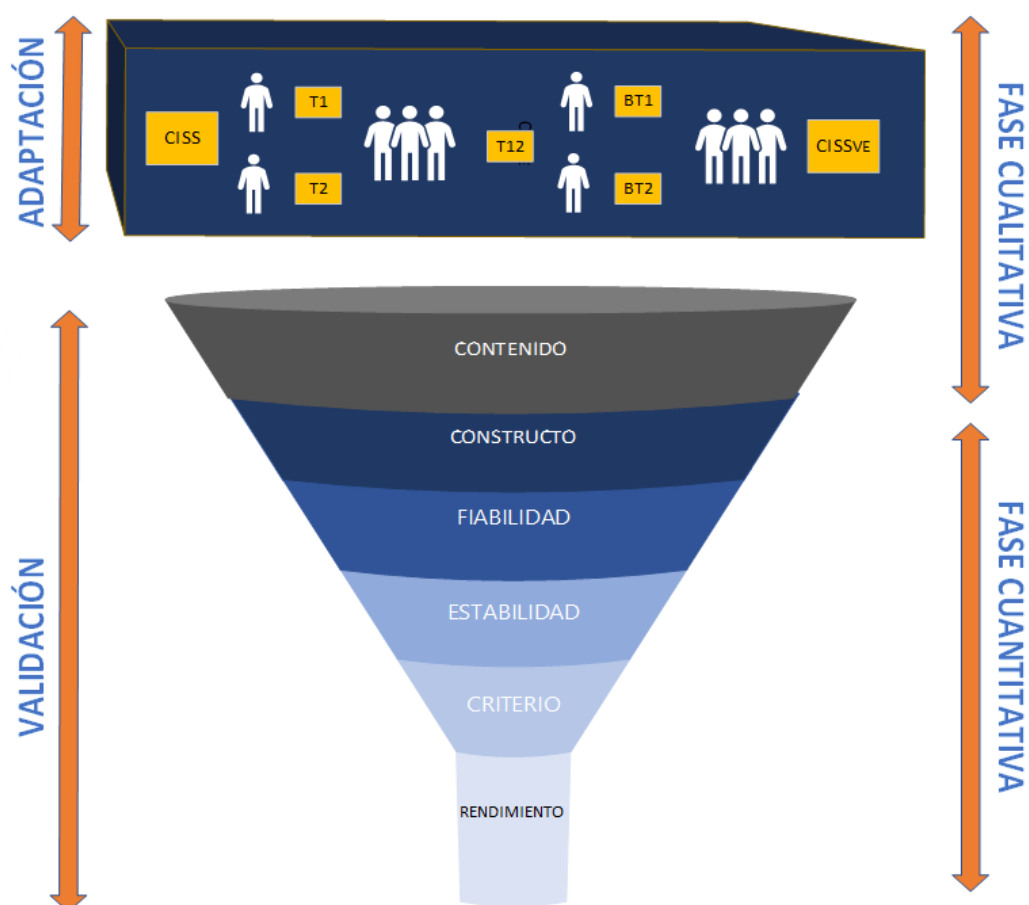


FIGURA 4.1 – PROCESO DE ADAPTACIÓN Y VALIDACIÓN DE LA ESCALA CISSV15

## MÉTODOS

Como primer paso antes de comenzar el estudio, se solicitó autorización a los autores de la versión original del CISS <sup>59,60</sup> para desarrollar la adaptación al español de este instrumento (CISSv15).

El estudio fue aprobado por el Comité Ético del Hospital Clínico San Carlos de Madrid (véase el 11.2 – anexo II) y el protocolo de estudio se adhirió a los principios de la Declaración de Helsinki. Todos los participantes dieron su consentimiento informado por escrito antes de participar. Para los participantes menores de 18 años, se obtuvo el consentimiento de su madre padre o tutor legal. Todos los niños mayores de 12 años también dieron su consentimiento antes de realizar cualquier prueba.

### 4.1 ADAPTACIÓN TRANSCULTURAL DE LA ESCALA CISSv15 AL ESPAÑOL

El proceso de adaptación propuesto por Gjersing, Caplehorn y Clausen <sup>122</sup> introducido anteriormente, comprende tres grandes bloques más una fase de pre-test, desde la traducción inicial de la escala CISSv15 hasta la versión final, designada como CISSve, como se puede ver de forma sintética en la Figura 4.2 y más detalladamente en la Figura 4.3. Estas figuras serán explicadas con más detalle en los siguientes apartados.

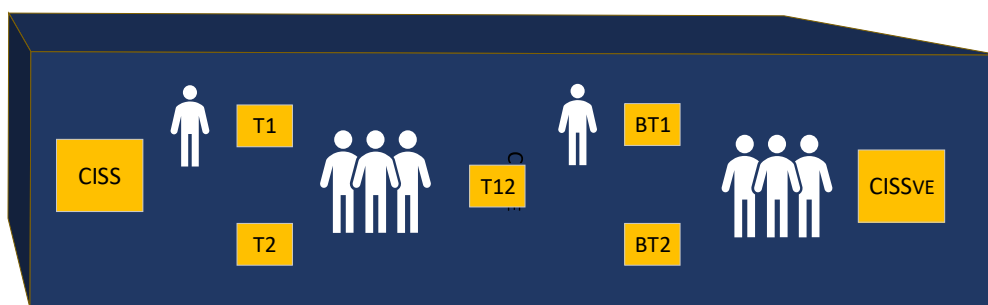


FIGURA 4.2 – PROCESO GENERAL DE ADAPTACIÓN DE LA ESCALA CISSv15

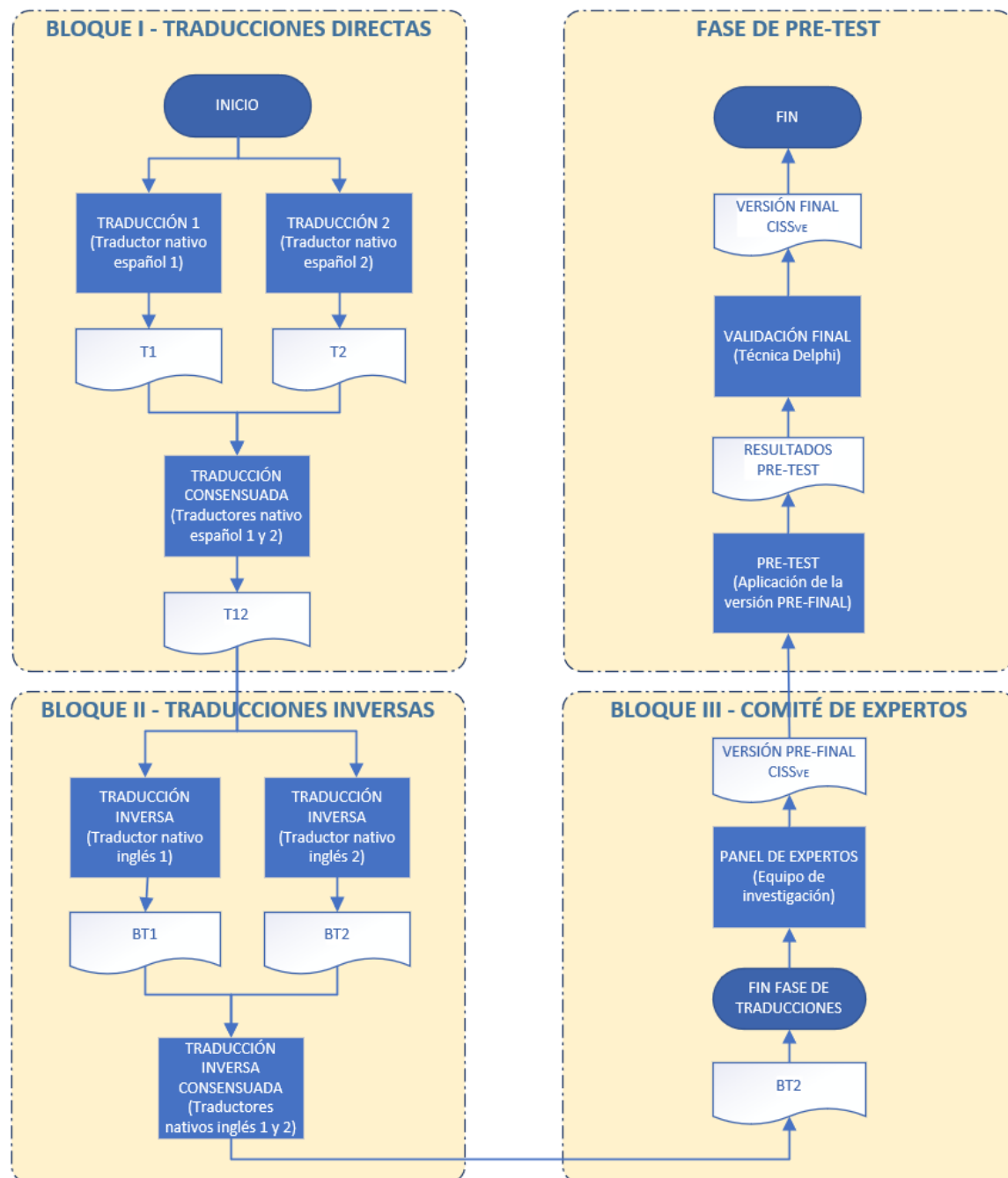


FIGURA 4.3— PROCESO DETALLADO DE ADAPTACIÓN DE LA ESCALA CISSV15

#### 4.1.1 BLOQUE I – TRADUCCIONES DIRECTAS

En el primer bloque, como se puede observar en la Figura 4.4, se llevan a cabo 2 traducciones directas. Estas traducciones fueron realizadas por dos traductores independientes, con lengua materna, la lengua destino (español) y con nivel muy alto en el idioma origen (inglés).

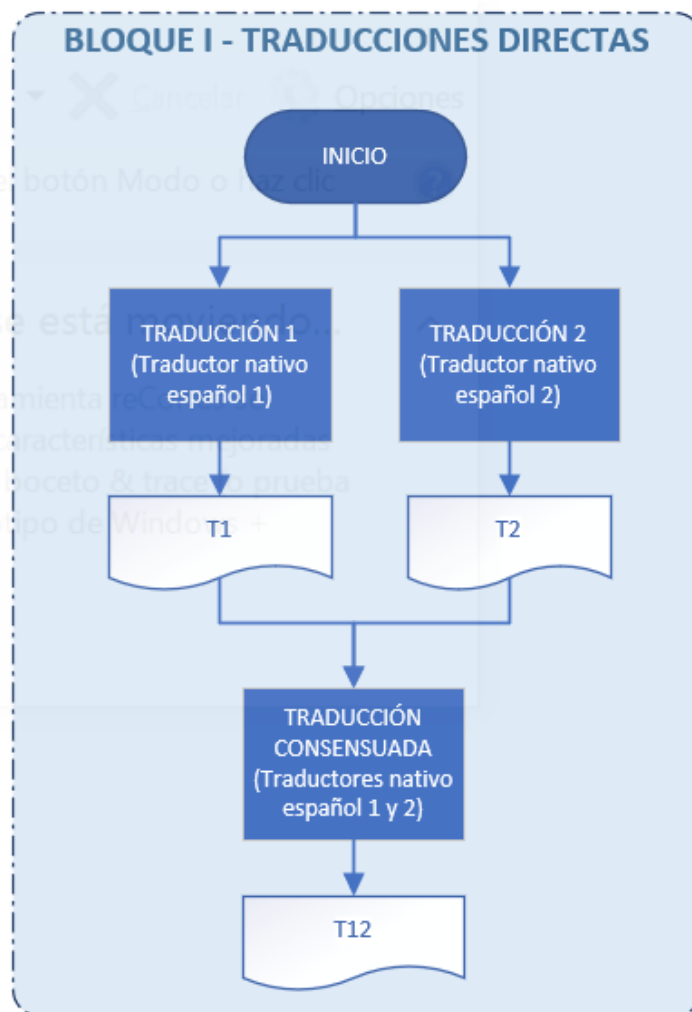


FIGURA 4.4 – BLOQUE I – TRADUCCIONES DIRECTAS

Se sigue la recomendación de Beaton et al. <sup>124</sup> que sugiere que los traductores deben pertenecer a ámbitos distintos. Uno de los dos traductores pertenece al grupo de investigación y el otro es un ejecutivo del sector bancario sin conocimientos en la materia sobre la que versa el cuestionario CISSv15.

Cada traductor, independientemente, partiendo de la versión original en la lengua origen (inglés) obtuvo una versión traducida a la lengua destino (español) del cuestionario CISSv15, incluyendo las instrucciones, los ítems y las opciones de respuesta. Denominamos a cada una de las versiones obtenidas por cada traductor, respectivamente, T1 y T2.

Terminado este proceso, se realizó una reunión entre los dos traductores independientes, moderada por el grupo de investigación en la que se abrió una discusión sobre las discrepancias encontradas entre ambas traducciones y en la que se obtuvo T12, la versión consensuada en español de T1 y T2.

#### 4.1.2 BLOQUE II – TRADUCCIONES INVERSAS

Partiendo de la versión consensuada en español obtenida en el bloque anterior (T12), de nuevo, se disgrega el trabajo en dos. Ahora son 2 traductores independientes de lengua materna coincidente con la del cuestionario origen CISSv15 (inglés) los que llevaron a cabo la traducción inversa, es decir, del idioma de T12 (español), al idioma del cuestionario origen CISSv15 (inglés). Los traductores realizaron la traducción sin tener ningún tipo de conocimiento sobre los conceptos explorados con el fin de evitar el sesgo. La secuencia se puede observar en la Figura 4.5.

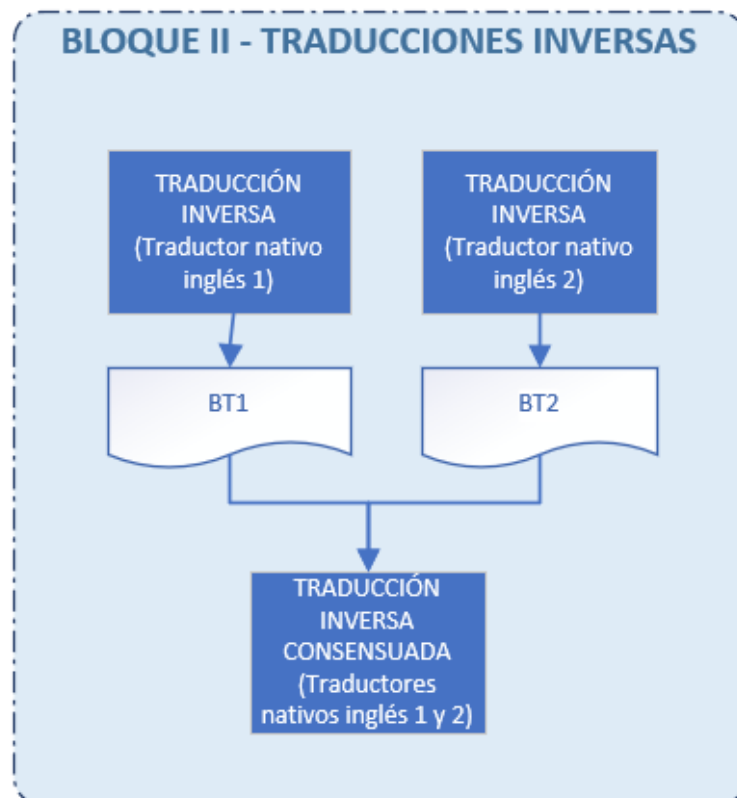


FIGURA 4.5 – BLOQUE II – TRADUCCIONES INVERSAS

De nuevo, uno de los dos traductores seleccionados fue un experto en la materia sobre la que versa el cuestionario, que ya había colaborado con Facultad de Óptica y Optometría de la Universidad Complutense de Madrid en otras ocasiones, mientras que el otro traductor era un profesional de un sector ajeno al ámbito sobre el que versa el cuestionario, concretamente se trataba de un profesional de la seguridad informática. En este caso se buscó un profesional sin ningún tipo de conocimiento sobre los conceptos explorados con el fin de obtener una versión exenta de cualquier tipo de sesgo que pudiera derivarse del hecho de desempeñar un trabajo afín al campo de estudio y por lo tanto estar condicionado por el conocimiento técnico y la familiaridad con la terminología propia de la materia sobre la que versan las preguntas de la escala.

## MÉTODOS

Las versiones obtenidas, en idioma inglés las denominamos BT1 y BT2 respectivamente, como abreviatura de la terminología más usada en la literatura inglesa para adaptación de cuestionarios, refiriéndose al término inglés “back translation”.

Realizado este proceso, se efectuó una reunión en la que intervinieron los dos traductores y el equipo de investigación, en la que se obtuvo la versión consensuada, BT12.

### 4.1.3 BLOQUE III – REUNIÓN DE EXPERTOS

Un panel de expertos que incluía a los cuatro traductores, y entre los que se encontraban un traductor profesional involucrado en el campo de la investigación, dos expertos en visión binocular y un miembro del equipo con experiencia en desarrollo de instrumentos PRO celebraron una reunión para consolidar las traducciones de las etapas anteriores (T1, T2, T12, BT1, BT2 y BT12) y crear la versión “pre-final” en español del CISS, designada CISS-versión española (CISSve). La Figura 4.6 ilustra el proceso.

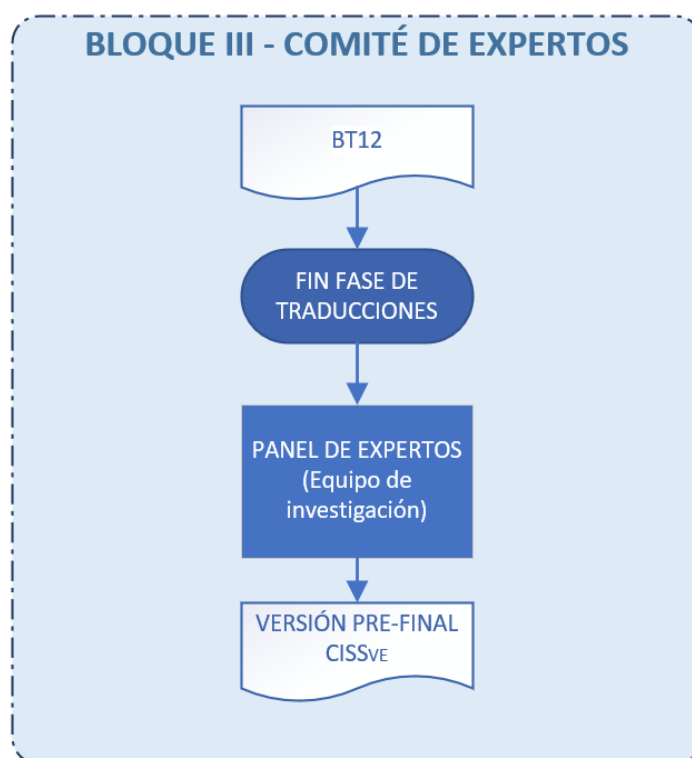


FIGURA 4.6 – BLOQUE III – REUNIÓN DE EXPERTOS

El panel de expertos resolvió las discrepancias por consenso y el resultado obtenido condujo a la versión preliminar del CISSve, correspondiente a la adaptación que se había obtenido como T12 (traducción directa consensuada).

#### 4.1.4 PRE-TEST

Para garantizar la idoneidad de esta versión consensuada, se realizó un pre-test a 43 hablantes nativos en español con edades comprendidas entre 9 y 30 años, y con objeto de asegurar la absoluta comprensión de todos los ítems de la escala CISSve por parte de los sujetos, se realizaron entrevistas cognitivas que fueron efectuadas utilizando una técnica de sondeo verbal.

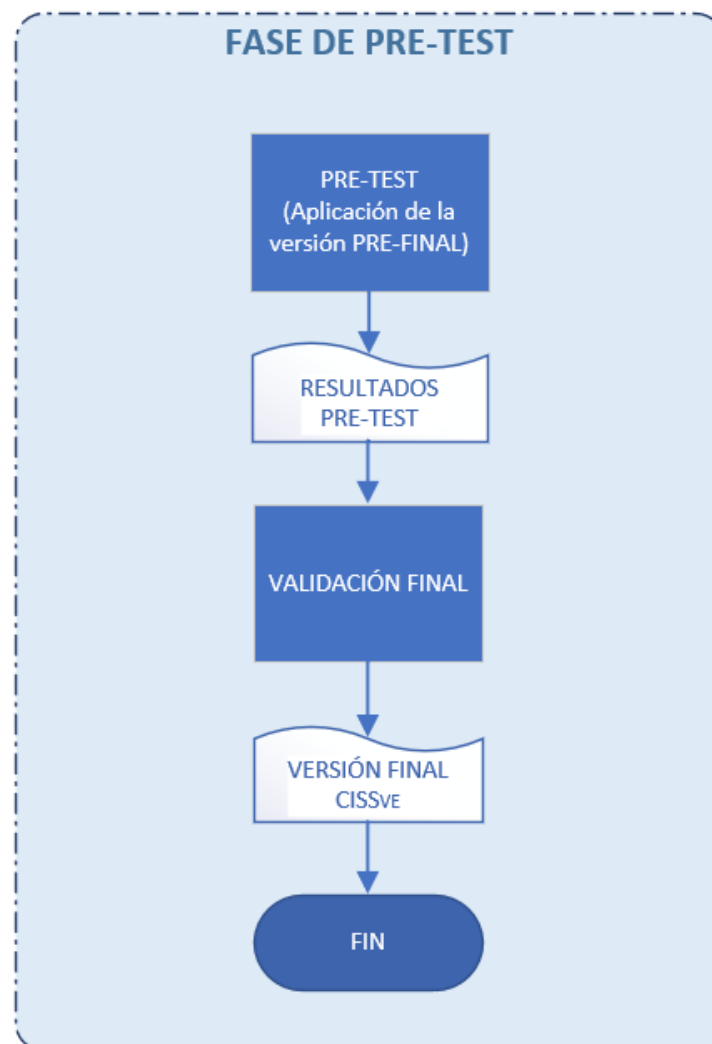


FIGURA 4.7 – FASE DE PRE-TEST

En Figura 4.7, podemos ver esquemáticamente el flujo correspondiente a esta fase y en la Figura 4.3, su ubicación en el proceso completo de adaptación ejecutado para la adaptación a español de CISSv15<sup>60</sup> siguiendo las directrices de Gjersing, Caplehorn y Clausen<sup>122</sup>.

## 4.2 VALIDACIÓN DE LA ESCALA CISSVe

### 4.2.1 CONSIDERACIONES GENERALES

Una vez disponible la escala adaptada a español CISSVe, el siguiente paso consiste en llevar a cabo la validación de esta. Este proceso consta de dos fases, una cualitativa y otra cuantitativa, como se puede ver en el esquema de la Figura 4.8.

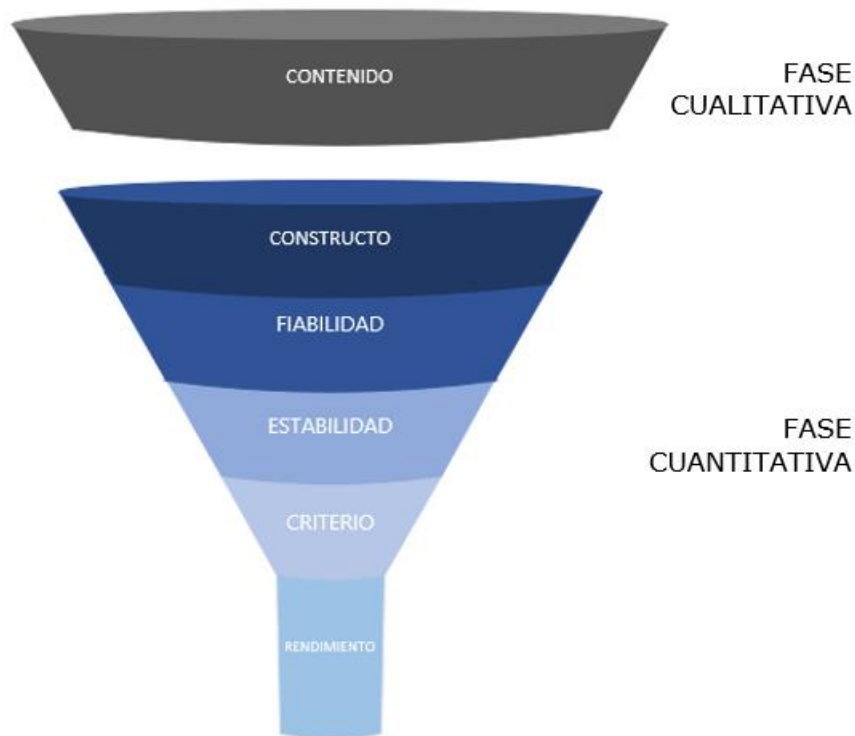


FIGURA 4.8 – FASES EN LA VALIDACIÓN DE UN INSTRUMENTO DOCUMENTAL

La fase cualitativa para la validación de instrumentos consiste en comprobar la validez del contenido de ese instrumento, es decir en asegurar que las preguntas y respuestas que contiene son las más adecuadas para conseguir medir la variable latente que se pretende.

La validación del constructo verifica que el comportamiento de los ítems es el adecuado y que el instrumento mide lo que debe medir y sólo eso. Primero se analiza desde el nivel estadístico descriptivo el comportamiento de los ítems a modo individual comprobando sus características: nivel de dificultad, discriminación, funcionamiento de las categorías de respuesta, y en su conjunto, se estudia cómo están distribuidos respecto a la muestra a analizar, lo adecuado que resulta el conjunto para medir esa muestra (“targeting”) y también, si conjuntamente miden la variable latente que queremos medir y solamente esa, es decir, si constituyen un constructo unidimensional.

Durante esta fase cuantitativa también se revisa la fiabilidad del instrumento (capacidad de discriminación), su estabilidad (repetibilidad) y su reproducibilidad (si el instrumento arroja

resultados consistentes al utilizarlo varias veces con la misma muestra y ante diferentes grupos muestrales dentro de una misma población). También se comprueba si el comportamiento es el mismo cambiando de examinador y pasando el cuestionario a la misma población (criterio).

También se realizó la validación convergente de la escala CISSve frente a la escala CVSS17, que mide una variable latente distinta a la que mide el CISSve pero relacionada con ella, y una validación divergente, relacionando las puntuaciones del CISSve con las de otro cuestionario que mide una variable sin relación aparente con la variable latente que se pretende medir, concretamente con el Warwick-Edinburgh Mental Well-being Scale (WEMWBS).

Como objetivo secundario se plantea la metodología para la validación del rendimiento, es decir la comparación entre los resultados obtenidos a través de las medidas que da el instrumento frente a la realidad, a través de una exploración clínica en gabinete, observando los falsos positivos y falsos negativos y analizando el comportamiento del instrumento frente a la valoración clínica del especialista.

#### 4.2.2 FASE CUALITATIVA. VALIDACIÓN DE CONTENIDO

La fase inicial en el proceso de validación, tanto para un instrumento de nueva creación, como para la adaptación transcultural a otro idioma de uno ya existente, es la validación de contenido.

Esta fase consiste en asegurar que las preguntas y respuestas que contiene son las más adecuadas para conseguir medir la variable latente que se pretende.

Para realizar este proceso se utilizan diferentes técnicas:

- Consulta a los sujetos participantes en el pre-test
- Consulta con expertos
- Revisión de la bibliografía relacionada con la variable latente que se pretende medir

##### 4.2.2.1 CONSULTA A LOS SUJETOS PARTICIPANTES EN EL PRE-TEST

El CISSve es una adaptación de un instrumento ya existente diseñado para evaluar aspectos subjetivos, por lo que resulta muy recomendable recoger la opinión que los propios sujetos de interés nos puedan ofrecer sobre la variable a medir.

Para recopilar esta información es recomendable llevar a cabo una investigación cualitativa utilizando grupos de discusión (“focus groups”) o entrevistas con informadores clave (“key informant interviews”) <sup>136</sup>.

En nuestro caso, se optó por llevar a cabo este proceso de acuerdo con las recomendaciones descritas en el apartado 1b la guía COSMIN<sup>96</sup> (“Estándares para evaluar la calidad de un estudio de entrevista cognitiva u otra prueba piloto realizada para evaluar la comprensibilidad y exhaustividad), realizando entrevistas individuales en primer lugar con cada participante. Adicionalmente se organizó un grupo de discusión para todas las personas del colectivo que

## MÉTODOS

había completado las respuestas a la escala. También se realizó una reunión con los padres de los menores que contestaron al cuestionario, con el objetivo de recabar sus comentarios.

Las respuestas de los usuarios sirvieron para evaluar la inteligibilidad y claridad de las preguntas formuladas.

### 4.2.2.2 CONSULTA A EXPERTOS

También nos dirigimos a Ópticos-Optometristas de nuestro entorno personal, profesional y académico para que revisaran si las palabras presentes en la escala CISSve para caracterizar las diferentes sintomatologías describían correctamente las molestias visuales y oculares que presentaban habitualmente sus pacientes con síntomas de insuficiencia de IC.

Siete Ópticos-Optometristas españoles (con  $7 \pm 5$  años de experiencia clínica) colaboraron en esta parte del estudio.

Este método sirvió para recopilar información y comentarios sobre los ítems de la escala, así como la posible redundancia que podían inferir algunos de ellos y la prescindibilidad de otros en anamnesis realizadas en la práctica profesional para formular diagnósticos tentativos.

### 4.2.2.3 REVISIÓN BIBLIOGRÁFICA

Mediante una búsqueda bibliográfica realizada en diferentes bases de datos (Medline, PsycINFO, EMBASE, IME y IBECs) en enero de 2018, restringiendo la búsqueda al periodo temporal comprendido entre 1980 y 2018 y a los idiomas español e inglés, se localizaron estudios publicados relacionados con la insuficiencia de convergencia en que se analizaban las distintas sintomatologías presentadas ante esta disfunción binocular a partir de la descripción que los pacientes hacían de ella <sup>13,30,41,42,44,50-55,57</sup>.

Esta revisión se llevó a cabo con una estructura taxonómica, en primer lugar, revisando los de ámbito más general y seguidamente los que describían síntomas.

También se realizó una búsqueda en marzo de 2018 en la base de datos de instrumentos PRO Proquolid, restringiendo la búsqueda al periodo comprendido entre 1980 y 2018 y a los idiomas español e inglés, con el objeto de localizar otras herramientas que evaluaran constructos relacionados con la insuficiencia de convergencia.

Posteriormente en febrero de 2020 se realizó una revisión adicional utilizando la base de datos de COSMIN<sup>96</sup> para revisiones sistemáticas de instrumentos "PRO" restringiendo la búsqueda entre los años 1980 y 2020, el tipo de instrumento a cuestionario y el tipo a PRO.

Las palabras clave que se utilizaron en la búsqueda fueron: "convergence", "insufficiency", "symptom", "survey", "dysfunction", "binocular", "vision", "anomalies".

### 4.2.3 FASE CUANTITATIVA. VALIDACIÓN DEL CONSTRUCTO, FIABILIDAD Y VALIDEZ CONVERGENTE Y DIVERGENTE. REPETIBILIDAD

Llegados a este punto, entramos en la parte cuantitativa del proceso de validación.

#### 4.2.3.1 ESTRATEGIA DEL ANÁLISIS

Para la generación de datos descriptivos, la comparación de las muestras, la evaluación de la repetibilidad y el análisis factorial, se utilizó el paquete de IBM SPSS Statistics versión 25.0<sup>137</sup>.

Para realizar el análisis de las propiedades psicométricas de la escala, se aplicó el método de Rasch, apoyando los cálculos principalmente en el paquete WINSTEPS (Versión 4.0.1)<sup>107</sup>. Puntualmente, para algunos cálculos, gráficos y verificaciones también se utilizaron otras herramientas: La suite R Studio (Versión 1.3.951)", con los paquetes de "R" "eRm (Versión 1.0)", "lrt (Versión 0.2.5)", "difR" (Versión 5.1)<sup>138</sup>, Jmetrik (Versión 4.1.1)<sup>139</sup> y ConQuest (Versión 5.36)<sup>140</sup>.

#### 4.2.3.2 DESCRIPCIÓN DE LA MUESTRA

Para llevar a cabo la validación de la escala CISSve y el análisis de la repetibilidad, la muestra se obtuvo de cuatro fuentes diferentes en Madrid, España. Una escuela de primaria (CEIP Vargas Llosa, Madrid), un instituto de secundaria (IES Juan Rodríguez Villanueva), un centro Universitario (Facultad de Óptica y Optometría de la Universidad Complutense de Madrid) y una compañía tecnológica (DXC Technology).

Ningún sujeto recibió compensación alguna por su colaboración.

El único criterio de inclusión que se estipuló fue que la edad de los sujetos estuviera comprendida entre 9 y 30 años.

La elección de este rango de edad queda justificada en base a la coincidencia con la de los grupos de estudio que fueron tomados por los autores de la escala original, CISSv15 cuando realizaron la validación de ese instrumento<sup>59-61</sup>. Es por ello que tratándose de una adaptación transcultural, se dé consistencia al estudio, seleccionando grupos muestrales de características similares.

Los criterios de exclusión utilizados fueron:

- Lengua materna diferente de la del cuestionario administrado
- Haber sido sometido a una cirugía visual previa (no refractiva)
- Tener una disfunción visual o neurológica
- Tomar medicamentos que pudieran afectar a la visión
- Tener cualquier tipo de disfunción que impidiera al sujeto leer o entender las preguntas del cuestionario (en papel)

Asumiendo un error máximo en la puntuación de un 5% (3 puntos), con un intervalo de confianza del 95% se calculó que la muestra mínima necesaria para llevar a cabo el estudio en

## MÉTODOS

una población en la que se desconociera la prevalencia a priori requería 384 sujetos. Como se revisó en el apartado **¡Error! No se encuentra el origen de la referencia.**, de acuerdo con la literatura científica, la prevalencia referida en investigaciones anteriores oscilaba en un amplio rango dependiendo de cada estudio (0.3%-33%)<sup>38,41,47,52,56-61</sup>, circunstancia que no permitió tomar ninguna asunción al respecto.

De acuerdo con los tamaños de muestra utilizados para la calibración de cuestionarios utilizados en estudios previos<sup>141,142</sup> que evaluaban otros constructos relacionados con la visión y en los que se utilizaba el análisis de Rasch, se fijó como objetivo distribuir el CISSve entre una muestra de unos 400 sujetos. En total, 458 sujetos (edad media, 16.02 años; rango de edad, 9–30 años; población femenina, 54.95%) completaron la versión española del cuestionario CISS (CISSve).

En cuanto al análisis del funcionamiento diferencial de los ítems muestrales (DIF), según las recomendaciones encontradas en la literatura científica<sup>143,144</sup> unos 200 sujetos resultan suficientes para obtener conclusiones estadísticamente significativas. En este trabajo se analizó el DIF tomando las respuestas de todos los sujetos de la muestra (458 sujetos), segmentados de acuerdo con los grupos de sexo y edad correspondientes, cumpliendo así el criterio.

Para el estudio de la repetitividad, se consideró un nivel de confianza de  $\pm 5\%$ , equivalente a un rango de  $\pm 3$  puntos (5% de la puntuación máxima de la escala – 60 puntos). De acuerdo con McAlinden et al.<sup>145</sup>, con estas premisas sería necesario administrar el cuestionario a al menos 192 sujetos. En total, 252 sujetos (todos los participantes excepto a los niños de educación primaria) contestaron por segunda vez los ítems de la escala CISSve, dejando entre ambas sesiones un periodo de entre 7 y 15 días.

En cuanto los requisitos para estudiar la validez convergente y la validez divergente del CISSve, se utilizó el software Granmo<sup>146</sup> (Institut Municipal d'Investigació Mèdica de Barcelona) como soporte para el cálculo de la mínima muestra necesaria para poder detectar una correlación de al menos 0,3 entre la puntuación registrada en la escala CISSve y las obtenidas en las escalas CVSS17 y WEMWBS, obteniendo un requerimiento de al menos 90 sujetos (Figura 4.9). En total, 252 sujetos de los 458 (todos los participantes excepto a los niños de educación primaria) respondieron a las tres escalas, CISSve, CVSS17 y WEMWBS

**FIGURA 4.9 – CÁLCULO DE LA MUESTRA NECESARIA PARA ESTUDIO DE VALIDEZ CONVERGENTE Y DIVERGENTE DE LA ESCALA CISSVe – GRANMO<sup>146</sup> (IMIM)**

#### 4.2.3.3 ANALISIS DE RASCH

El modelo utilizado para el análisis es el modelo de Rasch, un modelo englobado dentro de la teoría de respuesta al ítem (IRT) que transforma los valores de las puntuaciones brutas obtenidas por los sujetos y de las dificultades de los ítems a una escala común, de forma que, en esa escala transformada, un sujeto con una habilidad del mismo valor que la dificultad de un ítem tendrá el 50% de probabilidad de acertar ese ítem.

Según este modelo, la diferencia entre la habilidad de dos personas no depende de la dificultad de los ítems con los que se estima su capacidad. La probabilidad de que una persona responda con éxito a un ítem se podrá predecir de acuerdo a la relación existente entre la habilidad de la persona y la dificultad del ítem <sup>97</sup>.

Utilizando este análisis, tras la transformación que hace el modelo, los encuestados que tengan mayor nivel de síntomas (mayor puntuación) y los ítems de mayor dificultad (ítems menos puntuados, es decir, en los que sólo los sujetos más sintomáticos seleccionan las categorías de respuesta que indican una mayor percepción de ese síntoma) quedan ubicados en el lado positivo de la nueva escala y viceversa.

Debido a que el diseño de la escala original CISSv15 y consecuentemente de la escala adaptada a español CISSVe, está conferido por 15 ítems que se constituyen con categorías de respuesta iguales, se optó por utilizar el modelo RSM (Rating Scale Model) propuesto por Andrich <sup>3,6,136</sup>. Este modelo se utiliza cuando las categorías de respuesta son comunes a todos los ítems, es

## MÉTODOS

decir, todos los ítems tienen las mismas opciones de respuesta en una escala tipo <sup>88,147</sup>. Por lo tanto en nuestro caso era el más adecuado, a diferencia del PCM (Partial Credit Model)<sup>147</sup>, en que en el que las categorías de respuesta son propias de cada ítem, y pueden ser distintas en cada uno. En ese modelo, cada ítem per se constituye una escala independiente.

Recordemos que los cálculos de los parámetros del modelo se realizaron con paquete software Winsteps 4.0.1 (Winsteps.com, Beaverton, OR, USA) <sup>107</sup> y la evaluación de los resultados del modelo se realizó de acuerdo a los criterios propuestos por Khadka et al. <sup>114,115</sup>, que se pueden encontrar en 11.1 (Anexo I).

A través del método de Rasch, podemos analizar las propiedades psicométricas de la escala, evaluando las categorías de respuesta de los ítems de la escala, los estadísticos de ajuste al modelo de los ítems y de las personas, la dimensionalidad, el índice de separación de personas y los niveles de rendimiento de la escala CISSve, el targeting y el funcionamiento diferencial de los ítems por sexo y grupo de edad. También se evalúa su validez convergente y divergente, su repetitividad y los factores en que se agrupan los síntomas que componen la escala.

### 4.2.3.3.1 CATEGORÍAS DE RESPUESTA DE LOS ÍTEMS DE LA ESCALA CISSVE

Se evaluó el rendimiento de la estructura de la escala de calificación examinando el orden entre los umbrales de cada categoría respuesta.

Los umbrales de cada categoría de respuesta deben mantener un orden lógico de acuerdo con la intensidad de los síntomas y la separación entre categorías debe ser adecuada (1.4-5 logits) para considerar que todas ellas tienen una entidad propia capaz de diferenciarse claramente en nivel de la inmediatamente inferior y de la inmediatamente superior.

Según el criterio, el desorden de las categorías ocurre cuando las opciones de respuesta no siguen el orden jerárquico esperado <sup>6,148</sup>.

### 4.2.3.3.2 ESTADÍSTICOS DE AJUSTE AL MODELO DE LOS ÍTEMS Y DE LAS PERSONAS

El Infit y el Outfit son las métricas que el modelo de Rasch utiliza para evaluar la bondad del ajuste al modelo de los ítems y de las personas. Se calculan como la suma de los errores cuadráticos ponderada por su varianza, por lo que el valor teórico esperado es 1 en ambos casos y el intervalo que se considera como aceptable en cuanto a su adaptación al modelo es (0.5,1.5) según Khadka et al. <sup>114,115,148</sup>. Para muestras menores a 500 sujetos como sucede en nuestro caso, es (0,7-1,3).

El hecho de utilizar estos dos indicadores en lugar de uno solo para valorar la bondad del ajuste al modelo pretende diferenciar entre las zonas centrales y las periféricas de igualdad/diferencia de nivel ítem/sujeto. El Infit es más sensible a los errores en los ítems que están más cercanos en nivel a los sujetos medidos y el Outfit es más sensible a errores en los ítems de nivel muy dispar al del sujeto, es decir, los que son o muy fáciles, o muy difíciles para ese sujeto <sup>107,112</sup>.

Son indicadores muy potentes para poder detectar fácilmente desviaciones en el modelo y, por lo tanto, tomar decisiones referentes a los ítems. Por ejemplo, la eliminación de ciertos ítems que por su planteamiento resultan confusos y de los que se obtienen respuestas inesperadas. También pueden ser decisiones referentes a los sujetos, como por ejemplo la

eliminación de sujetos que responden al azar o no colaboran y responden las preguntas inadecuadamente, bien por falta de interés o por cualquier otra razón.

#### 4.2.3.3.3 DIMENSIONALIDAD

La escala se considera unidimensional cuando hay una única variable latente de interés y el nivel de esta variable latente es el foco de medición<sup>148</sup>.

Para evaluar la multidimensionalidad, utilizamos los resultados del análisis de componentes principales de Rasch (PCA) de residuos estandarizados, que busca patrones en la parte de los datos que no concuerda con las medidas esperadas por el análisis de Rasch, es decir, los datos inesperados. Se utilizan dos parámetros, la cantidad de varianza bruta explicada por la medida, por un lado, y por otro el autovalor (eigenvalue) de la varianza inexplicada en el primer contraste. Se considerará que el cuestionario es multidimensional cuando la varianza bruta explicada por las medidas es inferior al 50% y el autovalor del primer contraste del PCA es  $\geq 2$ . Esto indicaría que el cuestionario contiene un subconjunto de ítems que estarían midiendo otro concepto distinto<sup>114,115,148</sup>.

Cuando existen grupos de ítems que comparten los mismos patrones de datos inesperados, esos elementos probablemente también comparten un atributo en común, distinto al que mide la variable latente y al que nos referimos como una "dimensión secundaria" o "contraste"<sup>107</sup>.

Para evaluar la multidimensionalidad, también observamos el coeficiente de correlación desatenuado entre el primer y el segundo contraste obtenidos del análisis de PCA.

El valor de la correlación desatenuada, corresponde aproximadamente al valor de la correlación entre los dos contrastes sin error de medición. Cuando está cerca de 1, implica que las dos dimensiones miden realmente atributos de la misma variable latente. Según Linacre<sup>107</sup>, 0.82 debería ser usado como límite para considerar que los dos contrastes miden la misma variable, ya que indicaría que son el doble de dependientes que de independientes. Mientras que el valor 0.87, indicaría que el factor de dependencia sería triple al de independencia.

Debido a que el PCA de los residuos estandarizados es sólo indicativo de las dimensiones secundarias y no puede interpretarse como un análisis factorial habitual, también se realizó un análisis factorial exploratorio para identificar los componentes principales del CISSve y la cantidad de variabilidad explicada por cada uno de ellos.

Conceptualmente el análisis de las componentes principales es una técnica de reducción de variables, basada en poner un número más reducido de variables como combinación lineal "C" de las originales. Estas variables que se obtienen, son consecuencias o características de la variable latente que se está midiendo. En el análisis de las componentes principales se evalúa toda la varianza, la explicada por las medidas y la no explicada asignando pesos "w" a cada variable "Y". Por el contrario, el análisis factorial aborda la reducción de datos de una manera fundamentalmente diferente. Es un modelo para medir la variable latente "Y". Esta variable latente no se puede medir directamente con una única variable, es decir, la insuficiencia de convergencia no se puede medir directamente, pero sí se puede ver a través de las consecuencias que causa originando un conjunto de variables "Y<sub>i</sub>". Por ejemplo, tengo insuficiencia de convergencia y veo doble...

## MÉTODOS

El modelo de medición de análisis factorial para un modelo simple de un factor se parece al descrito en la Figura 4.10., en el que F, el factor latente, está causando las respuestas en las cuatro variables “Y<sub>i</sub>” medidas. Entonces las flechas van en la dirección opuesta a PCA.

Al igual que en PCA, las relaciones entre F y cada Y están ponderadas, y el análisis factorial está determinando los pesos óptimos (coeficientes “b<sub>1</sub>, b<sub>2</sub>, b<sub>3</sub>, b<sub>4</sub>”). En este modelo tenemos un conjunto de términos de error que son designados por los “u”. Esta es la varianza en cada Y que el factor no explica.

En la Figura 4.10 se puede ver gráficamente la diferencia existente entre ambas técnicas. Posteriormente en nuestro trabajo realizaremos un análisis factorial de nuestra escala.

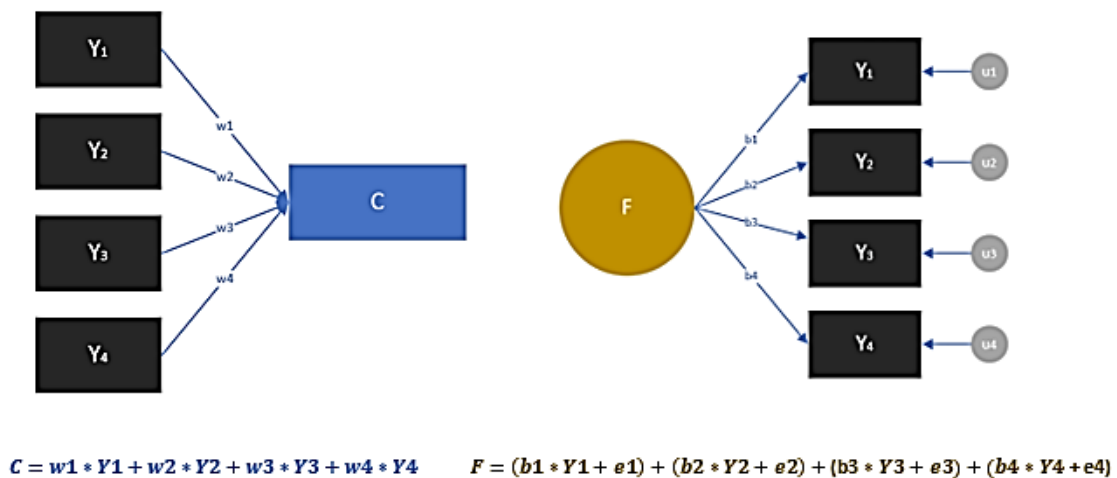


FIGURA 4.10 – DIFERENCIA ENTRE ANÁLISIS DE LAS COMPONENTES PRINCIPALES Y ANÁLISIS FACTORIAL

### 4.2.3.3.4 ÍNDICE DE SEPARACIÓN DE PERSONAS (PSI) Y NIVELES DE RENDIMIENTO DE CISSVE

El Índice de Separación de Personas (PSI), dentro del análisis de Rasch, es un indicador de confiabilidad, que tiene la misma función que el  $\alpha$  de Cronbach<sup>149</sup> en la teoría clásica del test, y presenta analogías con él tanto en sus valores como en la forma de interpretarlo<sup>6,107</sup>.

Este índice, cuyo valor mínimo aceptable es 2, se obtuvo a través del software Winsteps y es un indicador de la precisión de la medida que permite valorar hasta qué punto el instrumento puede distinguir entre los diferentes niveles de habilidad de los participantes.

El número de diferentes niveles de rendimiento se calculó de acuerdo con el método descrito por Wright<sup>150-153</sup>. Este método define niveles estadísticamente distintos de dificultad del elemento como estratos de dificultad con centros separados por tres errores de calibración, entonces este índice de separación, G, puede traducirse en el número de estratos de elementos definidos, H, mediante la ecuación  $H = (4G + 1)/3$ , donde G se calcula como el cociente entre la desviación estándar "verdadera" y el error de medición promedio<sup>51,152,154-156</sup>.

#### 4.2.3.3.5 ADECUACIÓN ESCALA-MUESTRA (“TARGETING”)

La medida en que la dificultad de los ítems está adaptada al nivel de las habilidades visuales de los participantes. Está definido como la diferencia entre la dificultad promedio de los ítems y el nivel promedio de los síntomas de los sujetos<sup>122</sup>.

Como ya hemos explicado, el modelo de Rasch sitúa en una misma escala de medida la habilidad de los sujetos y la dificultad de los ítems. Esto nos permite comprender rápidamente si un ítem es fácil o difícil para un sujeto determinado, o viceversa, si un sujeto es suficientemente hábil para poder resolver un ítem con solvencia.

Llevado a la interpretación sintomatológica, los sujetos más sintomáticos y los ítems más difíciles se situarán en la parte más positiva de la escala, y lo contrario ocurrirá con los sujetos que presenten menos síntomas y con los ítems más fáciles, que se posicionarán en la zona de valores más bajos en la escala.

El software estadístico utilizado, Winsteps 4.01<sup>107</sup>, nos permite tomar como referencia, ítems o sujetos y situar el cero de la escala donde consideremos más conveniente para realizar nuestro análisis. En nuestro caso, optamos por elegir como referencia (“0 de la escala”) el punto que representa la dificultad media de los ítems, eligiendo una desviación típica de 1 logit, de forma que concentremos el 68% de los ítems entre -1 y 1 logits, el 95% entre -2 y 2 y el 98% entre -3 y 3 logits.

Los sujetos quedarán mapeados en esa misma escala, encontrándose que estarán más o menos dispersos, en función del tipo de muestra de que se examine y más o menos centrados o desplazados respecto a los ítems en función de si la muestra tomada representa una población más o menos sintomática.

La medida que proporciona un ítem será más precisa en sujetos cuyo valor de “síntomas” esté más próximo a la “dificultad” de ese ítem y menos cuándo este más alejada. Por ello, nos encontraríamos en el caso ideal cuando la muestra a analizar estuviera centrada en el mismo punto que los ítems, y tuviera una dispersión similar. Es decir, la media de dificultad de los ítems coincidiera con la media de “los síntomas” de los sujetos, y la desviación típica de los ítems fuera similar a la de la muestra estudiada. Dado ese caso nos encontraríamos ante una adaptación ítem-sujeto perfecta o utilizando la nomenclatura utilizada en el análisis de Rasch, un “targeting” perfecto.

Khadka<sup>97</sup> definió este concepto como “la medida en que la dificultad del elemento, definida como el punto en la variable latente en la que la categoría más alta y más baja de respuesta a un ítem tienen la misma probabilidad de ser observados, coincide con el nivel de habilidades visuales de un participante”, esto es, la diferencia entre dificultad promedio de los ítems y la media de los niveles de síntomas de los sujetos. En nuestro trabajo, al haber elegido como caso de estudio una población general con edades comprendidas entre 9 y 30 años en las que la prevalencia de los síntomas asociados a disfunciones binoculares es baja, nos encontramos una muestra sesgada hacia la parte asintomática<sup>157</sup>, circunstancia consistente con lo esperado al tratarse de una población general.

Se realizó el mapa de Wright que consiste en representar gráficamente en una misma escala conjunta expresada en logit, tanto los ítems (de más a menos prevalentes) como los participantes (de menos a más sintomáticos). De manera que los ítems menos prevalentes y

## MÉTODOS

los sujetos más sintomáticos se localizan en la parte positiva de la escala y viceversa. Una diferencia  $>1$  indicaría una falta de emparejamiento entre ambos.

En nuestro contexto, los ítems más difíciles, que son los que tienden a recibir puntuaciones más bajas, se corresponden con síntomas menos prevalentes, mientras que las personas con mayor capacidad, que son las que consiguen puntuaciones más altas, se corresponderían con las más sintomáticas.

### 4.2.3.3.6 FUNCIONAMIENTO DIFERENCIAL DE LOS ELEMENTOS MUESTRALES (DIF) POR SEXO, GRUPO DE EDAD Y VERSIÓN CISS (CISSv15 VS CISSve)

Se examinó cada ítem para verificar si había alguna diferencia en la forma en que los subgrupos clínicamente relevantes respondieron a cada uno de ellos. Se compararon los siguientes subgrupos:

- Hombres vs mujeres
- Niños menores de 18 años [9-17] vs adultos jóvenes [18-30]

Además, dado que la prueba de DIF es una forma útil de validar las traducciones de los cuestionarios <sup>115,158,159</sup>, se realizó un análisis de DIF adicional para evaluar si los elementos de CISSve eran equivalentes a los de la escala original.

Para realizar el análisis DIF, se utilizó el software WINSTEPS que lo realiza basándose en dos métodos:

- Método de Mantel-Haenszel para estimar probabilidades, tamaño de DIF y significación de tablas cruzadas de observaciones en los dos grupos. Estima una medida del tamaño del DIF en logits (log odds) y su significancia, a través de observaciones en los dos grupos, como se observa en la Figura 4.11 <sup>107</sup>. Para la clasificación se utilizó el criterio del Educational Testing Service (ETS) <sup>160,161</sup>.

Categoría de DIF (ETS)	con "Contraste DIF" y diferencia estadísticamente significativa de DIF	
<b>C = Moderado a alto</b>	$ DIF  \geq 0,64$ logits	<b>prob (<math> DIF  \leq 0,43</math> logits <math>\leq 0.05</math> (doble cola)</b> <i>aproximadamente <math> DIF  \geq 0,43</math> logits + 2*DIF error std.</i>
<b>B = Pequeño a moderado</b>	$ DIF  \geq 0,43$ logits	<b>prob (<math> DIF  = 0</math> logits <math>\leq 0.05</math> (doble cola)</b> <i>aproximadamente <math> DIF  &gt; 2*DIF</math> error std.</i>
<b>A = Insignificante</b>		
C-, B- = DIF respecto al grupo focal   C+, B+ = DIF respecto al grupo de referencia		

FIGURA 4.11 – ESTIMACIÓN DEL TAMAÑO DEL DIF EN LOGITS SEGÚN MANTEL-HAENSZEL <sup>107</sup> (WINSTEPS)

- Método de la diferencia de logit (regresión logística). Es un método de clasificación en el que se estima si son diferentes o no los comportamientos a través de la diferencia entre las dificultades que da el modelo de Rasch para cada ítem en cada uno de los dos grupos, manteniendo todo lo demás constante <sup>94,107</sup>

El contraste DIF (es decir, la diferencia entre la dificultad del ítem entre los dos grupos) se define de acuerdo al criterio descrito en la Figura 4.12 <sup>114,123,162</sup>

NO DIF	Si CONTRASTE DIF < 0.50 logit	
DIF MÍNIMO	Si 0.50 logit < CONTRASTE DIF < 1.0 logit	
DIF NOTABLE	Si CONTRASTE DIF > 1.0 logit	

**FIGURA 4.12 – CRITERIO DE VALORACIÓN SOBRE LA EXISTENCIA DE DIFERENCIA DE DIFICULTAD ENTRE GRUPOS**

La calidad general de los datos psicométricos obtenidos en esta etapa (excepto los niveles de rendimiento) fue evaluada según los criterios de la guía propuesta por Khadka et al.<sup>97</sup> para la evaluación de la calidad de cuestionarios oftalmológicos que podemos encontrar en 11.1 (anexo I).

#### 4.2.3.3.7 ANÁLISIS FACTORIAL

Mediante el paquete SPSS 25.0, se realizó un análisis factorial utilizando el análisis de las componentes principales (PCA) con una rotación Varimax con normalización de Kaiser sobre los datos recopilados en los 438 participantes que respondieron todos los ítems de CISSve.

#### 4.2.3.3.8 VALIDEZ CONVERGENTE Y VALIDEZ DIVERGENTE

Tanto el cuestionario CISSve, como la escala de síntomas de visión asociados al uso de ordenador “Computer-Vision Symptom Scale” (CVSS17)<sup>51,127</sup> y la escala de bienestar mental Warwick-Edinburgh (WEMWBS)<sup>118,163</sup> fueron administrados en papel a todos los participantes, excepto a los niños de primaria, en total a 252 sujetos.

La validez convergente se evaluó estimando el coeficiente de correlación entre las puntuaciones obtenidas por los mismos sujetos en las escalas CISSve y CVSS17. La validez divergente se evaluó a través del coeficiente de correlación entre las puntuaciones de las escalas CISSve y WEMWBS.

Según Khadka et al.<sup>114</sup> un coeficiente de correlación entre las escalas superior a 0.3 se consideraría como una prueba de validez convergente, mientras que un coeficiente de correlación menor, constataría la divergencia entre las escalas.

Al indicar la prueba K-S utilizada en las puntuaciones CISSve, CVSS17 y WEMWBS una distribución no normal para todas las medidas, se tuvo que recurrir a la estadística no paramétrica calculando el coeficiente de correlación (Rho de Spearman) apoyados por el paquete estadístico SPSS 25.0<sup>137</sup>.

#### 4.2.3.3.9 REPETIBILIDAD

Los participantes volvieron a completar el CISSve en una segunda sesión entre 7 y 15 días después.

Para evaluar la repetibilidad de la escala CISSve se utilizó el método de Bland-Altman, calculando posteriormente el coeficiente de repetibilidad ( $COR=1.96*SD$ ), la diferencia media (DM), los límites de acuerdo al 95% ( $DM\pm COR$ ) y el p-valor (t-test muestras pareadas) que muestra si la DM es o no estadísticamente significativa ( $p<0.05$ )<sup>164</sup>.

Adicionalmente se calculó el coeficiente de correlación intraclase (ICC), para la prueba test-retest fijando un intervalo de confianza del 95%.

## MÉTODOS

### 4.2.3.4 APLICACIÓN DE LA ESCALA CISSVE A POBLACIÓN CLÍNICA

#### 4.2.3.4.1 COMPARACIÓN DE LAS RESPUESTAS AL CISSVE DE PACIENTES CON INSUFICIENCIA DE CONVERGENCIA DIAGNOSTICADOS CLÍNICAMENTE CON LOS DE UNA POBLACIÓN GENERAL

En este apartado se compararon las respuestas a la escala CISSve en una población sintomática diagnosticada clínicamente de acuerdo a los criterios considerados por el grupo de investigación “Convergence Insufficiency Treatment Trial” (CITT), es decir, exoforia en la visión próxima ( $\geq 4\Delta$  mayor que en visión lejana), insuficiente amplitud de fusión base temporal en cerca ( $\leq 15\Delta$  en la prueba de vergencia fusional positiva VFP a una distancia de 40 cm ó si falla el criterio de Sheard) y dificultad en las pruebas de punto próximo de convergencia (rotura  $\geq 6$  cm)<sup>74</sup> (Apartado 1.2.1), frente a los de una muestra de población general escogida aleatoriamente, con el único requisito de que estuviera pareada en cuanto a la edad con la población de la muestra sintomática y sin conocer a priori la puntuación alcanzada en la escala CISSve. Los sujetos se seleccionaron de entre los pacientes que contestaron a las preguntas de la escala durante la investigación para realizar la traducción al español y la validación posterior del instrumento.

El análisis realizado se acotó a una población comprendida en el rango de edad de 9-30 años. La elección de este rango de edad queda justificada en base a la coincidencia con la de los grupos de estudio que fueron tomados por los autores de la escala original, CISSv15 cuando se realizó la validación de ese instrumento y su adaptación transcultural<sup>59,60,74</sup>.

La muestra analizada (n=112, M=18.06, DT=6.88 años) contenía 55 pacientes diagnosticados clínicamente, de los cuales 11 fueron excluidos, 3 por no satisfacer los requisitos de edad (9-30 años) y otros 8 por haber contestado de forma arbitraria el test, y el resto fueron tomados aleatoriamente de entre la población general analizada durante la validación de CISSve con los criterios explicados anteriormente. Quedando por lo tanto una muestra sintomática (n=49, M=19,80, DT=9.13 años) y una no sintomática (n=52, M=14.02, DT=6.36 años)

Para soportar los cálculos durante el análisis se utilizaron el paquete estadístico SPSS 25.0<sup>137</sup>, y el paquete para análisis de Rasch Winsteps<sup>107</sup>

Se analizaron las diferencias entre las puntuaciones medias de ambas muestras para comprobar si eran estadísticamente significativas y por lo tanto la escala CISSve es adecuada para identificar correctamente los casos positivos

Utilizando los factores encontrados durante el trabajo de validación de la escala CISSve, se analizó la influencia de cada uno de ellos en la puntuación total de la escala CISSve para cada una de las muestras.

Se estudió si existían diferencias significativas en cuanto al sexo y en cuanto a la edad en cada una de las muestras

Se aplicó el análisis de Rasch con el software Winsteps y se compararon las medidas de cada ítem para cada tipo de población (sintomática diagnosticada clínicamente vs general).

#### 4.2.3.4.2 APLICACIÓN DEL CISSVE EN PACIENTES DIAGNOSTICADOS CLÍNICAMENTE DE INSUFICIENCIA DE CONVERGENCIA, ANTES Y DESPUÉS DE TRATAMIENTO CON TERAPIA VISUAL

En este apartado se analizan las puntuaciones obtenidas en la escala CISSve por una población sintomática diagnosticada clínicamente ( $n=49$ ,  $M=18.47$ ,  $DT=6.65$  años), parte de la cual ( $n=27$ ,  $M=16.22$ ,  $DT=7.10$  años) fue sometida posteriormente a tratamiento con terapia visual. El criterio diagnóstico utilizado para definir que un sujeto presenta IC fue el considerado por el grupo de investigación "Convergence Insufficiency Treatment Trial" (CITT)<sup>74</sup> especificados en el apartado anterior (4.2.3.4.1).

El rango de edad seleccionado fue de 9-30 años. La elección de este rango de edad queda justificada en base a su coincidencia con la de los grupos de estudio que fueron tomados por los autores de la escala original, CISSv15 cuando realizaron la validación de ese instrumento y su adaptación transcultural<sup>59-61,74,165</sup>.

De los 27 pacientes que participaron en el estudio y posteriormente se sometieron a terapia ( $n=27$ ,  $M=16.22$ ,  $DT=7.10$  años), 3 fueron excluidos por no completar el tratamiento, resultando finalmente un espacio muestral ( $n=24$ ,  $M=17.29$ ,  $DT=6.74$  años)

Los pacientes contestaron al CISSve antes y después de la terapia visual con objeto de poder analizar y evaluar el impacto de la misma en la percepción de los síntomas. Se realizó un análisis cluster de todas las respuestas al CISSve (incluyendo las pre y post terapia) para observar si después de la terapia, sujetos antes sintomáticos se desplazaban a la zona no sintomática.

Para soportar los cálculos durante el análisis se utilizó el paquete estadístico SPSS 25.0<sup>137</sup>, y el paquete para análisis de Rasch Winsteps<sup>107</sup>

#### 4.2.3.4.3 COMPARACIÓN DE LA PREVALENCIA DE LAS DIFERENTES TIPOLOGÍAS DE SÍNTOMAS POR GRUPOS DE EDAD, INCLUYENDO UNA MUESTRA DE PACIENTES PRÉSBITAS

En este apartado se analizan las puntuaciones obtenidas en la escala CISSve por una pequeña muestra de sujetos présbitas que participaron en la investigación ( $n=29$ ,  $M=60.86$ ;  $DT=13.82$ ), y que no se habían considerado por estar fuera del rango de edad que cubre este estudio (9-30 años).

Con el objeto de comparar a este grupo (mayores de 40 años) con los otros dos grupos estudiados en este trabajo (niños de 9-17 años y adultos jóvenes de 18 a 30 años), se observó el comportamiento de las puntuaciones de estos sujetos présbitas sobre la escala CISSve en cuanto a las diferentes tipologías de síntomas determinadas por los factores encontrados al realizar el análisis factorial (Apartado 4.2.3.3.7),

Para realizar la comparación entre grupos se agregaron las puntuaciones obtenidas en todos los ítems correspondientes a cada factor para cada grupo de edad, y se representaron gráficamente en un diagrama de barras con el objeto de observar cualitativamente la diferencia en prevalencia de cada tipo de síntomas para cada grupo de edad.

### 4.3 COMPARACIÓN ENTRE LAS PROPIEDADES DE CISSv15 Y CISSve

En este apartado se realizó una comparación entre las propiedades de la escala adaptada a español, CISSve, y las de la versión original del mismo, CISSv15.

Para llevar a cabo esta comparación de una forma sistemática, se siguió el mismo esquema que se describió en la sección 4.2.3.3, utilizando el método de Rasch, apoyando los cálculos principalmente en el paquete Winsteps<sup>107</sup> y en el software estadístico SPSS v25.00. Puntualmente, para algunos cálculos y verificaciones también se utilizaron otras herramientas, los paquetes de “R”, “eRm”, “lrm” y “difR”<sup>138</sup>, Jmetrik<sup>139</sup>, y ConQuest<sup>140</sup>.

La muestra utilizada para analizar la versión adaptada al español, CISSve como indicamos en el apartado de métodos 4.2.3.2, se obtuvo de cuatro fuentes diferentes en Madrid, España. Una escuela de primaria (CEIP Vargas Llosa, Madrid), un instituto de secundaria (IES Juan Rodríguez Villanueva), un centro universitario (Facultad de Óptica y Optometría de la Universidad Complutense de Madrid) y una compañía tecnológica (DXC Technology).

Respecto a la utilizada para el análisis del CISSv15 original en su versión inglesa, se obtuvo de un centro educativo en Swindon (Wiltshire, UK). La edad de los sujetos evaluados estuvo comprendida entre 9 y 30 años y los sujetos no recibieron compensación alguna por su colaboración.

Los criterios de exclusión utilizados fueron los mismos que se emplearon para la muestra en la que se analizó la versión CISSve:

- Lengua materna diferente de la del cuestionario administrado
- Haber sido sometido a una cirugía visual previa (no refractiva)
- Tener una disfunción visual o neurológica
- Tomar medicamentos que pudieran afectar a la visión
- Tener cualquier tipo de disfunción que impidiera al sujeto leer o entender las preguntas del cuestionario

En total, 458 sujetos de habla hispana (edad media, 16.02 años; rango de edad, 9–30 años; población femenina, 54.95%) completaron la versión española del cuestionario CISS (CISSve) y 218 de habla inglesa (edad media, 15.81 años; rango de edad, 12–20 años; población femenina, 37.61%) completaron la versión original del mismo (CISS v15).

La evaluación de los resultados de la comparación entre ambas escalas se realizó de acuerdo a los criterios propuestos por Khadka et al.<sup>114 94</sup>, que se pueden encontrar en 11.1 (anexo I).





# RESULTADOS



## 5 RESULTADOS

En este apartado se describen todos los resultados obtenidos a través de la metodología descrita en la sección anterior correspondiente a los métodos.

### 5.1 ADAPTACIÓN TRANSCULTURAL DE LA ESCALA CISSv15

#### 5.1.1 BLOQUE I – TRADUCCIONES DIRECTAS

Se llevaron a cabo 2 traducciones directas del cuestionario CISSv15. Estas traducciones fueron realizadas por dos traductores independientes bilingües.

Cada traductor, independientemente, partiendo de la versión original en la lengua origen (inglés) obtuvo una versión traducida a la lengua destino (español). Denominamos a cada una de las versiones obtenidas por cada traductor, respectivamente, T1 y T2.

En la Figura 5.1 podemos ver la versión de partida, CISSv15 <sup>60</sup>.

**Convergence Insufficiency Symptom Survey**

Name \_\_\_\_\_ DATE \_\_\_/\_\_\_/\_\_\_

**Clinician Instructions:** Read the following subject instructions and then each item exactly as written. If subject responds with "yes" - please qualify with frequency choices. **Do not give examples.**

**Subject Instructions:** Please answer the following questions about how your eyes feel when reading or doing close work.

		Never	(not very often) Infrequently	Sometimes	Fairly often	Always
1.	Do your eyes feel tired when reading or doing close work?					
2.	Do your eyes feel uncomfortable when reading or doing close work?					
3.	Do you have headaches when reading or doing close work?					
4.	Do you feel sleepy when reading or doing close work?					
5.	Do you lose concentration when reading or doing close work?					
6.	Do you have trouble remembering what you have read?					
7.	Do you have double vision when reading or doing close work?					
8.	Do you see the words move, jump, swim or appear to float on the page when reading or doing close work?					
9.	Do you feel like you read slowly?					
10.	Do your eyes ever hurt when reading or doing close work?					
11.	Do your eyes ever feel sore when reading or doing close work?					
12.	Do you feel a "pulling" feeling around your eyes when reading or doing close work?					
13.	Do you notice the words blurring or coming in and out of focus when reading or doing close work?					
14.	Do you lose your place while reading or doing close work?					
15.	Do you have to re-read the same line of words when reading?					
		x 0	x 1	x 2	x 3	x 4

**TOTAL SCORE** \_\_\_\_\_

FIGURA 5.1 – ESCALA CISSV15

La Figura 5.2, muestra la primera traducción directa, realizada por CP, perteneciente al grupo de investigación, nacionalidad española y nivel bilingüe de inglés (versión T1).

**Encuesta sobre síntomas de insuficiencia de convergencia**

Nombre \_\_\_\_\_ Fecha \_\_/\_\_/\_\_

**Instrucciones al clínico:** Lea las siguientes instrucciones y después cada una de las preguntas exactamente cómo están escritas. Si el sujeto responde "sí" – por favor califique con una de las opciones de frecuencia

**No ponga ejemplos.**

**Instrucciones al sujeto:** Por favor, conteste las siguientes preguntas sobre cómo se encuentran sus ojos mientras lee o trabaja de cerca

**PREGUNTAS**

1. ¿Siente sus ojos cansados cuando lee o trabaja de cerca?
2. ¿Siente incomodidad en sus ojos cuándo lee o trabaja de cerca?
3. ¿Tiene Dolores de cabeza cuándo lee o trabaja de cerca?
4. ¿Se siente adormecido cuándo lee o trabaja de cerca?
5. ¿Pierde concentración cuándo lee o trabaja de cerca?
6. ¿Tiene problemas para recordar lo que ha leído?
7. ¿Experimenta vision doble cuando lee o trabaja de cerca?
8. ¿Ve palabras que se mueven, saltan, nadan o parecen flotar en la página cuando lee o trabaja de cerca?
9. ¿Le da la impresión de que lee lento?
10. ¿Le duelen alguna vez los ojos cuando lee o trabaja de cerca?
11. ¿Se le inflaman los ojos alguna vez cuando lee o trabaja de cerca?
12. ¿Siente sensación de tirantez alrededor de los ojos cuando lee o trabaja de cerca?
13. ¿Percibe las palabras borrosas ó que se enfocan y desenfocan cuando lee o trabaja de cerca?
14. ¿Se "pierde de línea" cuando lee o trabaja de cerca?
15. ¿Tiene que volver a leer la misma línea de palabras al leer?

**RESPUESTAS**

- Nunca
- (No muy a menudo) No es frecuente
- A veces
- Con bastante frecuencia
- Siempre

\_\_ x 0 \_\_ x 1 \_\_ x 2 \_\_ x 3 \_\_ x 4

**RESULTADO TOTAL** \_\_\_\_\_

FIGURA 5.2 – TRADUCCIÓN T1

La Figura 5.3, muestra la segunda traducción directa, en este caso realizada por un profesional del sector bancario, de nacionalidad española, bilingüe en inglés y ajeno al campo de estudio objeto de la escala CISS. Esta traducción se realizó de forma totalmente independiente a T1, tal y como marcan las buenas prácticas definidas por Gjersing, Caplehorn y Clausen <sup>122</sup>.

**Encuesta sobre síntomas de insuficiencia de convergencia**

Nombre \_\_\_\_\_ Fecha \_\_\_/\_\_\_/\_\_\_

**Instrucciones al profesional clínico:** Lea las siguientes instrucciones y después cada una de las preguntas tal y cómo están escritas. Si el sujeto responde "sí" – por favor califique con una opción de frecuencia

**No ponga ejemplos.**

**Instrucciones al paciente:** Por favor, conteste las siguientes preguntas sobre su sensación al leer o trabajar de cerca

**PREGUNTAS**

1. ¿Tiene sensación de cansancio en sus ojos al leer o trabajar de cerca?
2. ¿Tiene sensación de incomodidad en sus ojos al leer o trabajar de cerca?
3. ¿Tiene Dolores de cabeza al leer o trabajar de cerca?
4. ¿Tiene sensación de somnolencia al leer o trabajar de cerca?
5. ¿Pierde la concentración al leer o trabajar de cerca?
6. ¿Tiene problemas para acordarse de lo que ha leído?
7. ¿Ve doble al leer o trabajar de cerca?
8. ¿Ve palabras moverse, saltar, nadar o que parecen flotar en la página al leer o trabajar de cerca?
9. ¿Le parece que lee lento?
10. ¿Le duelen alguna vez los ojos al leer o trabajar de cerca?
11. ¿Se le inflaman los ojos alguna vez al leer o trabajar de cerca?
12. ¿Siente sensación de tirantez alrededor de los ojos al leer o trabajar de cerca?
13. ¿Percibe las palabras borrosas ó que van y vienen al leer o trabajar de cerca?
14. ¿Se "pierde de renglón" al leer o trabajar de cerca?
15. ¿Tiene que volver a leer la misma línea de palabras al leer?

**PUNTUACIÓN**

- Nunca
- (No muy a menudo) No es frecuente
- A veces
- Con bastante frecuencia
- Siempre

**PUNTUACIÓN TOTAL** \_\_\_\_\_

FIGURA 5.3 – TRADUCCIÓN T2

Como se aprecia, las dos traducciones directas presentan diferencias significativas tanto en el texto explicativo, como en varios de los ítems e incluso en las opciones de respuesta.

Llegados a este punto, el siguiente paso para alinear las versiones consistió en realizar una reunión entre los dos traductores independientes, moderada por el grupo de investigación en la que se abrió una discusión sobre las discrepancias y en la que se obtuvo T12, la versión consensuada en español de T1 y T2. Antes de realizar la reunión con los traductores, el equipo de investigación se reunió internamente para evaluar algunos criterios y preparar la sesión para que fuera efectiva.

Se transcribieron los documentos T1 y T2 a una hoja Excel estructurada de la siguiente forma:

- Se separaron la parte explicativa de los ítems, y también de las opciones de respuesta. En la primera columna se situó el contenido correspondiente al contenido de T1 y en la segunda el correspondiente a T2. En una tercera columna se situaron los comentarios y observaciones para que permitieran evaluar durante la reunión ciertos puntos a considerar importantes, con el objeto de que se tuvieran en cuenta. Por ejemplo, se discutió respecto a la idoneidad del “trato de usted” frente al lenguaje coloquial.
- Se anotó la necesidad de adecuación de ciertos ítems de la pura “literalidad” a un texto más inteligible y apropiado en español en expresiones difíciles de transcribir exactamente y donde se debe buscar que el sentido de la pregunta se correspondiera plenamente en español respecto a la versión inglesa, en lugar de prevalecer la literalidad que podría causar confusión. Podemos ver un ejemplo en el ítem 12, “Do you feel a "pulling" feeling around your eyes when reading or doing close work?”.
- También se marcaron en la hoja Excel las diferencias entre las dos versiones traducidas, T1 y T2, de forma que se facilitase la identificación de las mismas durante la reunión y así se asegurara que se llevaba a cabo la discusión y se decidiera sobre los aspectos señalados. En la Figura 5.4 se observa una captura de la hoja que fue utilizada con el objetivo de facilitar que la sesión de consenso de T12 se realizara de forma efectiva. Previamente a la reunión, se marcaron en rojo las partes en las que aparecían discrepancias significativas entre T1 y T2, en amarillo aquellas en las que aunque el significado fuera el mismo, había alguna palabra que a pesar de ser sinónima no coincidía y por último, en verde se señalaron las partes donde a pesar de no haber coincidencia absoluta, la diferencia entre las dos versiones era mínima (alguna preposición, la parte final en cada ítem “al leer o trabajar de cerca”, etc)

Encuesta sobre síntomas de insuficiencia de convergencia

Nombre \_\_\_\_\_ Fecha \_\_/\_\_/\_\_

Instrucciones al profesional clínico: Lea las siguientes instrucciones y después cada una de las preguntas tal y cómo están escritas.

Si el sujeto responde "sí" – por favor califique con una opción de frecuencia

No ponga ejemplos.

Instrucciones al paciente: Por favor, conteste las siguientes preguntas sobre su sensación al leer o trabajar de cerca

- Nunca
- (No muy a menudo) inusualmente
- A veces
- Bastante a menudo
- Siempre

1. ¿Tiene sensación de cansancio en sus ojos al leer o trabajar de cerca?
2. ¿Tiene sensación de incomodidad en sus ojos al leer o trabajar de cerca?
3. ¿Tiene Dolores de cabeza al leer o trabajar de cerca?
4. ¿Tiene sensación de somnolencia al leer o trabajar de cerca?
6. ¿Tiene problemas para acordarse de lo que ha leído?
7. ¿Ve doble al leer o trabajar de cerca?
8. ¿Ve palabras moverse, saltar, nadar o que parecen flotar en la página al leer o trabajar de cerca?
9. ¿Le parece que lee lento?
10. ¿Le duelen alguna vez los ojos al leer o trabajar de cerca?
11. ¿Se le inflaman los ojos alguna vez al leer o trabajar de cerca?
12. ¿Siente sensación de tirantez alrededor de los ojos al leer o trabajar de cerca?
13. ¿Percibe las palabras borrosas ó que van y vienen al leer o trabajar de cerca?
14. ¿Se "pierde de renglón" al leer o trabajar de cerca?
15. ¿Tiene que volver a leer la misma línea de palabras al leer?

\_\_ x 0 \_\_ x 1 \_\_ x 2 \_\_ x 3 \_\_ x 4

FIGURA 5.4 – PREPARATIVOS DE LA SESIÓN DE CONSENSO PARA OBTENCIÓN DE T12

Reunidos, el grupo de investigación y los traductores, se realizó una sesión de trabajo en la que se fueron revisando minuciosamente las diferencias entre ambas versiones para llegar a un acuerdo.

Se optó por el trato coloquial frente al formal por dos razones. En primer lugar, porque la población en la que se pretendía realizar el estudio estaba comprendida en una franja de edad en la que es el lenguaje que predomina, por lo que la inteligibilidad del mismo se maximizaría con esta premisa. El otro factor que sustentó esa decisión fue el hecho de que el lenguaje coloquial, en español crea más cercanía entre paciente y personal clínico, con lo que se espera que el cuestionario resulte más fácil de leer para el paciente y le predisponga más positivamente a contestar.

Podemos ver en la Figura 5.5, la versión T12, compendio de T1 y T2 tras la revisión y aprobación por parte del equipo de investigación.

**Escala sobre síntomas de Insuficiencia de Convergencia**

Nombre \_\_\_\_\_ Fecha \_\_/\_\_/\_\_

**Instrucciones al clínico:** Lee las siguientes instrucciones y después de cada pregunta tal y como está escrita Si el paciente responde "sí", por favor califica la respuesta con una opción de frecuencia. No pongas ejemplos

**Instrucciones al paciente:** Las siguientes preguntas se refieren a cómo te sientes mientras lees o trabajas de cerca

	Nunca	Rara Vez	A veces	Bastante a menudo	Siempre
1. ¿Notas tus ojos cansados?					
2. ¿Notas incomodidad en tus ojos?					
3. ¿Te duele la cabeza?					
4. ¿Te entra sueño?					
5. ¿Pierdes la concentración?					
6. ¿Te cuesta recordar lo que has leído?					
7. ¿Ves doble?					
8. ¿Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?					
9. ¿Te parece que lees lento?					
10. ¿Te duelen los ojos?					
11. ¿Se te irritan los ojos?					
12. ¿Tienes sensación de "tirantez" alrededor de los ojos?					
13. ¿Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?					
14. ¿Te pierdes de línea al leer?					
15. ¿Tienes que releer la misma línea de texto?					

FIGURA 5.5 – TRADUCCIÓN DIRECTA CONSENSUADA T12

La sesión se realizó por Skype, compartiendo la pantalla, con el objeto de proveer a todos los revisores visibilidad simultanea sobre el contenido, facilitando así las revisiones, discusiones y consensos. La experiencia resultó totalmente satisfactoria para todos los miembros del equipo

## RESULTADOS

consiguiéndose máxima eficiencia y facilitándose enormemente la compartición del material. Así lo manifestaron varios miembros del equipo durante la reunión.

Se tomaron notas de las discusiones llevadas a cabo durante la reunión y se enviaron minutas a los asistentes, dejando constancia de todos los acuerdos y del entregable alcanzado, la versión T12 (traducción directa consensuada).

### 5.1.2 BLOQUE II – TRADUCCIONES INVERSAS

Partiendo del entregable de la fase anterior, T12, la versión consensuada de T1 y T2, dos traductores nativos ingleses bilingües en español llevaron a cabo dos traducciones inversas de T12, es decir, de español a inglés (versión originaria de CISSv15). Las traducciones se realizaron de forma independiente con el objeto de no tener ningún tipo de condicionamiento.

Uno de los dos traductores fue seleccionado por su experiencia en la traducción de textos clínicos y su conocimiento sobre la praxis clínica, mientras que el otro se escogió por pertenecer a un ámbito ajeno al campo sobre el que versa el cuestionario.

La Figura 5.6 muestra las dos traducciones independientes, BT1 Y BT2 generadas en esta fase:

TRADUCTOR 1 (AB)	TRADUCTOR 2 (PS)
<b>SURVEY DESCRIPTION</b>	<b>SURVEY DESCRIPTION</b>
Convergence insufficiency symptom scale	Scale of symptoms of insufficient convergence
<b>INSTRUCTIONS</b>	<b>INSTRUCTIONS</b>
Instructions. Please respond to the following questions about how you feel when you are reading or doing other close work	Instructions for the clinic: Read the following instructions and then each question exactly how it is written. If the patient responds yes, please qualify the answer with how frequently. Don't put examples. Instructions for the patient: The following questions refer to how you feel while you read or work close-up
<b>ITEMS</b>	<b>ITEMS</b>
1. Do your eyes feel tired?	1. Do you feel your eyes are tired?
2. Do you have discomfort?	2. Do you feel discomfort in your eyes?
3. Do you get headaches?	3. Does your get headaches?
4. Do you get tired?	4. Do you feel tired?
5. Do you have difficulty concentrating?	5. Do you lose concentration?
6. Do you find it hard to remember what you've read?	6. Do you find it difficult to remember what you have read?
7. Do you see double?	7. Do you see double? (double vision)
8. Does it feel like the words move, mix together or float above the text?	8. Do you think the words move, they mix or float over the text?
9. Do you feel like you read slowly?	9. Do you think you read slowly?
10. Do your eyes hurt?	10. Do your eyes hurt?
11. Do your eyes itch?	11. Do your eyes get irritated?
12. Do you feel tightness around the eyes?	12. Do you have the sensation of tightness around the eyes?
13. Do the words blur or come in and out of focus?	13. Do you notice the words go blurry or focus and defocus
14. Do you lose your line?	14. Do you lose the line you are reading?
15. Do you have to re-read the same line?	15. Do you re-read the same line of text?
<b>REPLY CATEGORIES</b>	<b>REPLY CATEGORIES</b>
1. Never	1. Never
2. Hardly ever	2. Rarely
3. Sometimes	3. Sometimes
4. Often	4. Frequently
5. Always	5. Always

FIGURA 5.6 – TRADUCCIONES INVERSAS – BT1 Y BT2

A continuación, siguiendo las recomendaciones de Gjersing, Caplehorn y Clausen <sup>122</sup> se celebró una reunión a la que asistieron los traductores nativos ingleses bilingües en español que generaron las versiones BT1 y BT2 y el equipo de investigación, con el objeto de consensuar una versión inversa (BT12) como fruto de discutir las diferencias en cada una de las dos traducciones (BT1 y BT2).

Se encontraron pequeñas diferencias entre las traducciones de algunos de los ítems tal como se puede apreciar en la Figura 5.6. Para cada ítem, se acordó qué versión debía prevalecer. La mayoría de ellos, concretamente el 1,2,3,4,5,6, 7, 9, 10, 12, 14 y 15, fueron prácticamente idénticos en ambas traducciones. En algunos casos incluso resultaron ser idénticos, como por ejemplo el 7 y el 10, y en otros sólo existía algún pequeño matiz que los diferenciaba.

Como ejemplo, en algunos casos se cambia el verbo utilizado, “feel”-“have”, “get”-“have”, pero siempre con significado equivalente en el contexto en el que se utilizó, o se alteraron algunos adjetivos, “hard”-“difficult”, totalmente sinónimos, etc. En todo caso, en estos ítems hubo una coincidencia prácticamente total a todos los niveles (lingüístico, semántico y conceptual) en ambas retro traducciones.

El criterio tomado para elegir uno u otro consistió en hacer prevalecer el que transmitía más fielmente el contenido semántico de T12, y en caso de presentar equivalencia, el que presentó más literalidad en comparación con la versión origen, siempre en consenso y habiendo acuerdo entre todos los miembros del equipo.

En resumen. En todos estos ítems rápidamente se llegó a un acuerdo y siempre, ambos traductores coincidieron en que la traducción final escogida era la más idónea.

Sin embargo, los ítems 8, 11 y 13, presentaron diferencias significativas. Resultaron ser equivalentes semántica y conceptualmente, pero no idiomáticamente, “itch”-“irritated”, “mix”-“mix together”, “in and out of focus”-“blurry or of focus and defocus”. Por esta razón, se abrió el debate para decidir qué retro-traducción reflejaba mejor el significado de la versión origen T12. En todos estos casos, se optó finalmente por una versión revisada con elementos de ambas retro-traducciones, BT1 y BT2, pero distinta a las dos.

En la Figura 5.7 se puede apreciar el proceso a partir del cual se llegó a consenso. Sobre la tabla que contiene las dos traducciones inversas (BT1 y BT2), aparecen marcadas en verde las versiones de los ítems traducidos que fueron seleccionadas durante la revisión, y en naranja las que fueron consensuadas durante la misma. La Figura 5.8, muestra la versión BT12 acordada tras la reunión.

RESULTADOS

BT1 - Back translation (AB)	BT2 - Back translation (PS)
<b>Convergence insufficiency symptom scale</b>	<b>Scale of symptoms of insufficient convergence</b>
DT Item	DT Item
<b>Instructions.</b> Please respond to the following questions about how you feel when you are reading or doing other close work	<b>Instructions for the clinic:</b> Read the following instructions and then each question exactly how it is written. If the patient responds yes, please qualify the answer with how frequently. Don't put examples. <b>Instructions for the patient:</b> The following questions refer to how you feel while you read or work close-up.
1. Do your eyes feel tired?	1. Do you feel your eyes are tired?
2. Do you have discomfort?	2. Do you feel discomfort in your eyes?
3. Do you get headaches?	3. Does your get headaches?
4. Do you get tired?	4. Do you feel tired?
5. Do you have difficulty concentrating?	5. Do you lose concentration?
6. Do you find it hard to remember what you've read?	6. Do you find it difficult to remember what you have read?
7. Do you see double?	7. Do you see double? (double vision)
8. Does it feel like the words move, mix together or float above the text?	8. Do you think the words move, they mix or float over the text?
9. Do you feel like you read slowly?	9. Do you think you read slowly?
10. Do your eyes hurt?	10. Do your eyes hurt?
11. Do your eyes itch?	11. Do your eyes get irritated?
12. Do you feel tightness around the eyes?	12. Do you have the sensation of tightness around the eyes?
13. Do the words blur or come in and out of focus?	13. Do you notice the words go blurry or focus and defocus
14. Do you lose your line?	14. Do you lose the line you are reading?
15. Do you have to re-read the same line?	15. Do you re-read the same line of text?
<b>REPLY CATEGORIES</b>	<b>REPLY CATEGORIES</b>
1. Never	1. Never
2. Hardly ever	2. Rarely
3. Sometimes	3. Sometimes
4. Often	4. Frequently
5. Always	5. Always

FIGURA 5.7 – DIFERENCIAS ENTRE BT1 Y BT2

## BT12 - Agreement

<b>Convergence insufficiency symptom scale</b>
DT Item
<b>INSTRUCTIONS.</b> Please respond to the following questions about how you feel when you are reading or doing other close work
1. Do your eyes feel tired?
2. Do you have discomfort in your eyes?
3. Do you get headaches?
4. Do you feel tired?
5. Do you lose your concentration?
6. Do you find it difficult to remember what you have read?
7. Do you see double?
8. Does it feel like the words move, merge or float on the page?
9. Do you feel like you read slowly?
10. Do your eyes hurt?
11. Do your eyes feel sore?
12. Do you feel tightness around the eyes?
13. Do you notice the words blur or come in and out of focus?
14. Do you lose the line you are reading?
15. Do you have to re-read the same line?
<b>REPLY CATEGORIES</b>
1. Never
2. Rarely
3. Sometimes
4. Often
5. Always

FIGURA 5.8 – TRADUCCIÓN INVERSA CONSENSUADA (BT12)

## 5.1.3 BLOQUE III – REUNIÓN DE EXPERTOS

El siguiente paso según las recomendaciones de Gjersing, Caplehorn y Clausen <sup>122</sup> consistió en una reunión en la que un panel de expertos, entre los que se encontraban todos los traductores que generaron las versiones T1, T2, BT1 y BT2, el equipo de investigación, y otros expertos en la materia para ayudar en la evaluación de la escala, valoraron todas las traducciones y realizaron una comparación entre ambas para asegurar la equivalencia semántica conceptual e idiomática.

Con el objeto de asegurar que no se omitiera ninguna cuestión importante durante la reunión de expertos, de la que saldría la versión preliminar de la escala CISSve, el equipo de investigación realizó un trabajo previo, aunando ambas traducciones independientes en un único documento, e identificando las diferencias de forma que la reunión pudiera ser conducida de forma efectiva y cubriera todas las discrepancias entre BT1 y BT2, consiguiendo así una optimización del tiempo de los participantes a la vez que una cobertura total de los puntos de agenda.

La reunión se celebró utilizando una herramienta de colaboración (Skype), compartiendo la pantalla para facilitar la presencia de algunos de los miembros localizados fuera de Madrid que atendieron la reunión de forma remota. El equipo de investigación y uno de los traductores realizaron la revisión conjuntamente de forma presencial en el Departamento de Optometría y Visión de la Facultad de Óptica y Optometría.

## RESULTADOS

Durante la sesión, el equipo discutió la idoneidad de cada ítem propuesto por cada una de las dos versiones BT1 y BT2 y se clasificaron según su equivalencia conceptual con la versión original inglesa. Como resultado se marcaron los ítems según su grado de equivalencia de acuerdo a la tabla de la Figura 5.9 que resume la categorización propuesta por Praena Crespo et al.<sup>166</sup>.

Equivalencia A	La traducción inversa es lingüísticamente y/o semánticamente equivalente a la pregunta de la versión original	
Equivalencia B	La traducción inversa es lingüísticamente y/o semánticamente equivalente a la pregunta de la versión original, pero existen una o dos palabras con distinto significado	
Equivalencia C	Los ítems expresan un significado diferente, por tanto no son equivalentes	

**FIGURA 5.9 – DIFERENCIAS ENTRE BT1 Y BT2**

Reunidos, las dos traducciones inversas fueron comparadas (equivalencia semántica, conceptual, e idiomática). Cada uno de los ítems fue clasificado según su equivalencia conceptual de acuerdo con la Figura 5.10 y se señalaron las diferencias con el objetivo de llegar a un consenso en los ítems que presentaban diferencias.

BT1 (Back Translation 1)	BT2 (Back Translation 2)	BT1 vs BT2
1. Do your eyes feel tired?	1. Do you feel your eyes are tired?	
2. Do you have discomfort?	2. Do you feel discomfort in your eyes?	
3. Do you get headaches?	3. Does your get headaches?	
4. Do you get tired?	4. Do you feel tired?	
5. Do you have difficulty concentrating?	5. Do you lose concentration?	
6. Do you find it hard to remember what you've read?	6. Do you find it difficult to remember what you have read?	
7. Do you see double?	7. Do you see double? (double vision)	
8. Does it feel like the words move, mix together or float above the text?	8. Do you think the words move, they mix or float over the text?	
9. Do you feel like you read slowly?	9. Do you think you read slowly?	
10. Do your eyes hurt?	10. Do your eyes hurt?	
11. Do your eyes itch?	11. Do your eyes get irritated?	
12. Do you feel tightness around the eyes?	12. Do you have the sensation of tightness around the eyes?	
13. Do the words blur or come in and out of focus?	13. Do you notice the words go blurry or focus and defocus	
14. Do you lose your line?	14. Do you lose the line you are reading?	
15. Do you have to re-read the same line?	15. Do you re-read the same line of text?	
REPLY CATEGORIES	REPLY CATEGORIES	
1. Never	1. Never	
2. Hardly ever	2. Rarely	
3. Sometimes	3. Sometimes	
4. Often	4. Frequently	
5. Always	5. Always	

**FIGURA 5.10 – COMPARACIÓN ENTRE ÍTEMS DE BT1 Y BT2 REALIZADA POR EL EQUIPO DE INVESTIGACIÓN**

CISS-V15 (Original)	T12 (Direct Translation - Consensuada)	BT1 (Back Translation 1)	BT2 (Back Translation 2)	BT12 (Back traslation - Consensuada)
<p><b>Convergence Insufficiency Symptom Survey</b></p> <p><b>Instructions.</b> Please answer the following questions about how your eyes feel when reading or doing close work</p>	<p><b>Escala sobre síntomas de insuficiencia de convergencia</b></p> <p><b>Instrucciones.</b> Las siguientes preguntas se refieren a cómo se sienten mientras lees o trabajas de cerca</p>	<p><b>Convergence insufficiency symptom scale</b></p> <p><b>Instructions.</b> Please respond to the following questions about how you feel when you are reading or doing other close work</p>	<p><b>Scale of symptoms of insufficient convergence</b></p> <p><b>Instructions for the clinic:</b> Read the following instructions and then each question exactly how it is written. If the patient responds yes, please qualify the answer with how frequently. Don't put examples. <b>Instructions for the patient.</b> The following questions refer to how you feel while you read or work close-up.</p>	<p><b>Convergence insufficiency symptom scale</b></p> <p><b>Instructions.</b> Please respond to the following questions about how you feel when you are reading or doing other close work</p>
1. Do your eyes feel tired when reading or doing close work?	1. ¿Notas tus ojos cansados?	1. Do your eyes feel tired?	1. Do you feel your eyes are tired?	1. Do your eyes feel tired?
2. Do your eyes feel uncomfortable when reading or doing close work?	2. ¿Notas incomodidad en tus ojos?	2. Do you have discomfort?	2. Do you feel discomfort in your eyes?	2. Do you have discomfort in your eyes?
3. Do you have headaches when reading or doing close work?	3. ¿Te duele la cabeza?	3. Do you get headaches?	3. Does your get headaches?	3. Do you get headaches?
4. Do you feel sleepy when reading or doing close work?	4. ¿Te entra sueño?	4. Do you get tired?	4. Do you feel tired?	4. Do you feel tired?
5. Do you lose concentration when reading or doing close work?	5. ¿Pierdes la concentración?	5. Do you have difficulty concentrating?	5. Do you lose concentration?	5. Do you lose your concentration?
6. Do you have trouble remembering what you have read?	6. ¿Te cuesta recordar lo que has leído?	6. Do you find it hard to remember what you've read?	6. Do you find it difficult to remember what you have read?	6. Do you find it difficult to remember what you have read?
7. Do you have double vision when reading or doing close work?	7. ¿Ves doble?	7. Do you see double?	7. Do you see double? (double vision)	7. Do you see double?
8. Do you see the words move, jump, swim or appear to float on the page when reading or doing close work?	8. ¿Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?	8. Does it feel like the words move, mix together or float above the text?	8. Do you think the words move, they mix or float over the text?	8. Does it feel like the words move, merge or float on the page?
9. Do you feel like you read slowly?	9. ¿Te parece que lees lento?	9. Do you feel like you read slowly?	9. Do you think you read slowly?	9. Do you feel like you read slowly?
10. Do your eyes ever hurt when reading or doing close work?	10. ¿Te duelen los ojos?	10. Do your eyes hurt?	10. Do your eyes hurt?	10. Do your eyes hurt?
11. Do your eyes ever feel sore when reading or doing close work?	11. ¿Se te irritan los ojos?	11. Do your eyes itch?	11. Do your eyes get irritated?	11. Do your eyes feel sore?
12. Do you feel a "pulling" feeling around your eyes when reading or doing close work?	12. ¿Tienes sensación de "tirantez" alrededor de los ojos?	12. Do you feel tightness around the eyes?	12. Do you have the sensation of tightness around the eyes?	12. Do you feel tightness around the eyes?
13. Do you notice the words blurring or coming in and out of focus when reading or doing close work?	13. ¿Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?	13. Do the words blur or come in and out of focus?	13. Do you notice the words go blurry or focus and defocus	13. Do you notice the words blur or come in and out of focus?
14. Do you lose your place while reading or doing close work?	14. ¿Te pierdes de línea al leer?	14. Do you lose your line?	14. Do you lose the line you are reading?	14. Do you lose the line you are reading?
15. Do you have to re-read the same line of words when reading?	15. ¿Tienes que releer la misma línea de texto?	15. Do you have to re-read the same line?	15. Do you re-read the same line of text?	15. Do you have to re-read the same line?
<b>REPLY CATEGORIES</b>	<b>CATEGORIAS DE RESPUESTA</b>	<b>REPLY CATEGORIES</b>	<b>REPLY CATEGORIES</b>	<b>REPLY CATEGORIES</b>
1. Never	1. Nunca	1. Never	1. Never	1. Never
2. Infrequently	2. Muy pocas veces	2. Hardly ever	2. Rarely	2. Rarely
3. Sometimes	3. Algunas veces	3. Sometimes	3. Sometimes	3. Sometimes
4. Fairly Often	4. Muchas veces	4. Often	4. Frequently	4. Often
5. Always	5. Siempre	5. Always	5. Always	5. Always

FIGURA 5.11 – REVISIÓN CONJUNTA DE TODAS LAS TRADUCCIONES POR PARTE DEL PANEL DE EXPERTOS

## RESULTADOS

Se observó que en los ítems 10 y 11 de la escala CISSv15 en su versión original aparecía “ever” en el enunciado. Situación que levantó un coloquio, y controversia entre algunos miembros del panel, ya que “ever” es un adverbio que en el contexto en el que se utiliza significa “alguna vez” y la frecuencia es algo que debería quedar teóricamente desligado del enunciado del ítem, ya que precisamente el objeto de la escala Likert elegida para captar el grado de sintomatología trata de gradar la frecuencia con la que ocurren esos síntomas. El introducir un “ever” tendría sentido en un ítem dicotómico en el que se esperase una respuesta “sí” ó “no”, pero no en uno del que se espera una respuesta sobre una variable ordinal con 5 categorías de respuesta diferentes que indican diferentes grados de sintomatología.

En la Figura 5.12 se puede observar la tabla generada con la puntuación asignada a cada uno de los ítems de acuerdo a la clasificación descrita en la Figura 5.9 y tras la evaluación del panel de expertos.

TRADUCCION INVERSA CONSENSUADA - BT12	CISS-15	CLASIFICACIÓN ITEMS - PANEL
SURVEY DESCRIPTION		SURVEY DESCRIPTION
Convergence insufficiency symptom scale	Convergence Insufficiency Symptom Survey	N/A
INSTRUCTIONS		INSTRUCTIONS
Instructions. Please respond to the following questions about how you feel when you are reading.		Instructions: Please answer the following questions about how your eyes feel when you are reading. (No literal pero semánticamente equivalente)
ITEMS	ITEMS	ITEMS
1. Do your eyes feel tired?	Do your eyes feel tired when reading or doing close work?	A
2. Do you have discomfort in your eyes?	Do your eyes feel uncomfortable when reading or doing close work?	A
3. Do you get headaches?	Do you have headaches when reading or doing close work?	A
4. Do you feel tired?	Do you feel sleepy when reading or doing close work?	A (No literal pero semánticamente equivalente)
5. Do you lose your concentration?	Do you lose concentration when reading or doing close work?	A
6. Do you find it difficult to remember what you have read?	Do you have trouble remembering what you have read?	A
7. Do you see double?	Do you have double vision when reading or doing close work?	A
8. Does it feel like the words move, merge or float on the page?	Do you see the words move, jump, swim or appear to float on the page when reading?	A (No literal pero semánticamente equivalente)
9. Do you feel like you read slowly?	Do you feel like you read slowly?	A
10. Do your eyes hurt?	Do your eyes ever hurt when reading or doing close work?	A
11. Do your eyes feel sore?	Do your eyes ever feel sore when reading or doing close work?	A
12. Do you feel tightness around the eyes?	Do you feel a "pulling" feeling around your eyes when reading or doing close work?	A (No literal pero semánticamente equivalente)
13. Do you notice the words blur or come in and out of focus?	Do you notice the words blurring or coming in and out of focus when reading?	A
14. Do you lose the line you are reading?	Do you lose your place while reading or doing close work?	A (No literal pero semánticamente equivalente)
15. Do you have to re-read the same line?	Do you have to re-read the same line of words when reading?	A
1. Never	Never	A
2. Rarely	Infrequently	A (No literal pero semánticamente equivalente)
3. Sometimes	Sometimes	A
4. Often	Fairly Often	A (No literal pero semánticamente equivalente)
5. Always	Always	A

**FIGURA 5.12 – REVISIÓN DE LA EQUIVALENCIA DE LAS TRADUCCIONES LLEVADA A CABO POR EL PANEL DE EXPERTOS Y CATEGORIZACIÓN DE LOS ÍTEMES DE BT12 SEGÚN SU CORRESPONDENCIA CON LA VERSIÓN ORIGINAL**

Como se puede apreciar, pese a las pequeñas diferencias en algunas de las palabras no esenciales dentro de la descripción de algunos de los ítems, todos se encontraron lingüística y semánticamente equivalentes y por lo tanto se clasificaron con la “Equivalencia A”, quedando justificada esta calificación con un comentario significando la sinonimia entre el vocablo presentado en la traducción inversa consensuada, BT12, y la versión original, CISSv15.

Finalmente, se optó por traducciones de ítems provenientes de ambos traductores, 8 del primer traductor y 6 del segundo respectivamente. También se ajustaron otros 3, llegando a acuerdo durante la reunión (Figura 5.11)

Dada esa situación y tras revisar una vez más el proceso completo y el contenido semántico de la traducción directa consensuada, T12 (Figura 5.5), verificando que todos los ítems guardaban los principios de claridad y concreción necesarios, se decidió seguir adelante con esta versión como versión “pre-final” adaptada a español de la escala CISSv15. A esta se le denominó versión preliminar del CISS - versión española (CISSve).

#### 5.1.4 PRE-TEST

Con objeto de localizar cualquier posible inconsistencia y asegurar que la interpretación que hacen los sujetos sobre los ítems del CISSve no ofrece dudas, se realizó un pre-test siguiendo las recomendaciones de Gjersing, Caplehorn y Clausen <sup>122</sup> y Beaton et al <sup>124</sup>. Se aplicó la versión preliminar obtenida en la fase anterior, correspondiente con la traducción consensuada T12 a un grupo de 83 personas de distinto nivel educativo para valorar el comportamiento inicial de la escala.

En este punto se añadió la opción “no entiendo la pregunta” a las respuestas de cada ítem con objeto de evaluar posibles ítems confusos y valorar su sustitución o la supresión de cualquier expresión dentro de alguno de ellos que pudiera sugerir ambigüedad o confusión.

La muestra estuvo compuesta por trabajadores de la compañía DXC, empresa tecnológica del sector de las telecomunicaciones y por niños pertenecientes al entorno de la urbanización Puerta de Madrid de Majadahonda donde existe una población infantil muy numerosa y se cuenta con fácil acceso a los niños que colaboran con el estudio.

En todos los casos, junto a la versión preliminar de la escala CISSve, se entregó el consentimiento informado (anexo III). El documento se preparó en dos versiones diferentes, una para los adultos y otra para los padres de los menores, con objeto de informar de que la colaboración se realizaba de forma voluntaria y sin ánimo de lucro, explicar su propósito, y la conformidad con la utilización de los datos para este estudio, tal y como exige la Ley Orgánica de Protección de Datos (Ley Orgánica 15/99, 13 diciembre) <sup>167</sup>.

De la muestra inicial recogida, con 83 registros, se suprimieron 40 por no cumplir los criterios de edad, pero se guardaron los datos para poder ser utilizados en otro momento de esta investigación o de investigaciones futuras. Se trataba de sujetos que se descartaron de la muestra reclutada por ser menores de 9 o mayores de 30, que caían fuera del rango estudiado por sus autores en la versión originaria del CISSv15 <sup>59 60</sup>.

Utilizando la metodología propuesta inicialmente por Rasch <sup>103</sup> y extrapolada a su modelo poltómico por Andrich <sup>6,148</sup>, se hizo una primera exploración del funcionamiento del CISSve con objeto de obtener un análisis preliminar. El trabajo se realizó con el soporte de las herramientas Winsteps 4.01 <sup>107</sup>, Jmetrik <sup>139</sup> y SPSS v.25 <sup>137</sup> que proporcionan la capacidad de cómputo y análisis estadístico necesarios para realizar los cálculos requeridos por el modelo y comprender el comportamiento inicial del CISSve a través de la información proporcionada por la pequeña muestra utilizada en esta fase.

## RESULTADOS

Se evaluaron los resultados de acuerdo a los criterios de Khadka et al.<sup>97</sup>. En primer lugar y antes de comenzar el análisis, examinamos si había sujetos con más de un 33% de respuestas no válidas para excluirlos si se diera el caso. Todos los sujetos que respondieron al cuestionario contestaron todas las preguntas. Con objeto de poder recabar el “feedback” sobre la claridad de las preguntas y analizar si existía algún ítem con alguna descripción confusa, todos los cuestionarios fueron aplicados individualmente.

A nivel descriptivo, la puntuación en la versión prefinal de la escala CISSve, que se obtuvo para la muestra de sujetos elegida (n=43), fue media=-2,21; DT=1,46 logits. Es decir, se trata de una muestra dentro de una población general, no sintomática a priori en la mayoría de sus casos, con valores de dispersión bastante elevados.

A continuación se muestra el resumen inicial que proporciona el software Winsteps 4.01<sup>107</sup> en la Figura 5.13.

AN-LISIS PRELIMINAR DEL CISS									
Person	83 INPUT		43 MEASURED		INFIT		OUTFIT		
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	9.3	15.0	-2.21	.55	1.00	.1	.95	.0	
P.SD	8.2	.0	1.46	.14	.40	.9	.50	.9	
REAL RMSE	.57	TRUE SD	1.35	SEPARATION	2.35	Person	RELIABILITY	.85	
Item	15 INPUT		15 MEASURED		INFIT		OUTFIT		
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	26.7	43.0	.00	.32	1.01	.0	.95	-.1	
P.SD	12.2	.0	1.00	.06	.37	1.4	.38	1.3	
REAL RMSE	.32	TRUE SD	.95	SEPARATION	2.95	Item	RELIABILITY	.90	

FIGURA 5.13 – DESCRIPTIVOS DE LA VERSIÓN PRELIMINAR DE CISSVE

En ella, se observa que en una escala logit, en la que fijamos arbitrariamente el “0 de escala” en la media de dificultad de los ítems y donde se toma como unidad la desviación típica de la distribución de dificultad de los ítems de la escala, se obtuvo un desalineamiento medio de -2.21 logits respecto a la media de dificultad de los ítems expresada en logits, lo que indica que los sujetos quedan desplazados hacia la parte no sintomática. Esto es, la población analizada a priori muestra ser en más de un 95%, asintomática, ya que el alejamiento del “0 de la escala” es superior a dos desviaciones típicas.

Los valores de confiabilidad, calculados en base a los errores cuadráticos que comparan lo observado frente a lo esperado por el modelo, arrojan resultados altos, como se puede ver tanto en los valores de separación expresados en escala logit, como en términos porcentuales a través del coeficiente de fiabilidad, semejante al  $\alpha$  de Cronbach que nos muestra Winsteps. Podemos observar un 90% (Ítem Reliability) para los ítems y un 85% para las personas (Person Reliability).

Buscando un poco más de detalle, a través del software estadístico IBM SPSS, obtenemos los detalles sobre la distribución de las variables que representan edad y puntuación que se pueden observar en la Figura 5.14.

Edad			Puntuación		
N	Válido	43	N	Válido	43
	Perdidos	0		Perdidos	0
Media		15,60	Media		9,3256
Error estándar de la media		,686	Error estándar de la media		1,27055
Mediana		15,00	Mediana		7,0000
Desv. Desviación		4,499	Desv. Desviación		8,33158
Asimetría		1,356	Asimetría		1,944
Error estándar de asimetría		,361	Error estándar de asimetría		,361
Curtosis		1,571	Curtosis		3,367
Error estándar de curtosis		,709	Error estándar de curtosis		,709
Mínimo		9	Mínimo		1,00
Máximo		28	Máximo		37,00

FIGURA 5.14 – DESCRIPCIÓN DE LA MUESTRA QUE PARTICIPA EN PRE-TEST DE CISSVE Y SU PUNTUACIÓN

Por simple inspección de estos resultados, se puede ver, a través de la mediana que la mitad de los sujetos que participaron en esta fase eran menores de 15 años y también que la mitad de todos los que contestaron, arrojaron una puntuación muy baja (inferior a 7 puntos). Un coeficiente de asimetría en ambos casos ratifica este hecho, significando en el caso de la edad que la distribución de los sujetos está más concentrada en la parte más joven. Por su parte, las puntuaciones son bajas en la mayoría de los casos, como es lógico, teniendo en cuenta que se trata de una población general de edad joven y conocemos que la prevalencia de la IC es muy baja en esos rangos de edad. La curtosis, muestra un apuntalamiento positivo, reflejando de nuevo la misma idea de concentración en torno a edades jóvenes y puntuaciones bajas.

En cuanto a la desviación típica nos da más información. En el caso de la edad, constata el hecho de que la franja de edad se mueve en valores de  $15 \pm 4.5$  años, mientras que, en el caso de la puntuación con una dispersión proporcionalmente mucho más alta, refleja que los pocos casos sintomáticos, claramente se separan de la media.

Veámoslo a través de un enfoque más gráfico, el que nos proporciona la Figura 5.15, que muestra un diagrama de Wright obtenido con Winsteps que enfrenta horizontalmente los ítems frente a los sujetos y da una primera idea sobre el mapa de dificultad de los ítems frente a la habilidad de los sujetos que se analizan.

## RESULTADOS

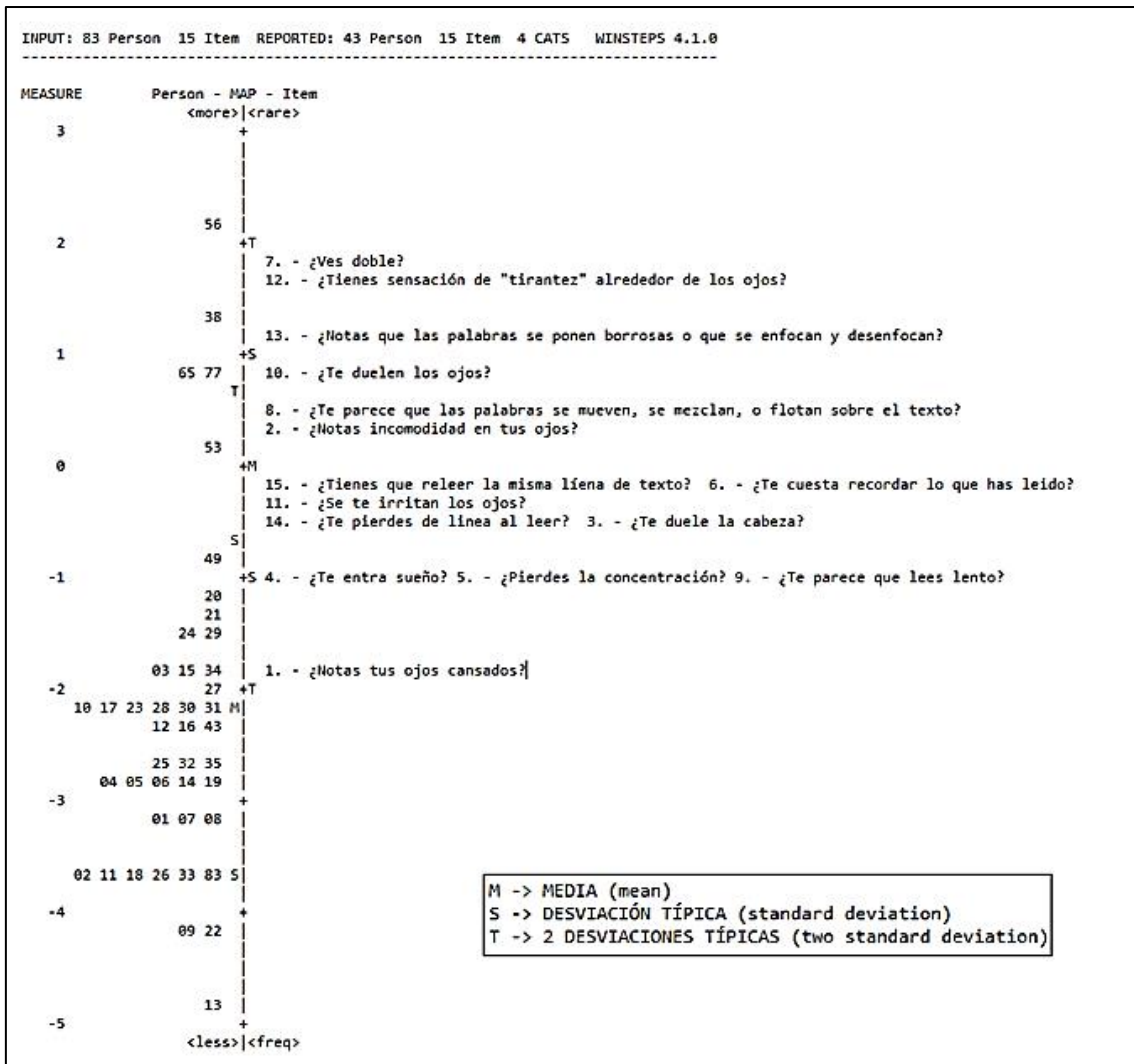


FIGURA 5.15 – MAPA DE WRIGHT DEL PRE-TEST DE LA ESCALA CISSVE

En la Figura 5.15, M (mean), indica la media. En el lado derecho de los ítems y en el lado izquierdo de los sujetos. Del mismo modo, S (standard deviation) se refiere a la desviación típica.

De nuevo, observamos gráficamente lo que ya habíamos visto numéricamente.

- 1) Se ha fijado el “0 de la escala” arbitrariamente en la media de dificultad de los ítems.
- 2) La unidad de la escala logit (1 logit) se corresponde con desviación típica de la distribución de dificultad de los ítems.
- 3) La sintomatología de los sujetos está desplazada hacia la parte negativa, indicando que se trata de una población predominantemente asintomática. Observamos que incluso el ítem más fácil (es decir, el ítem 1 en el que los sujetos menos sintomáticos marcan de media una categoría más alta en comparación con los otros ítems), es difícil para un sujeto de habilidad media, dentro de los que han completado este pre-test, ya que su nivel queda por encima de la media de los sujetos.
- 4) La dispersión existente entre las puntuaciones obtenidas por los sujetos es mayor que la existente entre la suma de puntuaciones de todos los sujetos para cada ítem, es

decir, el CISSve está optimizado para medir síntomas de IC en un segmento de la población, y la muestra que se está analizando cubre un rango mayor.

- 5) La diferencia en logits entre ítem y sujeto que observamos, indica la mayor o menor probabilidad de éxito para ese ítem, tal y como se muestra en la Figura 5.16.

Logit Difference	Probability of Success	Probit Difference		Logit Difference	Probability of Success	Probit Difference
5.0	99%	2.48		-5.0	1%	-2.48
4.6	99%	2.33		-4.6	1%	-2.33
4.0	98%	2.10		-4.0	2%	-2.10
3.0	95%	1.67		-3.0	5%	-1.67
2.2	90%	1.28		-2.2	10%	-1.28
2.0	88%	1.18		-2.0	12%	-1.18
1.4	80%	0.85		-1.4	20%	-0.85
1.1	75%	0.67		-1.1	25%	-0.67
1.0	73%	0.62		-1.0	27%	-0.62
0.8	70%	0.50		-0.8	30%	-0.50
0.69	67%	0.43		-0.69	33%	-0.43
0.5	62%	0.31		-0.5	38%	-0.31
0.4	60%	0.25		-0.4	40%	-0.25
0.2	55%	0.13		-0.2	45%	-0.13
0.1	52%	0.06		-0.1	48%	-0.06
0.0	50%	0.00		-0.0	50%	-0.00

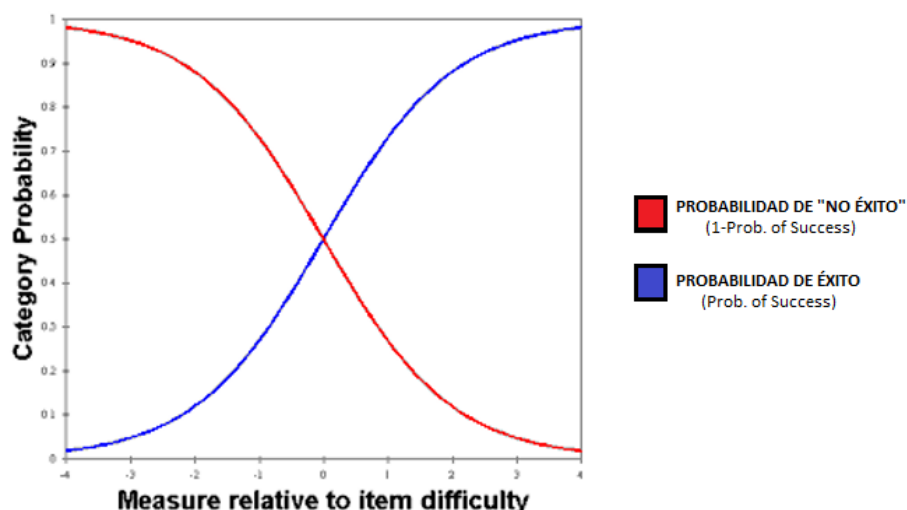


FIGURA 5.16 – CORRESPONDENCIA ENTRE DISTANCIA ÍTEM-SUJETO EN ESCALA LOGIT Y PROBABILIDAD DE ÉXITO PARA EL ÍTEM DE UN DETERMINADO SUJETO

A continuación, se revisó si había personas en nuestra muestra con un Outfit por encima de 3.0. No hubo ninguna. En el caso de que hubiese habido alguna, se habría eliminado, ya que significaría que habría contestado de forma inadecuada.

Se comprobó el ajuste al modelo de Rasch de las puntuaciones de los sujetos. En la Figura 5.17, puede observarse que aparecen 9 sujetos fuera del rango óptimo recomendado por Khadka et al.<sup>97,114</sup> (0.7-1.3) para el “infit” y el “outfit”, lo que indica un ligero desajuste con los valores que se podría esperar sobre ellos de acuerdo con el modelo de Rasch.

RESULTADOS

TABLE 6.1 AN-LISIS PRELIMINAR DEL CISS ZOU993WS.TXT Apr 26 2020 13:59  
 INPUT: 83 Person 15 Item REPORTED: 43 Person 15 Item 4 CATS WINSTEPS 4.1.0  
 Person: REAL SEP.: 2.35 REL.: .85 ... Item: REAL SEP.: 2.95 REL.: .90

Person STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT  MNSQ ZSTD	OUTFIT  MNSQ ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Person		
19	5	15	-2.90	.54	1.92	1.8	2.52	2.1	A-.11	.41	60.0	70.6	19
33	3	15	-3.59	.65	1.90	1.6	2.15	1.4	B .10	.34	80.0	81.1	33
38	32	15	1.32	.38	1.36	1.1	2.11	2.4	C .28	.58	66.7	51.8	38
7	4	15	-3.21	.58	1.84	1.6	1.70	1.1	D .12	.38	73.3	76.3	07
56	37	15	2.14	.44	1.83	1.8	1.49	1.1	E .44	.53	53.3	63.0	56
15	10	15	-1.77	.43	1.69	1.6	1.51	1.3	F .56	.50	53.3	60.6	15
17	8	15	-2.17	.46	1.38	1.0	1.61	1.4	G .08	.47	40.0	61.5	17
65	28	15	.77	.36	1.37	1.1	1.35	1.1	H .27	.59	40.0	48.6	65
49	16	15	-.81	.38	1.34	1.0	1.33	1.0	I .51	.56	40.0	52.1	49
4	5	15	-2.90	.54	1.92	1.8	2.52	2.1	J .44	.41	86.7	70.6	04
23	8	15	-2.17	.46	1.19	.6	1.31	.8	K .47	.47	53.3	61.5	23
43	7	15	-2.39	.48	1.29	.8	.91	.0	L .67	.46	73.3	65.0	43
14	5	15	-2.90	.54	1.19	.6	.97	.1	M .49	.41	73.3	70.6	14
20	14	15	-1.11	.39	1.12	.4	1.09	.4	N .52	.54	40.0	55.9	20
21	13	15	-1.26	.40	1.10	.4	1.08	.4	O .57	.54	60.0	56.4	21
30	8	15	-2.17	.46	1.10	.4	1.02	.2	P .40	.47	66.7	61.5	30
11	3	15	-3.59	.65	1.09	.4	.95	.2	Q .21	.34	73.3	81.1	11
77	28	15	.77	.36	1.06	.3	1.09	.4	R .45	.59	40.0	48.6	77
83	3	15	-3.59	.65	1.09	.4	1.07	.4	S .18	.34	73.3	81.1	83
3	10	15	-1.77	.43	1.04	.2	1.02	.2	T .25	.50	40.0	60.6	03
6	5	15	-2.90	.54	1.02	.2	.74	-.3	U .59	.41	80.0	70.6	06
8	4	15	-3.21	.58	.98	.1	.77	-.2	V .34	.38	66.7	76.3	08
53	23	15	.13	.36	.86	-.3	.91	-.2	u .53	.59	53.3	47.8	53
12	7	15	-2.39	.48	.85	-.3	.75	-.4	t .57	.46	53.3	65.0	12

FIGURA 5.17 – AJUSTE AL MODELO DE RASCH DE LAS PUNTUACIONES DE LOS SUJETOS

Recordemos que tanto el Infit como el Outfit se calculan como suma y ponderación de los errores cuadráticos. El Infit es más sensible a los errores en los ítems que están en el mismo nivel que los sujetos medidos y el Outfit es más sensible a errores en los ítems de nivel muy dispar al del sujeto medido, es decir, o muy fáciles, o muy difíciles para ese sujeto.

Sin entrar a nivel de ítem, hacemos un primer resumen de las características generales de este cuestionario, CISSve basándonos en la guía de Khadka et al <sup>97,114</sup>.

Según se aprecia en la Figura 5.18, este primer chequeo nos indica que el Infit y el Outfit tienen un ajuste ideal al modelo, nivel A en la guía de Khadka, ya que los valores medios están muy cercanos a 1. También se observa que el cuestionario está bastante sesgado hacia la parte menos sintomática de la muestra. La diferencia sujetos-ítems es de -2.21, que corresponde con un nivel B en la guía de Khadka. El índice de separación es de 2.35, que está por encima del mínimo y también se corresponde con un nivel B en la guía de Khadka.

SUMMARY OF 43 MEASURED Person

	TOTAL SCORE		MEASURE	MODEL S.E.	INFIT		OUTFIT	
	SCORE	COUNT			MNSQ	ZSTD	MNSQ	ZSTD
MEAN	9.3	15.0	-2.21	.52	1.00	.1	.95	.0
P.SD	8.2	.0	1.46	.13	.40	.9	.50	.9
S.SD	8.3	.0	1.48	.13	.41	1.0	.51	.9
MAX.	37.0	15.0	2.14	1.04	1.92	1.8	2.52	2.4
MIN.	1.0	15.0	-4.88	.36	.42	-1.9	.25	-1.8
REAL RMSE	.57	TRUE SD	1.35	SEPARATION	2.35	Person	RELIABILITY	.85
MODEL RMSE	.53	TRUE SD	1.36	SEPARATION	2.55	Person	RELIABILITY	.87
S.E. OF Person MEAN = .23								

DELETED: 40 Person  
 Person RAW SCORE-TO-MEASURE CORRELATION = .97  
 KRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .93 SEM = 2.20

SUMMARY OF 15 MEASURED Item

	TOTAL SCORE		MEASURE	MODEL S.E.	INFIT		OUTFIT	
	SCORE	COUNT			MNSQ	ZSTD	MNSQ	ZSTD
MEAN	26.7	43.0	.00	.30	1.01	.0	.95	-.1
P.SD	12.2	.0	1.00	.06	.37	1.4	.38	1.3
S.SD	12.7	.0	1.03	.06	.38	1.5	.39	1.4
MAX.	54.0	43.0	1.77	.42	2.05	3.3	1.66	1.8
MIN.	9.0	43.0	-1.83	.23	.56	-2.3	.36	-2.2
REAL RMSE	.32	TRUE SD	.95	SEPARATION	2.95	Item	RELIABILITY	.90
MODEL RMSE	.30	TRUE SD	.95	SEPARATION	3.15	Item	RELIABILITY	.91
S.E. OF Item MEAN = .27								

FIGURA 5.18 – RESUMEN GENERAL DE AJUSTE DE ÍTEM Y SUJETOS EN LA VERSIÓN PRELIMINAR DE CISSVE

Como podemos ver numéricamente, en la Figura 5.19 y gráficamente en la Figura 5.20, en cuanto a las categorías de respuesta, la guía de Khadka menciona que las categorías de respuesta deben estar ordenadas y que la distancia entre los umbrales (Andrich thresholds) debe estar entre 1.4 y 5.0. Se cumplen ambos criterios recomendados, observándose que las categorías de respuesta aparecen ordenadas y separadas aproximadamente 1.5 logits, con lo que concluimos que la escala discrimina bien, asignándole un nivel A en la guía de Khadka.

TABLE 3.2 AN-LISIS PRELIMINAR DEL CISS ZOU993WS.TXT Apr 26 2020 13:59  
 INPUT: 83 Person 15 Item REPORTED: 43 Person 15 Item 4 CATS WINSTEPS 4.1.0

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY	OBSERVED	OBSVD	SAMPLE	INFINIT	OUTFIT	ANDRICH	CATEGORY		
LABEL	SCORE	COUNT	%AVRGE	EXPECT	MNSQ	MNSQ	THRESHOLD	MEASURE	
0	0	360	56	-3.17	-3.14	1.06	1.04	NONE   (-3.00)	0
1	1	198	31	-1.63	-1.71	.92	.82	-1.82   -.85	1
2	2	58	9	-.12	-.06	1.12	1.15	.30   .97	2
3	3	29	4	1.56	1.64	1.00	.92	1.52   ( 2.80)	3

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY	STRUCTURE	SCORE-TO-MEASURE	50% CUM.	COHERENCE	ESTIM			
LABEL	MEASURE	S.E.	AT CAT. ----ZONE----	PROBABLY	M->C C->M RMSR	DISCR		
0	NONE		( -3.00) -INF -2.08	81% 81%	.4203	0		
1	-1.82	.10	-.85 -2.08 .11	-1.93	57% 62%	.5194	1.08	1
2	.30	.16	.97 .11 1.98	.19	44% 48%	.9201	.93	2
3	1.52	.26	( 2.80) 1.98 +INF	1.74	67% 34%	.9115	.90	3

M->C = Does Measure imply Category?  
 C->M = Does Category imply Measure?

FIGURA 5.19 – DISTRIBUCIÓN DE LAS CATEGORÍAS DE RESPUESTA DE LA ESCALA CISSVE

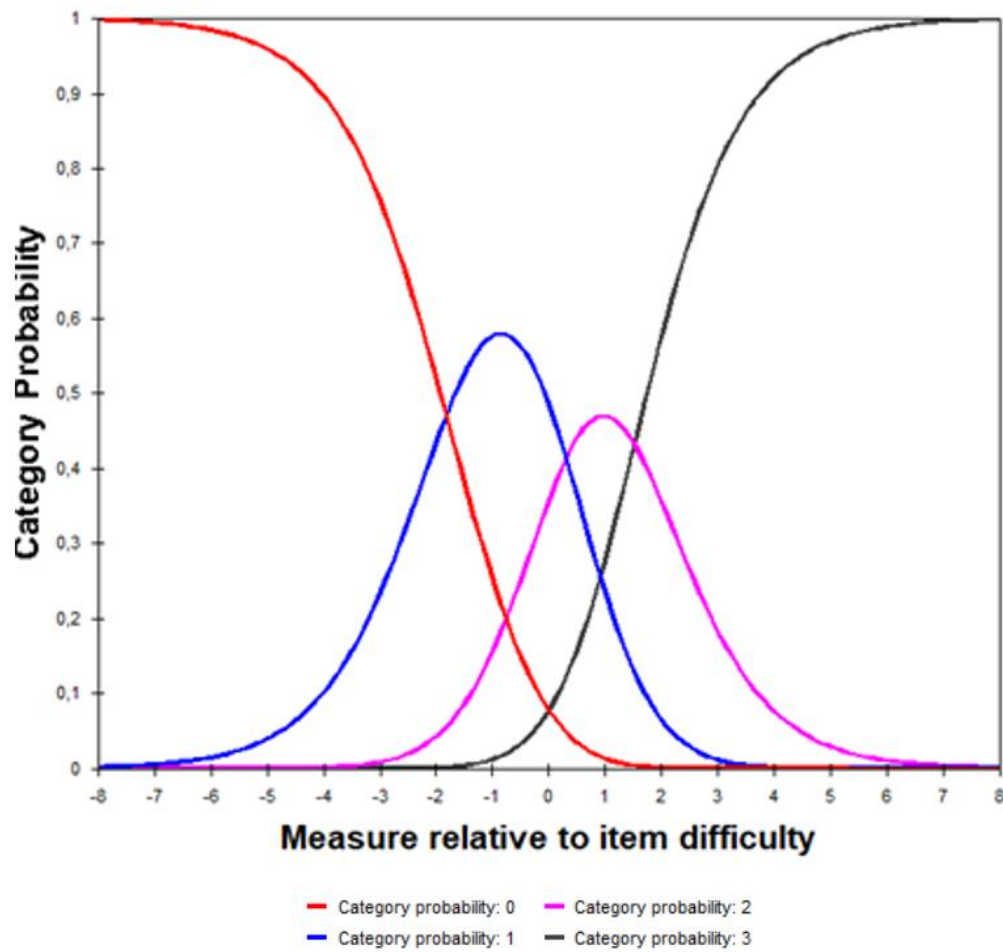


FIGURA 5.20 – DISTRIBUCIÓN GRÁFICA DE LAS CATEGORÍAS DE RESPUESTA DE LA ESCALA CISSVE

Destacar la peculiaridad de que la muestra analizada es prácticamente asintomática, y eso hace que aparezcan sólo 4 categorías, pues ningún sujeto en ningún ítem marcó la opción más sintomática.

Se revisó el ajuste de cada ítem al modelo de Rasch. Como puede observarse en la Figura 5.21, prácticamente todos los ítems están en el intervalo recomendado (0.7-1.3) o lo rozan encontrándose muy cerca, pero el ítem 2, se sale ampliamente, arrojando un Infit de 2.05 que se corresponde con un nivel B en la guía de Khadka.

Recordemos que el Infit es más sensible a respuestas de sujetos con sintomatologías parejas a la dificultad del ítem y el Outfit es más sensible a las respuestas de los sujetos con sintomatologías más distantes a la dificultad del ítem (muy sintomáticos en comparación con la dificultad de ese ítem, o muy poco sintomáticos para la dificultad de ese ítem). Valores inferiores a 0.7 significan que el ítem es demasiado predecible mientras que los que superan 1.3, significan que el ajuste no es óptimo cuando se compara lo observado con lo esperado por el modelo.

TABLE 10.1 ANALISIS PRELIMINAR DEL CISS ZOU498MS.TXT Sep 14 2021 18: 9  
 INPUT: 83 Person 15 Item REPORTED: 43 Person 15 Item 4 CATS WINSTEPS 4.1.0  
 Person: REAL SEP.: 2.35 REL.: .85 ... Item: REAL SEP.: 2.95 REL.: .90

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT ZSTD	OUTFIT ZSTD	PTMEASUR-CORR.	AL-EXP.	AL-OBS%	EXACT MATCH EXP%	Item
2	20	43	.41	.30 2.05	3.3 1.66	1.8 A.32	.69	60.5	70.7	2.	-Notas incomodidad en tus ojos?
14	31	43	-.47	.26 1.35	1.5 1.46	1.8 B.54	.72	53.5	60.8	14.	-Te pierdes de línea al leer?
1	54	43	-1.83	.23 1.12	.6 1.41	1.7 C.59	.70	51.2	55.4	1.	-Notas tus ojos cansados?
9	39	43	-.99	.25 1.32	1.4 1.33	1.4 D.63	.72	60.5	57.3	9.	-Te parece que lees lento?
11	28	43	-.25	.27 1.20	.9 1.29	1.2 E.69	.71	62.8	62.6	11.	-Se te irritan los ojos?
15	26	43	-.10	.28 1.17	.8 1.05	.3 F.65	.71	67.4	63.2	15.	-Tienes que releer la misma línea de texto?
10	16	43	.81	.33 1.00	.1	-.7 G.79	.67	79.1	75.8	10.	-Te duelen los ojos?
8	19	43	.50	.31 .93	-.2	-.1 H.82	.69	74.4	71.7	8.	-Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
7	9	43	1.77	.42 .84	-.3	-.3 I.75	.62	88.4	86.2	7.	-Ves doble?
13	13	43	1.17	.36 .72	-.9	.83	.65	83.7	80.3	13.	-Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
6	26	43	-1.10	.28 .81	-.8	.72	.71	72.1	63.2	6.	-Te cuesta recordar lo que has leído?
5	40	43	-1.05	.25 .80	-.9	.78	.72	55.8	57.2	5.	-Pierdes la concentración?
12	10	43	1.60	.40 .62	-1.1	.80	.63	90.7	84.9	12.	-Tienes sensación de "tirantez" alrededor de los ojos?
3	32	43	-.54	.26 .56	-2.3	.63	.72	74.4	60.5	3.	-Te duele la cabeza?
4	38	43	-.93	.25 .61	-2.0	.58	.72	69.8	57.4	4.	-Te entra sueño?
MEAN	26.7	43.0	.00	.30 1.01	.0	.95	-.1	69.6	67.2		
P.SD	12.2	.0	1.00	.06 .37	1.4	.38	1.3	12.0	10.1		

FIGURA 5.21 – AJUSTE DE CADA ÍTEM AL MODELO DE RASCH

## RESULTADOS

Corroboramos este hallazgo con el programa JMETRIK repasando la discriminación de los ítems y encontrando que efectivamente, el ítem 2 para esta muestra tiene una discriminación insuficiente, de acuerdo con los criterios definidos por Ebel <sup>168</sup>, al presentar un coeficiente de 0,18, por debajo del umbral como se puede ver en la Figura 5.22. Al realizarse este análisis preliminar en una pequeña muestra, no se pudo sacar una conclusión definitiva respecto a este ítem, pero se tuvo en cuenta para observarlo en un análisis posterior.

Escala de discriminación de Ebel, Elosua

MUY BUENA	>0,4
BUENA	0,29<a<0,39
DEFICIENTE	0,20<a<0,29
INSUFICIENTE	<0,19

	SCORE	DIFIC	SD	DISCRIM
item2	Overall	0,4651	0,6305	0,1881
	0.0 (0.0)	0,6047	0,4947	-0,2325
	1.0 (1.0)	0,3256	0,4741	-0,0361
	2.0 (2.0)	0,0698	0,2578	0,2710
	3.0 (3.0)	0,0000	0,0000	NaN
	4.0 (4.0)	0,0000	0,0000	NaN

FIGURA 5.22 – ÍNDICE DE DISCRIMINACIÓN DEL ÍTEM 2 EN LA ESCALA PRELIMINAR CISSVE

## 5.2 VALIDACIÓN DE LA ESCALA CISSve

### 5.2.1 CONSIDERACIONES GENERALES

Una vez disponible la escala adaptada a español CISSve, el siguiente paso fue la validación de este nuevo instrumento. Este proceso consta de dos fases, una cualitativa y otra cuantitativa, como se puede ver en la sección de métodos en el esquema de la Figura 4.1.

En la primera fase se evalúa el contenido de los ítems para valorar su idoneidad en cuanto al propósito que se pretende, es decir, medir síntomas de IC y en la segunda se realiza un análisis cuantitativo en el que se analizan mediante el método de Rasch <sup>103,147</sup> las propiedades de la escala y la calidad de la misma, evaluando la validez del constructo que la sustenta y la fiabilidad y repetibilidad que presenta.

### 5.2.2 FASE CUALITATIVA. VALIDACIÓN DE CONTENIDO

La búsqueda en la base de datos de instrumentos PRO, Proquolid, no arrojó más resultados a parte de los bien conocidos CIRS y CISS en sus distintas versiones, pero sí un análisis del

mismo, el Convergence insufficiency Treatment Trial, que analizaba los síntomas con las preguntas del CISS <sup>10</sup>.

También se encontró un análisis sobre la validez del CISSv15 realizada por Catherine Arenas-Mejía <sup>169</sup> en el que se reportaba una sensibilidad de un 83,3% y una especificidad de un 81,6% para un grupo de 50 personas de entre 15 y 35 años (n=50, M=24,42, DT=5,57 años). En este trabajo, los participantes en primer lugar fueron evaluados clínicamente, siendo sometidos a pruebas de motilidad ocular, punto próximo de convergencia y pruebas de fusión, obteniéndose tras el examen clínico un diagnóstico positivo de IC en 12 de los 50 sujetos examinados. En cuanto al instrumento, el criterio que se estableció para considerar a los sujetos como sintomáticos fue el hecho de superar un punto de corte fijado en 20 puntos en la escala CISS-v15 sin hacer distinción de la edad. Resultaron positivos 10 de los 12 pacientes sintomáticos y 7 de los 38 asintomáticos, concluyendo por lo tanto en los valores de sensibilidad (83.3%) y especificidad (81.6%) referidos anteriormente.

En la revisión bibliografía para verificar la idoneidad de los ítems del se encontró que el CISSve cubría la mayoría del espectro de síntomas de IC registrados en la literatura científica <sup>13,30,41,42,44,50-55,57</sup>.

Las entrevistas individuales llevadas a cabo con los pacientes que completaron el pre-test corroboraron que todos los ítems eran perfectamente claros e inteligibles y que no existían problemas para poder completar adecuadamente el cuestionario. Adicionalmente, tanto el grupo de discusión en la empresa DXC como los padres de los participantes menores de edad, corroboraron la inteligibilidad y claridad del contenido de la escala, tanto en su parte explicativa como en las opciones de respuesta.

Algunos de los profesionales Ópticos-Optometristas consultados, refirieron algunos comentarios de valor, como la posible redundancia que podían inferir algunos ítems y la posibilidad de prescindir de otros en anamnesis realizadas en la práctica profesional para hacer prediagnósticos rápidos. Concretamente, el ítem 2, el ítem 4, el ítem 8 y el ítem 12 fueron cuestionados como apropiados para el diagnóstico al redundar con otros e introducir, bajo su punto de vista, complejidad innecesaria en las preguntas realizadas en las anamnesis.

Bajo la opinión del equipo de investigación, cada ítem de entre los planteados por el CISSve (adaptado del CISSv15) introduce ligeros matices respecto a los demás y pese a que las observaciones sobre la posible redundancia o falta de claridad apuntadas por algunos profesionales se valoró considerablemente, se estimó que el análisis de Rasch detectaría durante las fases siguientes de la validación esas posibles duplicidades, localizándolas en la escala logit. También se estimó que los ítems considerados por algunos profesionales como prescindibles por tener poco valor diagnóstico, serían detectados gracias a los valores de Infit y Outfit del modelo de Rasch que conseguiría cuantificar esta apreciación en caso de ser cierta.

## RESULTADOS

### 5.2.3 FASE CUANTITATIVA. VALIDACIÓN DEL CONSTRUCTO CON ANALISIS DE RASCH, VALIDEZ CONVERGENTE Y DIVERGENTE Y REPETIBILIDAD

#### 5.2.3.1 ESTRATEGIA DEL ANÁLISIS

En este apartado se presentan los resultados obtenidos del análisis de las propiedades psicométricas de la escala CISSve mediante el análisis de Rasch realizado con ayuda del paquete de software Winsteps 4.0.1 (Winsteps.com, Beaverton, OR, USA) <sup>107</sup>. La evaluación de los resultados del modelo se resumen de acuerdo a los criterios propuestos por Khadka et al. <sup>97,114</sup>, que se pueden encontrar en 11.1 (anexo I).

#### 5.2.3.2 DESCRIPCIÓN DE LA MUESTRA

El análisis realizado en nuestro trabajo estuvo referido a una población general comprendida en el rango de edad de 9-30 años.

El conjunto de la muestra (n=458, M=16.02, DT=5.68 años) utilizado para el estudio de la escala CISSve fue recogido de diversos orígenes, una empresa tecnológica (n=15, M=25.07, DT=4.44 años), un colegio (n=162, M=9.92, DT=0.83 años), un instituto (n=95, M=15.56, DT=1.69 años) y una universidad (n=158, M=21.47, DT=2.02 años). Las muestras tomadas a población inglesa con objeto de comparar las propiedades de la versión adaptada del instrumento, CISSve, con las de la original, CISSv15, fueron obtenidas en un colegio público en Reading, Reino Unido (n=182, M=15,82, D=1,62 años).

En primer lugar y antes de comenzar el análisis, se descartó a los sujetos con más de 33% de respuestas no válidas y a los que de forma evidente hubieran contestado de forma arbitraria a los ítems de la escala. De entre los 458 sujetos que participaron en la prueba rellenando el cuestionario CISSve, la inmensa mayoría, 439 cumplieron los criterios de inclusión, al haber contestado al mínimo número de preguntas requerido (al menos 10) y no haberlo hecho de forma arbitraria. Para identificar a los sujetos que respondieron al azar, nos ayudamos en primer lugar de las tablas "Person Measures", "Person fit order" y "Person keyform unexpected" de Winsteps <sup>107</sup>, revisando las respuestas de los participantes en las que los valores de infit y outfit eran significativamente mayores que 1 y en las que por lo tanto se infería un mal ajuste al modelo. Posteriormente, a través de la tabla "Person keyform unexpected" se veía el detalle persona a persona de las diferencias entre las respuestas esperadas según el modelo frente a lo que se observaba en las respuestas. De ese conjunto de personas se descartaron las personas con respuestas manifiestamente incoherentes. Por ejemplo, patrones 0-4-0-4-0-4..., 3-3-3-3-3-3, etc.

En la Figura 5.23 se puede observar la muestra final seleccionada (439 sujetos de los 458 que contestaron el cuestionario).

Person	458 INPUT	439 MEASURED			INFINIT	OUTFIT		
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	15.1	14.9	-1.44	.42	1.00	.0	.99	.01
P.SD	10.0	.3	1.18	.25	.45	1.3	.47	1.2
REAL RMSE	.48	TRUE SD	1.07	SEPARATION	2.22	Person RELIABILITY		.83
Item	23 INPUT	15 MEASURED			INFINIT	OUTFIT		
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	443.2	437.3	.00	.06	1.02	.1	.99	-.21
P.SD	130.7	1.7	.50	.01	.18	2.9	.18	2.51
REAL RMSE	.06	TRUE SD	.50	SEPARATION	7.69	Item RELIABILITY		.98

FIGURA 5.23 – DESCRIPCIÓN (WINSTEPS) DE LA MUESTRA QUE CONTESTÓ LA ESCALA CISSVE

A nivel descriptivo, las puntuaciones recogidas sobre la escala CISSve de nuestra muestra de sujetos fue (n=439; M= 15,1; DT=10,0), que en escala logit se corresponde con (n=439; M=-1.44; DT=1.18).

Los valores de confiabilidad ("reliability") proporcionados por Winsteps <sup>107</sup> que aparecen en azul en la Figura 5.23, se calculan de la siguiente manera:  $\text{Person Reliability} = (\text{Varianza real de las personas}) / (\text{Varianza real de las personas} + \text{Error medio de medición})$ . La "Varianza real de las personas" se refiere a la varianza en las medidas de habilidad o rasgo latente de los examinados. Esta es la variabilidad en las puntuaciones que se debe a las diferencias reales entre las habilidades o rasgos latentes de los examinados. El "Error medio de medición" es una medida del error asociado con las mediciones de las habilidades o rasgos latentes de los examinados. Este error puede ser debido a una variedad de factores, como por ejemplo a variabilidad en el rendimiento del examinado de un día a otro, errores en la administración del test o limitaciones en la precisión del test en sí mismo. Por lo tanto, el índice "Person Reliability" indica cuánta de la variabilidad total en las puntuaciones del cuestionario es debida a las diferencias reales entre las habilidades de los examinados. Un valor más alto de "Person Reliability" indica que una mayor proporción de la variabilidad de las puntuaciones se debe a las diferencias reales entre las habilidades (nivel de síntomas) de los sujetos examinado. Un valor de 83% indica que el test es preciso y confiable. Del mismo modo, la "Item Reliability" en el análisis de Rasch y en con software Winsteps se calcula como:  $\text{Item Reliability} = (\text{Varianza real de los ítems}) / (\text{Varianza real de los ítems} + \text{error medio de medición de los ítems})$ . La "Varianza real de los ítems" se refiere a la variabilidad en las medidas de dificultad de los ítems. Esta es la variabilidad en las puntuaciones se debe a las diferencias reales entre las dificultades de los ítems. El "error medio de medición de los ítems" es una medida del error asociado a las mediciones de dificultad de los ítems. Este error se puede deber a una variedad de factores, que van desde la variabilidad en la interpretación del ítem por parte de diferentes sujetos, errores en la calificación de las respuestas, o limitaciones en la precisión del test. Por lo tanto, la "Item Reliability" indica cuánta de la variabilidad total en las puntuaciones de los ítems se debe a las diferencias reales entre las dificultades de los ítems. Cuanto más alto sea, una mayor proporción de la variabilidad será debida a las diferencias reales de dificultad entre los ítems, lo que indicará que el test es más preciso y confiable en cuanto a la medición de las dificultades de los ítems. Un valor de un 98% indica una enorme confiabilidad.

## RESULTADOS

Obteniendo a través del software estadístico SPSS los detalles de la distribución de las variables para el total de la muestra (n=458), podemos observar en la Figura 5.24, que son muy parecidos a los que arroja Winsteps, reflejados en la Figura 5.23. Un valor de asimetría positivo en la edad nos indica que la muestra está algo más concentrada en la parte más joven de la muestra y una curtosis negativa nos dice que la distribución está bastante aplanada en su rango.

Respecto a la puntuación, el valor de la mediana nos informa de que más de la mitad de los sujetos que contestaron, obtuvieron una puntuación muy baja, por debajo de 13,5. El coeficiente de asimetría, siendo positivo, ratifica el hecho de que la muestra tiene más concentración en esa zona baja de puntuaciones, como es lógico al tratarse de una población general y conociendo el hecho de la baja prevalencia de la IC en poblaciones generales en esos rangos de edad para una muestra de carácter general <sup>38,41,47,52,56-61</sup>.

		EDAD	PUNTUACION
N	Válido	458	458
	Perdidos	2	2
Media		16,02	15,08
Error estándar de la media		,266	,471
Mediana		16,00	13,50
Moda		9	3
Desv. Desviación		5,689	10,071
Varianza		32,370	101,419
Asimetría		,317	,832
Error estándar de asimetría		,114	,114
Curtosis		-1,052	,447
Error estándar de curtosis		,228	,228
Rango		22	50
Mínimo		8	0
Máximo		30	50
Suma		7335	6905

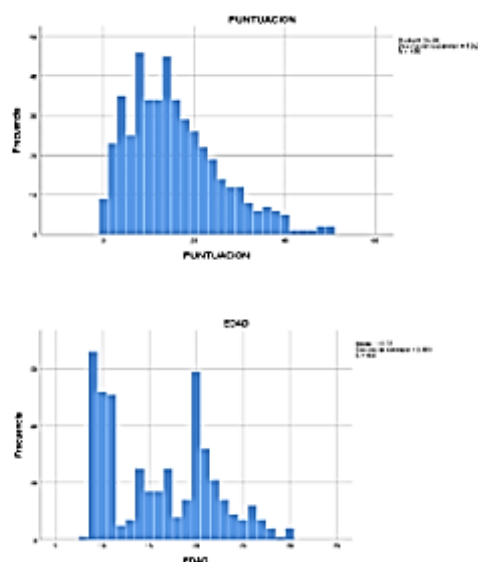


FIGURA 5.24 – DESCRIPCIÓN (SPSS) DE LA MUESTRA QUE CONTESTÓ LA ESCALA CISSVE

En 11.4 (anexo IV), se pueden encontrar varios gráficos y cuadros adicionales que analizan minuciosamente los estadísticos descriptivos de cada ítem, permitiendo observar el detalle de cada una de las variables que componen la escala.

### 5.2.3.3 ANALISIS DE RASCH

En este apartado, describimos los resultados obtenidos mediante el análisis de Rasch siguiendo la misma estructura secuencial que describimos en la sección de métodos:

5.2.3.3.1 CATEGORÍAS DE RESPUESTA DE LOS ÍTEMS DE LA ESCALA CISSVe

Examinando el orden que siguen los umbrales de las categorías de respuesta de la escala de calificación y la separación entre los mismos, podremos evaluar su rendimiento de acuerdo al criterio definido por Khadka et al.<sup>97,114</sup>

Empezaremos por revisar los resultados numéricos del análisis en la Figura 5.25, para pasar a la expresión gráfica de los mismos en la Figura 5.26.

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	OBSERVED SCORE	OBSVD COUNT	SAMPLE %	AVRGE	EXPECT	INFINIT MNSQ	OUTFIT MNSQ	ANDRICH THRESHOLD	CATEGORY MEASURE
0	0	2770	42	-2.12	-2.11	1.01	1.01	NONE	( -2.68)
1	1	1836	28	-1.29	-1.29	.96	.92	-1.32	-1.14
2	2	1211	18	-.62	-.67	.91	.87	-.55	-.05
3	3	578	9	-.10	-.08	1.07	1.12	.37	1.11
4	4	164	3	.42	.57	1.17	1.16	1.50	( 2.80)
MISSING		26	0	-1.59					

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

FIGURA 5.25 – RESUMEN DE LA ESTRUCTURA DE CATEGORÍAS DE RESPUESTA DE LA ESCALA CISSVE

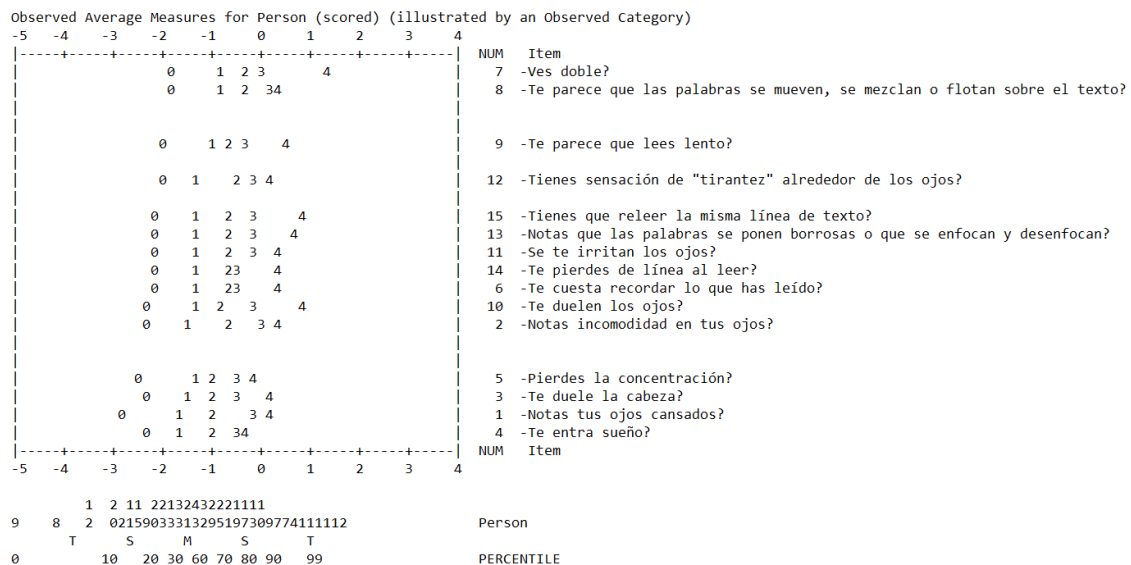


FIGURA 5.26 – CATEGORÍAS DE RESPUESTA POR ÍTEM DE LA ESCALA CISSVE

Tomando los valores medios de respuesta de los sujetos de la muestra, observamos que en todos los ítems, las categorías de respuesta guardan el orden, 0-1-2-3-4, confirmando que la estructura de la escala de medida es correcta y se ajusta perfectamente al modelo.

Como se puede observar en el resumen de la Figura 5.25, la mayoría de la muestra sitúa las respuestas en promedio, dentro de las categorías 0 y 1, confirmando otra vez el hecho de que la muestra elegida dentro de una población general de entre 9 y 30 años en promedio no presenta síntomas. En la misma tabla, también se aprecia que las categorías están ordenadas y

## RESULTADOS

que todas guardan distancias de más de un logit con sus aledañas, aunque entre la categoría 1 y la 2, y entre la categoría 2 y la 3 esta distancia es 1.1, ligeramente inferior a la recomendación que Khadka hace en su guía (1.4)<sup>97,114</sup>.

Gráficamente se ha podido observar este mismo hecho a través de las curvas de probabilidad de cada ítem. En la Figura 5.27 se muestra un ejemplo para el ítem 1, donde dependiendo de lo sintomáticos o asintomáticos que sean los sujetos, la probabilidad de estar en una categoría u otra, variará. En las gráficas se presentan ordenadas de menos a más prevalencia de los síntomas y describen, de izquierda a derecha, la frecuencia con la que los sujetos perciben los síntomas definidos en cada ítem, esto es, la curva de la izquierda representa la probabilidad de que el sujeto responda que “nunca” presenta un síntoma cuando se le pregunta por él, y la de la derecha representa cuándo contesta “siempre” cuando se le pregunta por ese síntoma.

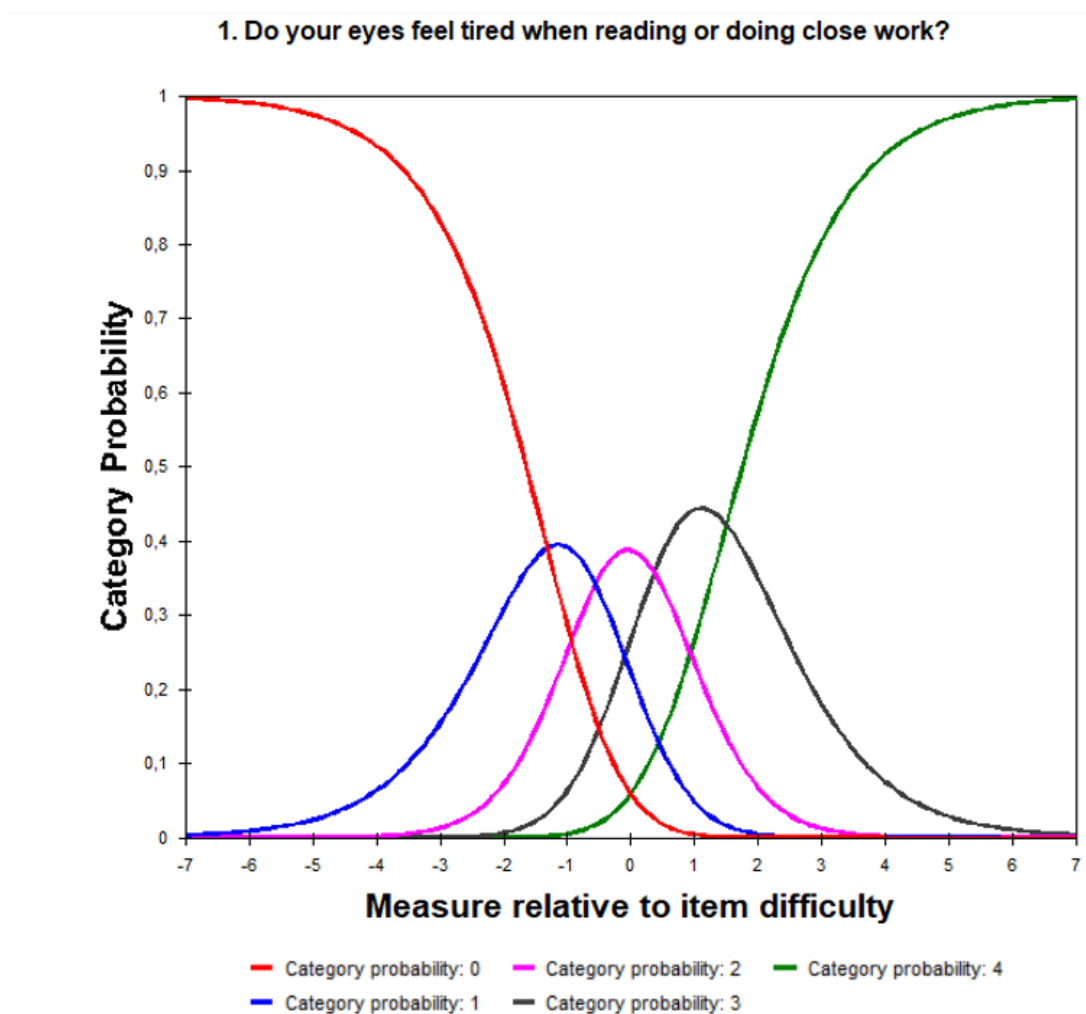


FIGURA 5.27 – CURVA DE PROBABILIDAD DEL ÍTEM 1 DE LA ESCALA CISSVE

### 5.2.3.3.2 ESTADÍSTICOS DE AJUSTE AL MODELO DE LOS ÍTEMS Y DE LAS PERSONAS

En nuestro análisis relativo a los ítems, podemos observar que prácticamente todos los ítems se ajustan al modelo, encontrándose que tanto Infit como Outfit presentan valores cercanos a

1 y se sitúan dentro del intervalo (0,7-1,3). Únicamente el ítem 1 tiene un valor de Infit de 0,58, y un valor de Outfit de 0,61, ligeramente por debajo de lo recomendado por Khadka et al.<sup>97,114</sup>, lo que indica simplemente que este ítem se podría considerar como demasiado predecible, por lo que aporta menos información de la que cabría esperar, tanto en el caso de pacientes poco sintomáticos, como se pueden comprobar a través del valor del Infit, como también en el caso de pacientes con mayor sintomatología (pacientes más alejados de la dificultad de ese ítem, esto es, -0,65 logits, en escala “logit”).

Veamos el resumen que nos hace el programa Winsteps en la Figura 5.28, colocando los ítems de la escala CISSve según su adaptación al modelo (de menos adaptados a más adaptados, es decir, de ítems con residuos más grandes a ítems con residuos más pequeños).

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT   MNSQ   ZSTD	OUTFIT   MNSQ   ZSTD	PTMEASUR-AL   CORR.   EXP.	EXACT MATCH   OBS%   EXP%	Item
4	650	435	-.71	.05	1.26   3.8   1.22	3.0   A   .60	.65   39.0   42.5	-Te entra sueño?	
3	593	439	-.53	.06	1.17   2.6   1.23	3.1   B   .57	.64   39.1   43.6	-Te duele la cabeza?	
12	359	437	-.27	.06	1.19   2.5   1.13	1.4   C   .55	.57   46.7   51.0	-Tienes sensación de "tirantez" alrededor de los ojos?	
6	456	439	-.09	.06	1.15   2.1   1.14	1.7   D   .57	.60   45.6   46.4	-Te cuesta recordar lo que has leído?	
7	202	437	1.04	.08	1.15   1.6   1.04	4   E   .48	.48   65.0   65.1	-Ves doble?	
9	311	439	.48	.07	1.15   1.9   1.13	1.4   F   .52	.55   51.9   54.2	-Te parece que lees lento?	
5	577	437	-.49	.06	1.09   1.4   1.08	1.2   G   .61	.64   42.5   43.6	-Pierdes la concentración?	
8	219	438	.93	.08	1.09   1.0   .89	-1.0   H   .55	.50   66.9   63.4	-Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?	
14	447	436	-.07	.06	1.06   .9   1.07	.9   G   .56	.60   48.9   46.5	-Te pierdes de línea al leer?	
11	428	438	.01	.06	1.01   .2   .95	-.6   F   .61	.60   50.1   47.9	-Se te irritan los ojos?	
15	406	437	.09	.06	.95   -.7   .93	-.8   E   .60	.59   50.5   48.4	-Tienes que releer la misma línea de texto?	
13	417	433	.03	.06	.92   -1.1   .93	-.8   D   .64	.59   48.8   48.2	-Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?	
10	468	436	-.14	.06	.80   -3.1   .78	-3.0   C   .66	.61   47.8   46.1	-Te duelen los ojos?	
2	484	439	-.18	.06	.77   -3.7   .74	-3.6   B   .67	.61   51.6   44.8	-Notas incomodidad en tus ojos?	
1	631	439	-0.65	.05	.58   -7.6   .61	-6.5   A   .72	.65   56.5   42.6	-Notas tus ojos cansados?	
MEAN	443.2	437.3	.00	.06	1.02   .1   .99	-.2	50.1   48.9		
P.SD	130.7	1.7	.50	.01	.18   2.9   .18	2.5	7.7   6.7		

FIGURA 5.28 – ESTADÍSTICAS DE AJUSTE AL MODELO DE LOS ÍTEMES

### 5.2.3.3.3 DIMENSIONALIDAD

Cuando la varianza explicada por las medidas de Rasch es  $\leq 50\%$  y/o el autovalor (eigenvalue) en el primer contraste es  $\geq 2.0$ , tenemos una indicación de subconjuntos de elementos que sugieren multidimensionalidad<sup>97,114,115</sup>.

En nuestro análisis de las componentes principales (PCA) con Winsteps v4.0.1, se observa que un 45.4% de la varianza queda explicada con las medidas del CISSve.

También se puede apreciar que el autovalor del primer contraste es 2.2, lo que señala que el CISSve está midiendo una segunda característica, al superar el valor de 2, criterio que determina la significancia de esa dimensión y hace sospechar la multidimensionalidad. El resto de contrastes arrojan valores por debajo de 2 como se puede ver en la Figura 5.29.

## RESULTADOS

TABLE 23.0 AN-LISIS DEL CISS ZOU876WS.TXT May 13 2020 15:10  
 INPUT: 458 Person 23 Item REPORTED: 439 Person 15 Item 5 CATS WINSTEPS 4.1.0

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

	Eigenvalue	Observed	Expected
Total raw variance in observations =	27.4956	100.0%	100.0%
Raw variance explained by measures =	12.4956	45.4%	46.1%
Raw variance explained by persons =	5.0550	18.4%	18.7%
Raw Variance explained by items =	7.4406	27.1%	27.5%
Raw unexplained variance (total) =	15.0000	54.6%	100.0%
Unexplned variance in 1st contrast =	2.2082	8.0%	14.7%
Unexplned variance in 2nd contrast =	1.6982	6.2%	11.3%
Unexplned variance in 3rd contrast =	1.4748	5.4%	9.8%
Unexplned variance in 4th contrast =	1.2205	4.4%	8.1%
Unexplned variance in 5th contrast =	1.1059	4.0%	7.4%

FIGURA 5.29 – ANÁLISIS DE LAS COMPONENTES PRINCIPALES (PCA) DE LA ESCALA CISSVE

En la Figura 5.30 se pueden ver los 6 ítems que constituyen la segunda característica encontrada.

CONTRAST 1 FROM PRINCIPAL COMPONENT ANALYSIS  
 STANDARDIZED RESIDUAL LOADINGS FOR Item (SORTED BY LOADING)

CON-TRAST	LOADING	MEASURE	INFIIT	OUTFIT	ENTRY	LOADING	MEASURE	INFIIT	OUTFIT	ENTRY
		MNSQ	MNSQ	MNSQ	NUMBER Item			MNSQ	MNSQ	NUMBER Item
1	.54	-.49	1.09	1.08	A 5 -Pierdes la concentración?	-.51	-.18	.77	.74	a 2 -Notas incomodidad en tus ojos
1	.49	-.09	1.15	1.14	B 6 -Te cuesta recordar lo que has	-.47	-.14	.80	.78	b 10 -Te duelen los ojos?
1	.45	.09	.95	.93	C 15 -Tienes que releer la misma li	-.46	-.65	.58	.61	c 1 -Notas tus ojos cansados?
1	.41	-.07	1.06	1.07	D 14 -Te pierdes de línea al leer?	-.40	.01	1.01	.95	d 11 -Se te irritan los ojos?
1	.36	.48	1.15	1.13	E 9 -Te parece que lees lento?	-.28	.27	1.19	1.13	e 12 -Tienes sensación de "tirantez
1	.29	-.71	1.26	1.22	F 4 -Te entra sueño?	-.24	.03	.92	.93	f 13 -Notas que las palabras se pon
						-.23	1.04	1.15	1.04	g 7 -Ves doble?
						-.18	-.53	1.17	1.23	H 3 -Te duele la cabeza?
						-.13	.93	1.09	.89	G 8 -Te parece que las palabras se

FIGURA 5.30 – ÍTEMS DE LA ESCALA CISSVE QUE CONSTITUYEN LA SEGUNDA CARACTERÍSTICA ENCONTRADA

El coeficiente de desatenuación de la correlación entre el primer contraste y el segundo fue de 0.88 como se observa en la Figura 5.31, indicando que las dos dimensiones comparten la mayor parte de la varianza medida sobre las personas.

Approximate relationships between the Person measures

PCA Contrast	Item Clusters	Pearson Correlation	Disattenuated Correlation	Pearson+Extr Correlation	Disattenuated+Extr Correlation
1	1 - 3	0.5335	0.8041	0.5750	0.8752
1	1 - 2	0.5150	0.8105	0.5549	0.8844
1	2 - 3	0.6469	1.0000	0.6716	1.0000

FIGURA 5.31 – COEFICIENTE DE DESATENUACIÓN ENTRE EL PRIMER Y SEGUNDO CONTRASTE

A continuación se incorpora un resumen de los resultados de los tests realizados para examinar la unidimensionalidad y cómo se utilizaron los datos aportados para justificar por qué el CISSve se puede considerar unidimensional.

- 1) Se realizó un análisis de las componentes principales. La varianza explicada fue de un 45.4%, lo cual sugiere multidimensionalidad.
- 2) Al examinar los autovalores de primer contraste, se encontró uno con autovalor superior a 2, lo cuál nos sugiere también multidimensionalidad.

- 3) Sin embargo, evaluando la correlación entre las dos dimensiones, obtenemos que el coeficiente de desatenuación entre el primer y segundo contraste fue de 0.88, hecho que descarta la multidimensionalidad.

A continuación se muestra a través del análisis de las componentes principales, cuáles son los ítems más interdependientes, y cuáles los que menos tienen en común como se puede observar en la Figura 5.32. Esta información es muy útil a nivel diagnóstico.

LARGEST STANDARDIZED RESIDUAL CORRELATIONS  
USED TO IDENTIFY DEPENDENT ITEM

CORREL	ENTRY ITEM NUMBER	ENTRY ITEM NUMBER
.26	7 -Ves doble?	
.24	8 -Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?	8 -Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
.23	1 -Notas tus ojos cansados?	13 -Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
.21	5 -Pierdes la concentración?	2 -Notas incomodidad en tus ojos?
.20	4 -Te entra sueño?	6 -Te cuesta recordar lo que has leído?
.18	14 -Te pierdes de línea al leer?	5 -Pierdes la concentración?
		15 -Tienes que releer la misma línea de texto?
- .28	5 -Pierdes la concentración?	10 -Te duelen los ojos?
- .26	1 -Notas tus ojos cansados?	14 -Te pierdes de línea al leer?
- .23	2 -Notas incomodidad en tus ojos?	14 -Te pierdes de línea al leer?
- .21	4 -Te entra sueño?	11 -Se te irritan los ojos?
- .21	6 -Te cuesta recordar lo que has leído?	10 -Te duelen los ojos?
- .21	2 -Notas incomodidad en tus ojos?	15 -Tienes que releer la misma línea de texto?
- .20	5 -Pierdes la concentración?	7 -Ves doble?
- .20	2 -Notas incomodidad en tus ojos?	5 -Pierdes la concentración?
- .20	8 -Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?	15 -Tienes que releer la misma línea de texto?
- .18	3 -Te duele la cabeza?	5 -Pierdes la concentración?
- .18	1 -Notas tus ojos cansados?	5 -Pierdes la concentración?
- .18	3 -Te duele la cabeza?	9 -Te parece que lees lento?
- .18	9 -Te parece que lees lento?	13 -Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
- .18	4 -Te entra sueño?	13 -Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?

FIGURA 5.32 – INTERDEPENDENCIA ENTRE ÍTEMS (ANÁLISIS PCA CON WINSTEPS)

5.2.3.3.4 ÍNDICE DE SEPARACIÓN DE PERSONAS (PSI) Y NIVELES DE RENDIMIENTO DE CISSve

Utilizando el software Winsteps<sup>107</sup>, se obtuvo el valor de PSI que indica el número de diferentes niveles de síntomas que es capaz de medir el CISSve en nuestra muestra, según el método descrito por Wright<sup>151,170</sup>.

Para el CISSve, se obtuvo un PSI de 2.22 o 2.34 si excluimos a los 9 sujetos con las respuestas menos adaptadas al modelo (Outfit >2.5) como se puede ver en la Figura 5.33, Figura 5.34 y Figura 5.35.

Person	458 INPUT	439 MEASURED			INFIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	15.1	14.9	-1.44	.42	1.00	.0	.99	.0
P.SD	10.0	.3	1.18	.25	.45	1.3	.47	1.2
REAL RMSE	.48	TRUE SD	1.07	SEPARATION	2.22	Person	RELIABILITY	.83

Item	23 INPUT	15 MEASURED			INFIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	443.2	437.3	.00	.06	1.02	.1	.99	-.2
P.SD	130.7	1.7	.50	.01	.18	2.9	.18	2.5
REAL RMSE	.06	TRUE SD	.50	SEPARATION	7.69	Item	RELIABILITY	.98

FIGURA 5.33 – PERSON SEPARATION INDEX (WINSTEPS)

SUMMARY OF 439 MEASURED (EXTREME AND NON-EXTREME) Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	15.1	14.9	-1.44	.39				
P.SD	10.0	.3	1.18	.25				
S.SD	10.1	.3	1.18	.25				
MAX.	50.0	15.0	1.73	1.83				
MIN.	.0	13.0	-5.34	.27				
REAL RMSE	.48	TRUE SD	1.07	SEPARATION	2.22	Person	RELIABILITY	.83
MODEL RMSE	.46	TRUE SD	1.08	SEPARATION	2.34	Person	RELIABILITY	.85
S.E. OF Person MEAN = .06								

Person RAW SCORE-TO-MEASURE CORRELATION = .93  
 CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .90 SEM = 3.22

FIGURA 5.34 – PERSON SEPARATION INDEX AND RELIABILITY – EXTREME AND NON-EXTREME (WINSTEPS)

SUMMARY OF 430 MEASURED (NON-EXTREME) Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	15.5	14.9	-1.35	.36	1.00	.0	.99	.0
P.SD	9.9	.3	1.04	.14	.45	1.3	.47	1.2
S.SD	9.9	.3	1.05	.14	.45	1.3	.47	1.2
MAX.	50.0	15.0	1.73	1.01	2.48	3.3	2.74	3.5
MIN.	1.0	13.0	-4.13	.27	.15	-4.1	.16	-4.0
REAL RMSE	.41	TRUE SD	.96	SEPARATION	2.34	Person	RELIABILITY	.85
MODEL RMSE	.39	TRUE SD	.97	SEPARATION	2.52	Person	RELIABILITY	.86
S.E. OF Person MEAN = .05								

MINIMUM EXTREME SCORE: 9 Person 2.1%  
 DELETED: 19 Person

FIGURA 5.35 -PERSON SEPARATION INDEX AND RELIABILITY - NON-EXTREME (WINSTEPS)

## RESULTADOS

Como se puede observar, un PSI de 2.22 muestra una confiabilidad del 83%, lo que indica que el CISSve es capaz de distinguir entre 3.22 niveles de síntomas.

Si descartamos de nuestra muestra los 9 sujetos que peor se ajustan al modelo, la media de los síntomas de la muestra se acerca 0.1 logits a la media de los ítems (vemos que pasa de -1.44 a -1.35), y la confiabilidad del test aumenta, pudiéndose apreciar este hecho a través de los indicadores person reliability y PSI. Tanto el valor de confiabilidad de las personas (person reliability), que pasa de 0.83 a 0.85, como el PSI (person separation index), que pasa de 2.22 a 2.34, mejoran. De forma análoga el  $\alpha$  de Cronbach también mejora desde 83% a 85%, que lleva a que el CISSve sea capaz de distinguir entre 3,3 niveles de síntomas.

Por ello, obtenemos que el CISSve, es capaz de distinguir entre 3,3 niveles diferentes de síntomas en nuestra muestra, de acuerdo a la ecuación  $[(4*Sep+1)/3]$ , donde Sep corresponde a 2,22. Esto supone un nivel B según las recomendaciones de Khadka <sup>97,114</sup>.

Si hacemos un análisis de la escala mediante el método de Wright que analiza la escala con independencia de la muestra estudiada, mucho más indicado para evaluar la calidad de la escala cuando las distribuciones de las muestras con las que contamos están algo sesgadas hacia uno de los lados <sup>127</sup>, encontramos que se pueden distinguir hasta 6.3 niveles de síntomas significativamente distintos, como se puede apreciar numéricamente en la Figura 5.36.

RESULTADOS

Puntuación	Medida	SE	Normalizado	SE	Frecuencia	%	Frecuencia Acumulada	% Acumulado	Percentil	Mínima medida para siguiente Nivel (Indep Muestra) WRIGHT	Razón	Nivel
0	-5,34	1,83	169	155	9	2	9	2,1	1	-5,34	SE FIJA Porque -1.16 > -4.13	NIVEL 1
1	-4,13	1	271	85	8	2	17	3,9	3	-1,17		
2	-3,43	0,71	331	61	12	3	29	6,6	5	-1,41		
3	-3,01	0,59	366	50	22	5	51	11,6	9	-1,49		
4	-2,71	0,51	391	43	10	2	61	13,9	13	-1,54		
5	-2,48	0,46	411	39	16	4	77	17,5	16	-1,57		
6	-2,28	0,43	428	36	10	2	87	19,8	19	-1,58		
7	-2,11	0,4	442	34	21	5	108	24,6	22	-1,59		
8	-1,96	0,38	455	32	21	5	129	29,4	27	-1,60		
9	-1,83	0,36	467	30	13	3	142	32,3	31	-1,61		
10	-1,71	0,34	477	29	19	4	161	36,7	35	-1,62		
11	-1,59	0,33	487	28	16	4	177	40,3	38	-1,62	Porque -1.62 < -1.59	NIVEL 2
12	-1,49	0,32	496	27	19	4	196	44,6	42	-0,67		
13	-1,38	0,31	504	27	20	5	216	49,2	47	-0,68		
14	-1,29	0,31	513	26	23	5	239	54,4	52	-0,68		
15	-1,2	0,3	520	25	14	3	253	57,6	56	-0,70		
16	-1,11	0,29	528	25	18	4	271	61,7	60	-0,71		
17	-1,02	0,29	535	25	16	4	287	65,4	64	-0,71		
18	-0,94	0,29	542	24	13	3	300	68,3	67	-0,71		
19	-0,86	0,28	549	24	14	3	314	71,5	70	-0,72		
20	-0,78	0,28	556	24	11	3	325	74	73	-0,72		
21	-0,7	0,28	562	23	11	3	336	76,5	75	-0,72	Porque -0.72 < 0.7	NIVEL 3
22	-0,63	0,27	568	23	10	2	346	78,8	78	0,08		
23	-0,55	0,27	575	23	12	3	358	81,5	80	0,08		
24	-0,48	0,27	581	23	6	1	364	82,9	82	0,08		
25	-0,41	0,27	587	23	7	2	371	84,5	84	0,08		
26	-0,33	0,27	593	23	8	2	379	86,3	85	0,08		
27	-0,26	0,27	600	23	3	1	382	87	87	0,08		
28	-0,19	0,27	606	23	8	2	390	88,8	88	0,08		
29	-0,12	0,27	612	23	5	1	395	90	89	0,08		
30	-0,05	0,27	618	23	6	1	401	91,3	91	0,08		
31	0,03	0,27	624	23	4	1	405	92,3	92	0,08		
32	0,1	0,27	630	23	4	1	409	93,2	93	0,08	Porque 0.08 < 0.10	NIVEL 4
33	0,17	0,27	637	23	4	1	413	94,1	94	0,86		
34	0,25	0,27	643	23	1	0	414	94,3	94	0,86		
35	0,32	0,28	649	23	5	1	419	95,4	95	0,88		
36	0,4	0,28	656	24	3	1	422	96,1	96	0,88		
37	0,48	0,28	662	24	4	1	426	97	97	0,88		
38	0,56	0,28	669	24	2	1	428	97,5	97	0,88		
39	0,64	0,29	676	24	2	1	430	97,9	98	0,89		
40	0,72	0,29	683	25	2	1	432	98,4	98	0,89		
41	0,8	0,29	690	25	0	0	432	98,4	98	0,89		
42	0,89	0,3	697	25	1	0	433	98,6	99	0,91		
43	0,98	0,3	705	26	1	0	434	98,9	99	0,91	Porque 0.91 < 0.98	NIVEL 5
44	1,07	0,31	713	26	0	0	434	98,9	99	1,84		
45	1,17	0,31	721	27	1	0	435	99,1	99	1,84		
46	1,27	0,32	730	27	0	0	435	99,1	99	1,86		
47	1,38	0,33	739	28	1	0	436	99,3	99	1,87		
48	1,49	0,34	748	29	1	0	437	99,5	99	1,89		
49	1,61	0,35	758	30	1	0	438	99,8	99	1,90		
50	1,73	0,36	769	31	1	0	439	100	99	1,92		
51	1,87	0,38	781	32	0	0	439	100	100	1,95		
52	2,02	0,4	793	34	0	0	439	100	100	1,98	Porque 1.98 < 2.02	NIVEL 6
53	2,18	0,42	807	35	0	0	439	100	100	3,18		
54	2,37	0,44	823	38	0	0	439	100	100	3,21		
55	2,58	0,48	841	41	0	0	439	100	100	3,27		
56	2,84	0,53	863	45	0	0	439	100	100	3,35		
57	3,15	0,6	890	51	0	0	439	100	100	3,46		
58	3,59	0,73	927	62	0	0	439	100	100	3,68		
59	4,32	1,02	988	86	0	0	439	100	100	4,21	Porque 4.21 < 4.32	NIVEL 7
60	5,54	1,83	1092	156	0	0	439	100	100	5,77		

FIGURA 5.36 – NIVELES DE SÍNTOMAS DE LA ESCALA CISSVe ESTADÍSTICAMENTE DIFERENTES CON INDEPENDENCIA DE LA MUESTRA DE ACUERDO AL MÉTODO DE WRIGHT

Esto se traduce en una fiabilidad de la escala de un 98% (basada en 6.3 niveles de síntomas), es decir, cuando hacemos el análisis con independencia de la muestra. Winsteps nos proporciona los datos necesarios de equivalencia entre la puntuación obtenida en la escala y su

## RESULTADOS

correspondiente medida en logits, con los errores standard asociados y el máximo número de niveles estadísticamente diferentes que se pueden obtener utilizando el método de Wright y Gothwal <sup>7,170</sup>. Recordemos conceptualmente que el error standard nos informa de los intervalos de confianza de la medida, es decir, los valores entre los que se encontraría la medida si escogiera otra muestra entre la población general (los cálculos de Winsteps asumen un  $p=0.05$ , es decir, que otorgan una confianza del 95%).

En la Figura 5.37 se puede ver representada la curva característica del test con el detalle de la correspondencia entre la puntuación obtenida en la escala CISSve y su nivel correspondiente de síntomas.

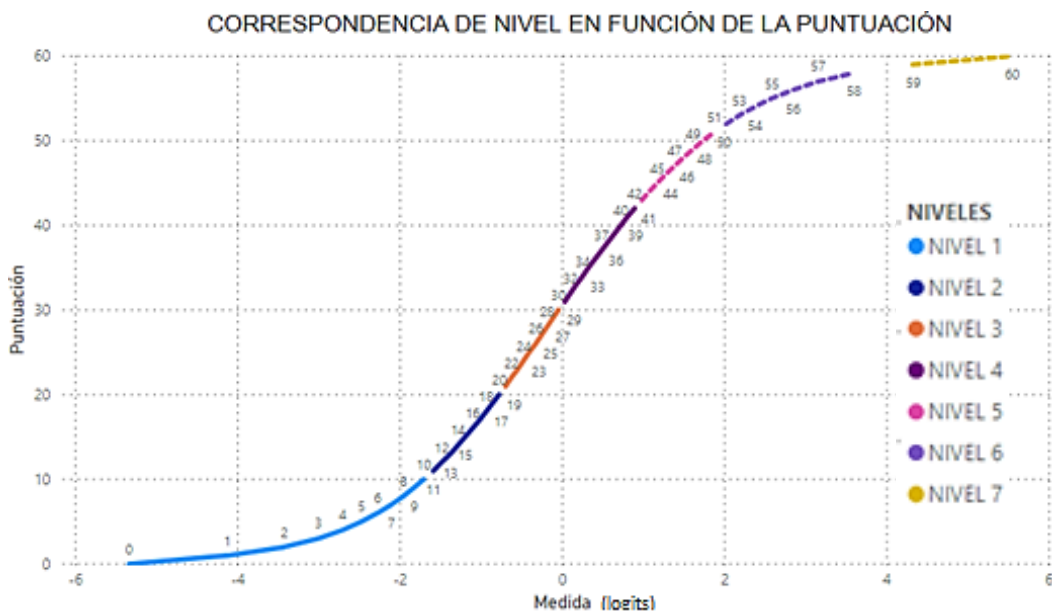


FIGURA 5.37 – CORRESPONDENCIA DE NIVEL DE SÍNTOMAS EN FUNCIÓN DE LA PUNTUACIÓN OBTENIDA EN LA ESCALA CISSVE

En el gráfico de la Figura 5.38 se observa la distribución de nuestra muestra por nivel de síntomas, dejando patente el hecho de que predominantemente nuestra muestra está en la parte no sintomática de la escala, es decir, prácticamente se usan sólo los 4 niveles más bajos (3.3), de los que es capaz de medir la escala CISSve.

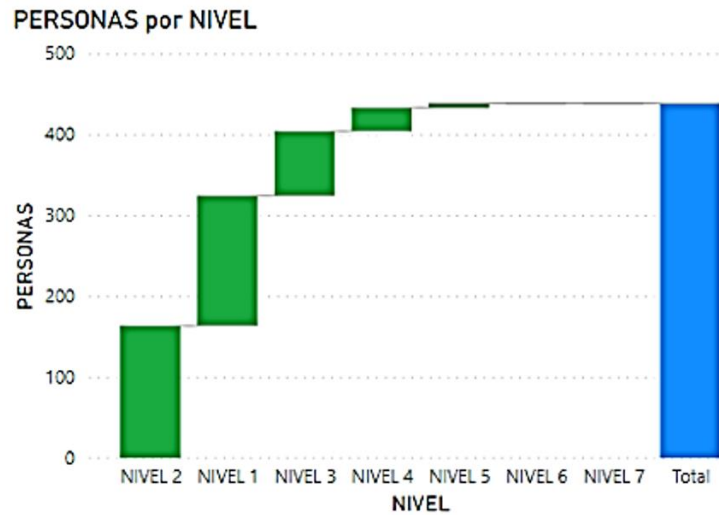


FIGURA 5.38 – DISTRIBUCIÓN DE LA MUESTRA EN CUANTO A SUS NIVELES DE SÍNTOMAS

Si observamos el gráfico de la Figura 5.39 podemos ver más nítidamente expresando en términos porcentuales, el hecho de que no hay prácticamente sujetos en los niveles altos de síntomas. La inmensa mayoría se encuentran en los niveles de síntomas 1 (35%), 2 (38%) y 3 (19%).

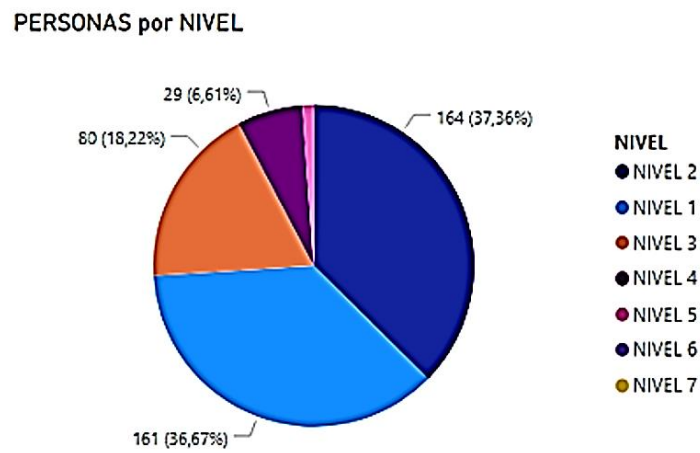


FIGURA 5.39 – DISTRIBUCIÓN PORCENTUAL DE LA MUESTRA EN LOS DIFERENTES NIVELES DE SÍNTOMAS

#### 5.2.3.3.5 ADECUACIÓN ESCALA-MUESTRA (“TARGETING”)

Como queda reflejado en la Figura 5.40, la muestra está 1.44 logits desplazada hacia la parte no sintomática (negativa) y la desviación típica de nuestra muestra se sitúa en 1.18, es decir, la muestra elegida tiene una dispersión ligeramente superior a la de los ítems (0,50).

RESULTADOS

SUMMARY OF 439 MEASURED (EXTREME AND NON-EXTREME) Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	15.1	14.9	-1.44	.39				
P.SD	10.0	.3	1.18	.25				
S.SD	10.1	.3	1.18	.25				
MAX.	50.0	15.0	1.73	1.83				
MIN.	.0	13.0	-5.34	.27				
REAL RMSE	.48	TRUE SD	1.07	SEPARATION	2.22	Person RELIABILITY	.83	
MODEL RMSE	.46	TRUE SD	1.08	SEPARATION	2.34	Person RELIABILITY	.85	
S.E. OF Person MEAN = .06								

Person RAW SCORE-TO-MEASURE CORRELATION = .93  
 CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .90 SEM = 3.22

FIGURA 5.40 – TARGETING DE LA ESCALA CISSVE PARA LA MUESTRA ESTUDIADA

Si observamos los valores medidos en nuestra muestra y descartamos a los 9 sujetos que peor se ajustan al modelo, observamos que la media de los síntomas de muestra se acerca 0.1 logits a la media de los ítems, tal y como se aprecia en la Figura 5.41, de forma que se reduce el desplazamiento entre sujetos e ítems a 1.35 logits, disminuyendo también la desviación típica a 1.04 logits.

SUMMARY OF 430 MEASURED (NON-EXTREME) Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	15.5	14.9	-1.35	.36	1.00	.0	.99	.0
P.SD	9.9	.3	1.04	.14	.45	1.3	.47	1.2
S.SD	9.9	.3	1.05	.14	.45	1.3	.47	1.2
MAX.	50.0	15.0	1.73	1.01	2.48	3.3	2.74	3.5
MIN.	1.0	13.0	-4.13	.27	.15	-4.1	.16	-4.0
REAL RMSE	.41	TRUE SD	.96	SEPARATION	2.34	Person RELIABILITY	.85	
MODEL RMSE	.39	TRUE SD	.97	SEPARATION	2.52	Person RELIABILITY	.86	
S.E. OF Person MEAN = .05								

MINIMUM EXTREME SCORE: 9 Person 2.1%  
 DELETED: 19 Person

FIGURA 5.41 – TARGETING DE LA ESCALA CISSVE PARA LA MUESTRA ESTUDIADA, DESCARTANDO LOS-SUJETOS QUE PEOR SE ADAPTAN AL MODELO (N=430, M=15.5, DT=9,9)

Observando ítems y sujetos en un mapa de Wright, que enfrenta horizontalmente los ítems y a los sujetos que presentan idéntico nivel de dificultad/síntomas, se puede ver de una forma muy rápida el ajuste de la muestra poblacional a la dificultad de los ítems y sacar rápidamente las mismas conclusiones a través del análisis gráfico que se ve en la Figura 5.43. Antes, en la Figura 5.42, se pueden ver dos histogramas enfrentados con las distribuciones de ítems y personas, que claramente constatan la misma idea. Para la muestra completa de 439 sujetos, existe un desplazamiento de 1.44 logits entre personas e ítems, hacia la parte menos sintomática.



## RESULTADOS

Podemos observar en la Figura 5.44, la tabla que muestra la probabilidad de que un paciente sea clasificado como “sintomático” por un ítem dependiendo de la distancia en logits del sujeto a ese ítem para el caso de un modelo dicotómico de escala con sólo dos categorías de respuesta (sintomático/no sintomático). Conceptualmente, un sujeto que esté alineado con un ítem tendrá un 50% de posibilidades de acertar ese ítem. Llevándolo a nuestro caso (modelo politómico con 5 categorías de respuesta), un sujeto alineado en sintomatología con la dificultad de un ítem (es decir, con una distancia sujeto-ítem expresada en logits de 0), tendría una probabilidad del 50% de elegir la categoría de sintomatología “media” (la categoría 2 en una escala Likert gradada de 0 a 4). Sin embargo, por ejemplo, un sujeto alejado -1.4 logits de ese mismo ítem (sujeto menos sintomático que el anterior) tendría solo un 20% de probabilidad de elegir la categoría de sintomatología “media” (2) de ese ítem, y aún menos probabilidad de elegir categorías más altas de sintomatología (3 o 4).

Logit Difference	Probability of Success	Probit Difference	Logit Difference	Probability of Success	Probit Difference
5.0	99%	2.48	-5.0	1%	-2.48
4.6	99%	2.33	-4.6	1%	-2.33
4.0	98%	2.10	-4.0	2%	-2.10
3.0	95%	1.67	-3.0	5%	-1.67
2.2	90%	1.28	-2.2	10%	-1.28
2.0	88%	1.18	-2.0	12%	-1.18
1.4	80%	0.85	-1.4	20%	-0.85
1.1	75%	0.67	-1.1	25%	-0.67
1.0	73%	0.62	-1.0	27%	-0.62
0.8	70%	0.50	-0.8	30%	-0.50
0.69	67%	0.43	-0.69	33%	-0.43
0.5	62%	0.31	-0.5	38%	-0.31
0.4	60%	0.25	-0.4	40%	-0.25
0.2	55%	0.13	-0.2	45%	-0.13
0.1	52%	0.06	-0.1	48%	-0.06
0.0	50%	0.00	-0.0	50%	-0.00

FIGURA 5.44 – PROBABILIDAD DE ÉXITO DE UN SUJETO SOBRE UN ÍTEM EN FUNCIÓN DE LA DISTANCIA EN LOGITS DE ESE SUJETO A ESE ÍTEM <sup>107</sup>

A continuación, se presenta un resumen del nivel de dificultad de cada ítem de la escala CISSve. Primero en la Figura 5.45 de forma numérica, y a continuación en la Figura 5.46, de forma gráfica. Los ítems están ordenados de más a menos difícil, significando esto, que los ítems más “difíciles”, serán capaces de discriminar a pacientes con mayor sintomatología, mientras que los más “fáciles” no separarán tan fácilmente a esa parte de la muestra del resto de la misma (predominantemente asintomática). Por ejemplo, un sujeto que tuviera pocos síntomas, en principio, podría notar que le entra sueño leyendo, o cansancio en los ojos al final del día, pero sin embargo, con una sintomatología pequeña o inexistente, sería más extraño que presentara visión doble, obteniendo por lo tanto puntuaciones más bajas en el ítem 7 (“ves doble?”) que en los otros dos. Sin embargo, en el caso de un sujeto con alto grado de sintomatología, no sería extraño que además de obtener puntuaciones altas en los ítems más fáciles, también presentara valores elevados en la característica que examina el ítem 7, ¿“ves doble?”, ya que, en ese caso, sería más probable que su disfunción binocular, le provocara una ruptura de la fusión. Por lo tanto, el ítem 7, al ser capaz de separar a sujetos con mayor sintomatología se encontraría en una parte más alta de la escala en logits que los otros dos

ítems, correspondiendo su valor en “logits” al de sujetos con sintomatologías mayores (es decir, también con valores en “logits” más elevados que los “menos sintomáticos”).

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT  MNSQ	OUTFIT  MNSQ	PTMEASUR-AL  CORR.	EXACT  OBS%	MATCH  EXP%	Item		
7	202	437	1.04	.08	1.15	1.61	1.04	.48	.48	65.0	65.1	-Ves doble?
8	219	438	.93	.08	1.09	1.01	.89	.55	.50	66.9	63.4	-Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
9	311	439	.48	.07	1.15	1.91	1.13	.52	.55	51.9	54.2	-Te parece que lees lento?
12	359	437	.27	.06	1.19	2.51	1.13	.55	.57	46.7	51.0	-Tienes sensación de "fritantez" alrededor de los ojos?
15	406	437	.09	.06	.95	-.71	.93	.60	.59	50.5	48.4	-Tienes que releer la misma línea de texto?
13	417	433	.03	.06	.92	-1.11	.93	.64	.59	48.8	48.2	-Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
11	428	438	.01	.06	1.01	.21	.95	.61	.60	50.1	47.9	-Se te irritan los ojos?
14	447	436	-.07	.06	1.06	.91	1.07	.56	.60	48.9	46.5	-Te pierdes de línea al leer?
6	456	439	-.09	.06	1.15	2.11	1.14	.57	.60	45.6	46.4	-Te cuesta recordar lo que has leído?
10	468	436	-.14	.06	.80	-3.11	.78	.66	.61	47.8	46.1	-Te duelen los ojos?
2	484	439	-.18	.06	.77	-3.71	.74	.67	.61	51.6	44.8	-Notas incomodidad en tus ojos?
5	577	437	-.49	.06	1.09	1.41	1.08	.61	.64	42.5	43.6	-Pierdes la concentración?
3	593	439	-.53	.06	1.17	2.61	1.23	.57	.64	39.1	43.6	-Te duele la cabeza?
1	631	439	-.65	.05	.58	-7.61	.61	.72	.65	56.5	42.6	-Notas tus ojos cansados?
4	650	435	-.71	.05	1.26	3.81	1.22	.60	.65	39.0	42.5	-Te entra sueño?
MEAN	443.2	437.3	.00	.06	1.02	.11	.99	-.21		50.1	48.9	
P.SD	130.7	1.7	.50	.01	.18	2.91	.18	2.51		7.7	6.7	

FIGURA 5.45 – CLASIFICACIÓN DE LOS ÍTEMS DE LA ESCALA CISSVE SEGÚN SU NIVEL DE DIFICULTAD

## RESULTADOS

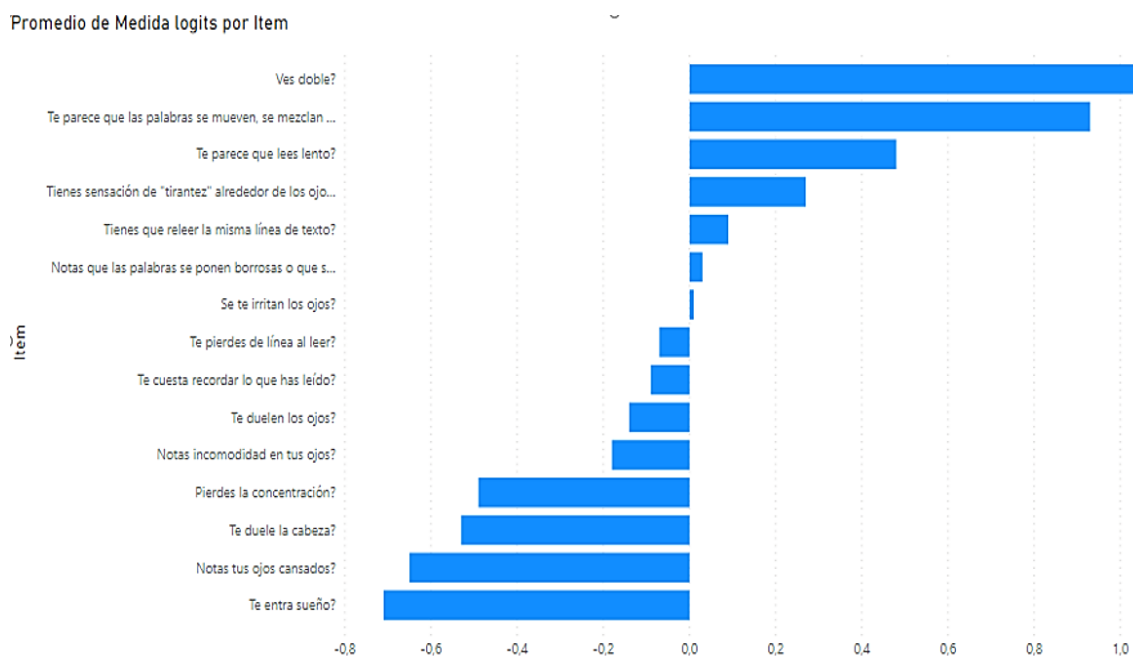


FIGURA 5.46— EXPRESIÓN GRÁFICA DE LOS ÍTEMS DE LA ESCALA CISSVE SEGÚN SU NIVEL DE DIFICULTAD

### 5.2.3.3.6 FUNCIONAMIENTO DIFERENCIAL DE LOS ELEMENTOS MUESTRALES (DIF) POR SEXO, GRUPO DE EDAD Y VERSIÓN CISS (CISSv15 VS CISSve)

Dentro del análisis, también se realizó un examen para comparar entre diferentes grupos de nuestra muestra y valorar si la escala se comporta de forma diferente, o por el contrario lo hace de forma similar, en los distintos grupos.

Las comparaciones fueron realizadas en cuanto al sexo y también por grupos de edad. Para llevar a cabo el análisis, se examinó cada ítem para poder determinar si existían diferencias entre alguno de los subgrupos, masculino versus femenino; adultos jóvenes (18-30 años) vs niños (9-17 años), examinando cómo cada uno de ellos respondió a cada ítem.

La prueba DIF es también una forma útil de validar traducciones de cuestionarios <sup>171</sup>, observando a través de este análisis si los ítems se comportan de forma equivalente a los de la versión original. Así pues, comparamos el comportamiento del CISSve con el de la versión original del CISSv15, descartando posibles problemas derivados de la adaptación que se verían si existiera DIF en alguno de los ítems dentro de los grupos comparados.

El DIF encontrado en uno de los ítems (ítem 14) se consideró un problema transcultural o traslacional para esa traducción en particular<sup>115,123,162</sup>. Siguiendo el método de diferencia de logit (regresión logística), los resultados observados sobre el DIF revelaron que no había diferencias notables de DIF en cuanto al sexo en ninguno de los ítems del CISSve, como se observa en la Figura 5.47.

Person CLASS/	OBSERVATIONS COUNT	AVERAGE	BASELINE EXPECT MEASURE	DIF SCORE MEASURE	DIF SCORE MEASURE	DIF SIZE	DIF S.E.	DIF t	Item Prob.	Number	Name
F	240	1.63	1.56	-.65	.07	-.73	.07	-1.18	.2400	1	-Notas tus ojos cansados?
M	190	1.27	1.35	-.65	-.09	-.53	.09	1.36	.1744	1	-Notas tus ojos cansados?
F	240	1.26	1.20	-.18	.06	-.26	.08	-1.03	.3063	2	-Notas incomodidad en tus ojos?
M	190	.96	1.03	-.18	-.07	-.07	.09	1.19	.2363	2	-Notas incomodidad en tus ojos?
F	240	1.59	1.47	-.53	.12	-.69	.07	-2.13	.0342	3	-Te duele la cabeza?
M	190	1.12	1.27	-.53	-.15	-.31	.09	2.47	.0142	3	-Te duele la cabeza?
F	239	1.60	1.62	-.71	-.02	-.69	.07	.30	.7661	4	-Te entra sueño?
M	187	1.43	1.41	-.71	.02	-.74	.08	-.35	.7266	4	-Te entra sueño?
F	239	1.39	1.44	-.49	-.05	-.43	.07	.82	.4130	5	-Pierdes la concentración?
M	189	1.30	1.24	-.49	.06	-.57	.09	-.97	.3315	5	-Pierdes la concentración?
F	240	1.10	1.13	-.09	-.03	-.04	.08	.55	.5832	6	-Te cuesta recordar lo que has leído?
M	190	1.01	.97	-.09	.04	-.15	.09	-.68	.4973	6	-Te cuesta recordar lo que has leído?
F	239	.49	.51	1.04	-.02	1.09	.10	.50	.6200	7	-Ves doble?
M	189	.45	.42	1.04	.03	.96	.12	-.66	.5106	7	-Ves doble?
F	239	.52	.55	.93	-.03	1.00	.10	.65	.5160	8	-Te parece que las palabras se mueven, se mezclan o flotan sobre?
M	190	.49	.46	.93	.04	.84	.12	-.84	.4015	8	-Te parece que las palabras se mueven, se mezclan o flotan sobre?
F	240	.67	.78	.48	-.11	.68	.09	2.22	.0277	9	-Te parece que lees lento?
M	190	.79	.65	.48	.14	.21	.10	-2.71	.0073	9	-Te parece que lees lento?
F	237	1.22	1.17	-.14	.05	-.20	.08	-.85	.3969	10	-Te duelen los ojos?
M	190	.94	1.00	-.14	-.06	-.05	.09	.97	.3311	10	-Te duelen los ojos?
F	239	1.08	1.07	.01	.01	-.01	.08	-.27	.7874	11	-Se te irritan los ojos?
M	190	.89	.91	.01	-.02	.04	.09	.29	.7724	11	-Se te irritan los ojos?
F	240	.91	.90	.27	.01	.27	.08	.00	1.000	12	-Tienes sensación de "tirantez" alrededor de los ojos?
M	188	.75	.76	.27	-.01	.27	.10	.00	1.000	12	-Tienes sensación de "tirantez" alrededor de los ojos?
F	238	1.05	1.05	.03	.01	.03	.08	.00	1.000	13	-Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
M	186	.89	.90	.03	-.01	.03	.09	.00	1.000	13	-Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
F	238	1.04	1.12	-.07	-.08	-.05	.08	1.50	.1348	14	-Te pierdes de línea al leer?
M	189	1.06	.95	-.07	.10	-.23	.09	-1.81	.0726	14	-Te pierdes de línea al leer?
F	240	1.03	1.02	.09	.02	.06	.08	-.33	.7431	15	-Tienes que releer la misma línea de texto?
M	188	.84	.86	.09	-.02	.12	.10	.36	.7190	15	-Tienes que releer la misma línea de texto?

FIGURA 5.47 – ANÁLISIS DEL DIF EN CUANTO AL SEXO (WINSTEPS)

## RESULTADOS

Gráficamente a través de la Figura 5.48, se puede observar como el tamaño del DIF no excede los 0.50 logits en ninguno de los ítems.

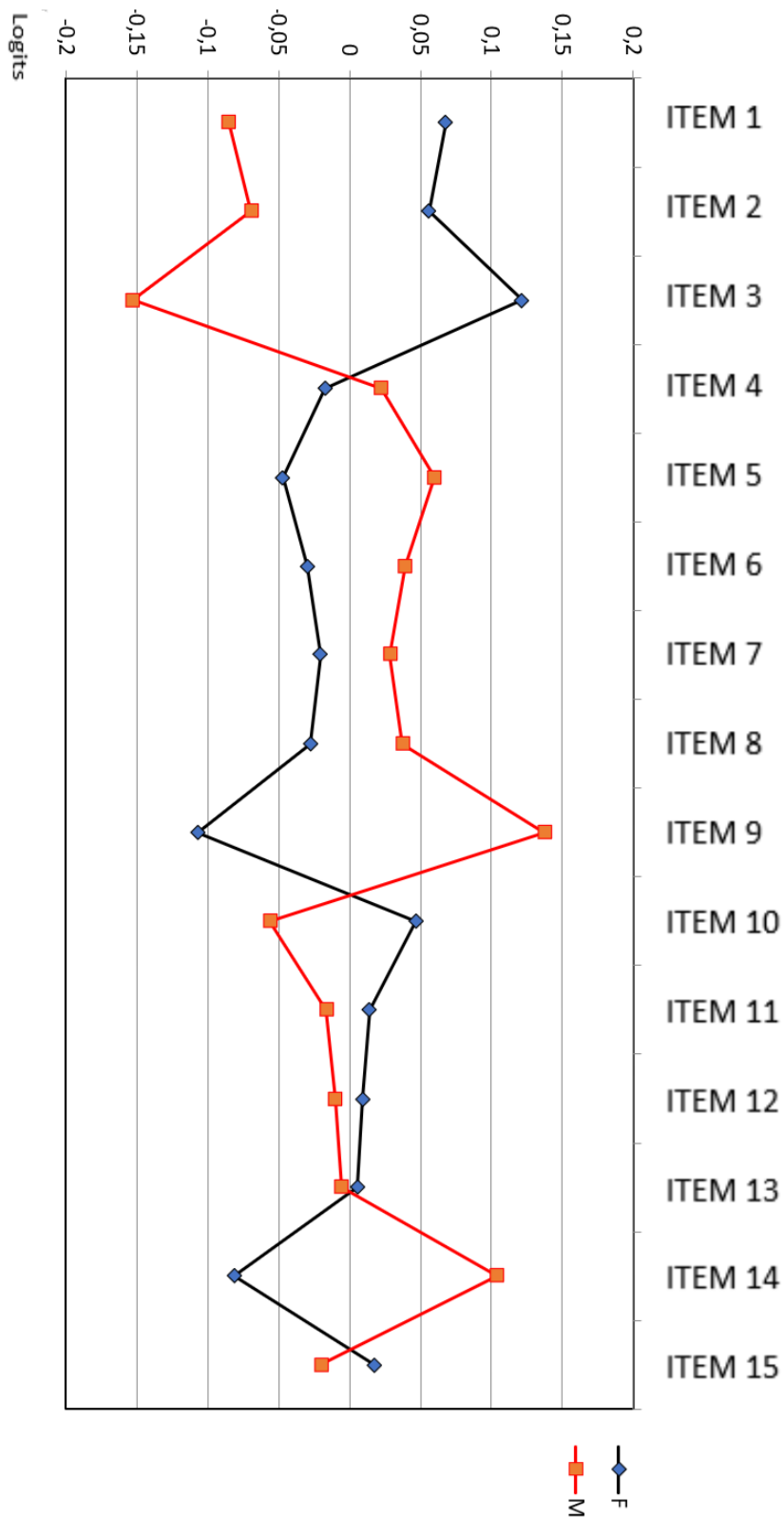


FIGURA 5.48 – REPRESENTACIÓN GRÁFICA DEL ANÁLISIS DEL DIF EN CUANTO AL SEXO (WINSTEPS)

En la comparación entre adultos jóvenes (18-30 años) y niños (9-17 años), sólo el ítem 14 (“¿te pierdes de línea al leer?”) mostró una desviación mínima según los criterios definidos por Linacre<sup>107</sup>, presentando un valor de 0.80. Este ítem resultó ser más difícil para adultos jóvenes que para niños, es decir, la puntuación fue más baja en adultos, lo que quiere decir que, como es lógico, los adultos jóvenes se perdían menos de línea al leer de media que los niños entre 9 y 17.

El ítem 2 (“Notas incomodidad en tus ojos”) se quedó en el límite, ya que como se puede ver en la Figura 5.49 el “DIF Size” tiene un valor de 0,43 y la prueba t, que evalúa si puede considerarse como válida la hipótesis del investigador H1, de que haya diferencia, corrobora que se podría afirmar, pues el p-valor es menor o muy próximo a 0.01.

Estas pequeñas diferencias resultan aceptables, considerando en primer lugar a nivel matemático que esta pequeña diferencia supone un impacto mínimo en la puntuación total constituida por la suma de las puntuaciones de los 15 ítems, y en segundo lugar que a nivel clínico, al fijarse diferentes umbrales en niños y adultos jóvenes de cara a su clasificarlos como sintomáticos o asintomáticos, la diferencia no resulta relevante, ya que queda absorbida por el propio criterio (16 puntos para niños<sup>59</sup> y 21 para adultos jóvenes<sup>60</sup>)

RESULTADOS

Person CLASS/	OBSERVATIONS COUNT	AVERAGE	BASELINE EXPECT MEASURE	DIF SCORE MEASURE	DIF SCORE MEASURE	DIF SITE	DIF S.E.	DIF t	Item Number	Name
AJ	181	1.81	1.68	-.59	.14	-.74	-.15	.08	1	-Notas tus ojos cansados?
ME	249	1.18	1.27	-.59	-.10	-.45	.13	.07	1	-Notas tus ojos cansados?
AJ	181	1.48	1.30	-.14	.18	-.36	-.22	.08	2	-Notas incomodidad en tus ojos?
ME	249	.84	.96	-.14	-.13	.07	.21	.08	2	-Notas incomodidad en tus ojos?
AJ	181	1.58	1.59	-.49	-.01	-.49	.00	.00	3	-Te duele la cabeza?
ME	249	1.21	1.20	-.49	.01	-.49	.00	.00	3	-Te duele la cabeza?
AJ	181	1.75	1.77	-.70	-.03	-.67	.03	.08	4	-Te entra sueño?
ME	245	1.38	1.36	-.70	.02	-.72	-.03	.07	4	-Te entra sueño?
AJ	181	1.55	1.59	-.48	-.04	-.44	.04	.08	5	-Pierdes la concentración?
ME	247	1.23	1.20	-.48	.03	-.52	-.04	.07	5	-Pierdes la concentración?
AJ	181	1.30	1.27	-.10	.04	-.15	-.05	.08	6	-Te cuesta recordar lo que has leído?
ME	249	.91	.94	-.10	-.03	-.06	.04	.08	6	-Te cuesta recordar lo que has leído?
AJ	181	.51	.59	.96	-.07	1.11	.16	.11	7	-Ves doble?
ME	246	.48	.42	.96	.06	.81	-.15	.10	7	-Ves doble?
AJ	181	.65	.63	.87	.03	.83	-.05	.10	8	-Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
ME	248	.43	.45	.87	-.02	.92	.04	.11	8	-Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
AJ	181	.80	.87	.45	-.07	.56	.11	.10	9	-Te parece que lees lento?
ME	249	.68	.63	.45	.05	.34	-.11	.09	9	-Te parece que lees lento?
AJ	181	1.27	1.26	-.09	.01	-.09	.00	.08	10	-Te duelen los ojos?
ME	246	.93	.93	-.09	.00	-.09	.00	.08	10	-Te duelen los ojos?
AJ	181	1.32	1.18	.01	.14	-.17	-.18	.08	11	-Se te irritan los ojos?
ME	248	.77	.87	.01	-.10	.19	.18	.09	11	-Se te irritan los ojos?
AJ	181	1.03	1.01	.24	.02	.22	-.03	.09	12	-Tienes sensación de "tiranterez" alrededor de los ojos?
ME	248	.73	.74	.24	-.01	.27	.02	.09	12	-Tienes sensación de "tiranterez" alrededor de los ojos?
AJ	181	1.12	1.17	.03	-.05	.09	.06	.09	13	-Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
ME	244	.89	.86	.03	.04	-.03	-.06	.08	13	-Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
AJ	181	.93	1.25	-.08	-.32	.36	.44	.09	14	-Te pierdes de línea al leer?
ME	246	1.16	.92	-.08	.24	-.44	-.36	.08	14	-Te pierdes de línea al leer?
AJ	181	1.16	1.11	.10	.05	.04	-.06	.09	15	-Tienes que releer la misma línea de texto?
ME	246	.78	.82	.10	-.03	.16	.06	.08	15	-Tienes que releer la misma línea de texto?

FIGURA 5.49 - ANÁLISIS DEL DIF EN CUANTO A LA EDAD (WINSTEPS)

Se puede ver la representación gráfica de la expresión de los tamaños del DIF en la Figura 5.50 que muestra la perspectiva visual de lo analizado.

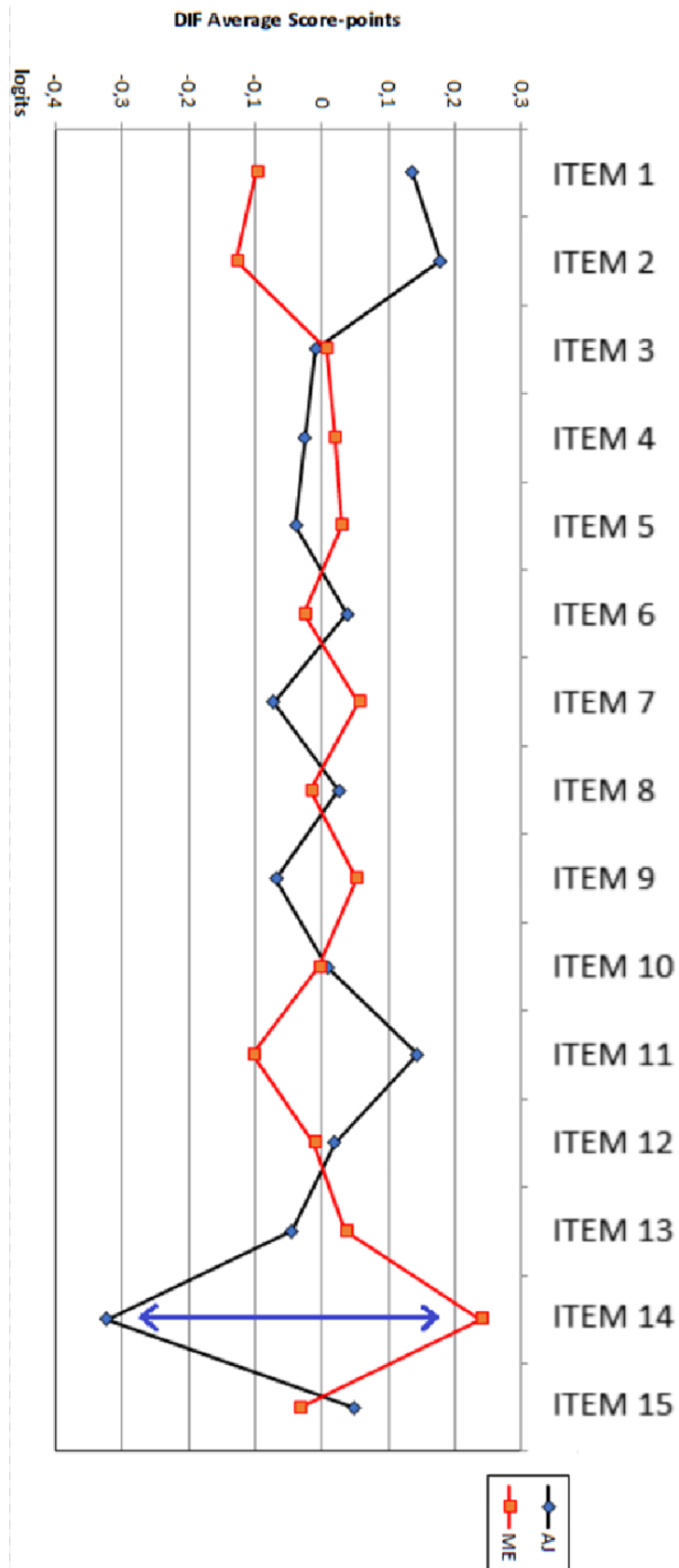


FIGURA 5.50 – REPRESENTACIÓN GRÁFICA DEL ANÁLISIS DEL DIF EN CUANTO A LA EDAD (WINSTEPS)

## RESULTADOS

Al realizar el mismo análisis diferencial que nos facilita Winsteps a través del método de Mantel que es la aplicación a muestras politómicas del método de Mantel-Haenszel y según Linacre <sup>107</sup> provee una aproximación más exacta que el método de Rasch-Welch cuando las muestras presentan respuestas completas (como es nuestro caso) <sup>107,172</sup>, se observaron hallazgos parecidos, aunque algunas pequeñas diferencias se hicieron más visibles.

En el análisis en el que se compararon los resultados en cuanto al sexo se puede ver que prácticamente no hay DIF en ninguno de los ítems (véase la Figura 5.51). Sólo en el ítem 3 (“Te duele la cabeza?”) se puede ver un p-valor de 0.003 que indica que podemos rechazar con un error de un 0.3% la hipótesis nula (H0), que no hay diferencia entre la forma en que contestan sujetos de sexo masculino y sujetos de sexo femenino a ese ítem, y por lo tanto aceptar la hipótesis del investigador (H1), que nos dice que existen diferencias entre cómo contestan a ese ítem los sujetos de sexo diferente.

La comparación entre adultos jóvenes (18-30) y niños (9-17) por este método (véase Figura 5.52), desvela que tanto en el ítem 1 (“¿Notas tus ojos cansados?”), como en el ítem 2 (“¿Notas incomodidad en tus ojos?”) como en el ítem 14 (“¿Te pierdes de línea al leer?”), la significancia de la prueba chi-cuadrado de homogeneidad, arroja un p-valor de 0.0005 y que en consecuencia, se puede afirmar con un error inferior a un 0.05%, que la hipótesis del investigador H1 de que haya diferencia entre las respuestas de adultos jóvenes y niños en esos ítems, es verdadera.

DIF class/group specification is: DIF=58M1

Person Obs-Exp	DIF	Person Obs-Exp	DIF	Person Obs-Exp	DIF	Person Obs-Exp	DIF	JOINT	Rasch-Welch	Mantel	Size	Active	Item					
CLASS/	Average	MEASURE	S.E.	CLASS/	Average	MEASURE	S.E.	CONTRAST	S.E.	t	d.f.	Prob.	Chi-sq	Prob.	CUMULON	Slices	Number	Name
F	.07	-.72	.07	M	-.09	-.53	.09	-.20	.11	-1.80	385	.0724	5.8450	.0156	-.52	38	1	-Horas tus ojos cansados?
F	.06	-.26	.08	M	-.07	-.07	.09	-.19	.12	-1.57	389	.1175	1.9557	.1620	-.30	38	2	-Horas incomodidad en tus ojos?
F	.12	-.69	.07	M	-.18	-.31	.09	-.37	.11	-3.26	389	.0012	13.3429	.0003	-.74	38	3	-Te duele la cabeza?
F	-.02	-.69	.07	M	.02	-.74	.08	.05	.11	.46	394	.6460	.1799	.0711	.06	38	4	-Te entra sueño?
F	-.05	-.42	.07	M	.06	-.57	.09	.14	.11	1.27	398	.2038	1.2876	.2565	.22	38	5	-Pierdes la concentración?
F	-.03	-.04	.08	M	.04	-.15	.09	.10	.12	.87	398	.8827	.3049	.5809	.11	38	6	-Te cuesta recordar lo que has leído?
F	-.02	1.09	.10	M	.03	.96	.12	.13	.16	.82	396	.4106	.8439	.3583	.22	38	7	-Tea doble?
F	-.02	1.00	.10	M	.04	.84	.12	.16	.15	1.06	400	.8991	2.0018	.1571	.36	38	8	-Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
F	-.11	.68	.09	M	.14	.21	.10	.47	.13	3.50	413	.0005	10.3824	.0013	.72	38	9	-Te parece que lees lento?
F	.05	-.20	.08	M	-.06	-.05	.09	-.15	.12	-1.29	389	.1972	1.0508	.3083	-.22	38	10	-Te duelen los ojos?
F	.01	-.01	.08	M	-.02	.04	.09	-.05	.12	-.40	392	.6928	.0020	.9640	-.01	38	11	-Se te irritan los ojos?
F	.01	.27	.08	M	-.01	.27	.10	.00	.13	.00	387	1.000	.0011	.9732	-.01	38	12	-Tienes sensación de "vibrar" alrededor de los ojos?
F	.01	.02	.08	M	-.01	.02	.09	.00	.12	.00	385	1.000	.0163	.8982	-.02	38	13	-Horas que las palabras se ponen borrosas o que se enfocan y desenfocan?
F	-.08	.05	.08	M	.10	-.22	.09	.28	.12	2.35	402	.0194	2.7262	.0987	.35	38	14	-Te pierdes de línea al leer?
F	.02	.06	.08	M	-.02	.12	.10	-.06	.12	-.49	386	.6268	.1080	.7424	-.07	38	15	-Tienes que releer la misma línea de texto?
M	-.09	-.52	.09	F	.07	-.72	.07	.20	.11	1.80	395	.0724	5.8450	.0156	.52	38	1	-Horas tus ojos cansados?
M	-.07	-.07	.09	F	.06	-.26	.08	.19	.12	1.57	389	.1175	1.9557	.1620	.30	38	2	-Horas incomodidad en tus ojos?
M	-.15	-.31	.09	F	.12	-.69	.07	.37	.11	3.26	389	.0012	13.3429	.0003	.74	38	3	-Te duele la cabeza?
M	.02	-.74	.08	F	-.02	-.69	.07	-.05	.11	.46	394	.6460	.1799	.0711	-.08	38	4	-Te entra sueño?
M	.06	-.57	.08	F	-.05	-.42	.07	-.14	.11	-1.27	398	.2038	1.2876	.2565	-.22	38	5	-Pierdes la concentración?
M	.04	-.15	.09	F	-.03	-.04	.08	-.10	.12	-.87	398	.8827	.3049	.5809	-.11	38	6	-Te cuesta recordar lo que has leído?
M	.03	.96	.12	F	-.02	1.09	.10	-.13	.16	-.82	396	.4106	.8439	.3583	-.22	38	7	-Tea doble?
M	.04	.84	.12	F	-.03	1.00	.10	-.16	.15	-1.06	400	.8991	2.0018	.1571	-.36	38	8	-Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
M	.14	.21	.10	F	-.11	.68	.09	-.47	.13	-3.50	413	.0005	10.3824	.0013	-.72	38	9	-Te parece que lees lento?
M	-.06	-.05	.09	F	.05	-.20	.08	.15	.12	1.29	389	.1972	1.0508	.3083	.22	38	10	-Te duelen los ojos?
M	-.02	.04	.09	F	.01	-.01	.08	.05	.12	.40	382	.6928	.0020	.9640	.01	38	11	-Se te irritan los ojos?
M	-.01	.27	.10	F	.01	.27	.08	.00	.13	.00	387	1.000	.0011	.9732	.01	38	12	-Tienes sensación de "vibrar" alrededor de los ojos?
M	-.01	.03	.09	F	.01	.03	.08	.00	.12	.00	385	1.000	.0163	.8982	.02	38	13	-Horas que las palabras se ponen borrosas o que se enfocan y desenfocan?
M	.10	-.22	.09	F	-.08	.05	.08	-.28	.12	2.35	402	.0194	2.7262	.0987	-.35	38	14	-Te pierdes de línea al leer?
M	-.02	.12	.10	F	.02	.06	.08	.06	.12	-.49	386	.6268	.1080	.7424	.07	38	15	-Tienes que releer la misma línea de texto?

Width of Mantel slice: MHSlice = .010 leontes

FIGURA 5.51 – ANÁLISIS DEL DIF EN CUANTO AL SEXO MEDIANTE EL MÉTODO DE MANTEL (WINSTEPS)

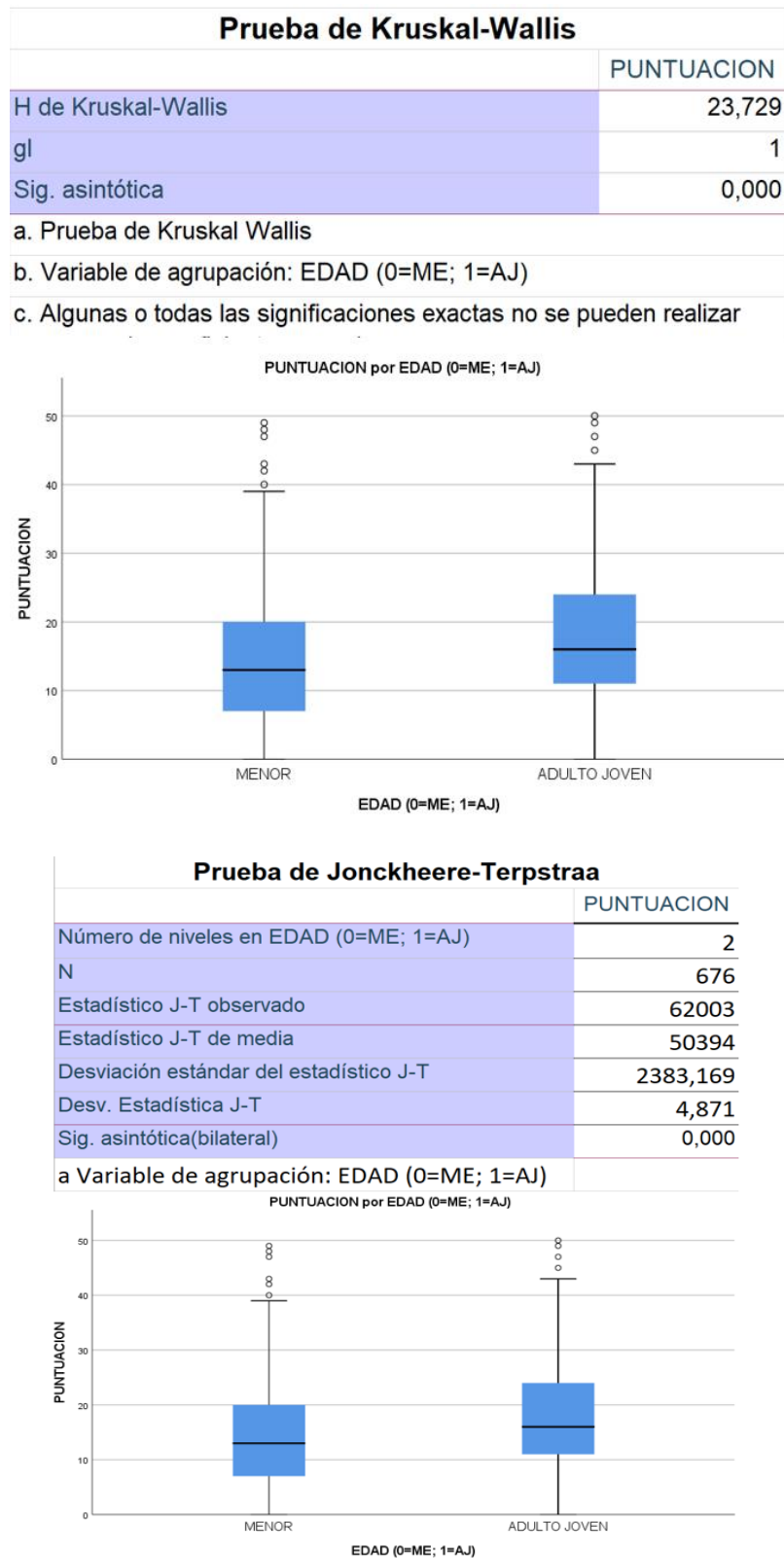
DIF class/group specification is: DIF=539M2

Person Obs-Exp	DIF	DIF	Person Obs-Exp	DIF	DIF	DIF	JOINT	Reach-Weich	Mantel	Size	Active Item							
CLASS/	Average	MEASURE	S.E.	CLASS/	Average	MEASURE	S.E.	CONTRAST	S.E.	t	d.f.	Prob.	Chi-sq	Prob.	CUMULOR	Slices	Number	Name
AT	.14	-.74	.08	ME	-.10	-.45	.07	-.29	.11	-2.65	409	.0083	12.0095	.0005	-.76	35	1	-Notas tus ojos cansados?
AT	.18	-.36	.08	ME	-.13	-.07	.08	-.42	.12	-3.66	419	.0003	16.5793	.0000	-.91	35	2	-Notas incomodidad en tus ojos?
AT	-.01	-.49	.08	ME	.01	-.49	.07	.00	.11	.00	405	1.0000	.3515	.5532	-.12	35	3	-Te duele la cabeza?
AT	-.03	-.67	.08	ME	.02	-.72	.07	.05	.11	.50	401	.6182	1.4155	.2341	-.24	35	4	-Te entra sueño?
AT	-.04	-.44	.08	ME	.03	-.52	.07	.08	.11	.78	404	.4384	.5338	.4650	-.14	35	5	-Pierdes la concentración?
AT	.04	-.15	.08	ME	-.03	-.06	.08	-.09	.12	-.77	412	.4407	.8692	.3512	-.19	35	6	-Te cuesta recordar lo que has leído?
AT	-.07	1.11	.11	ME	.06	.81	.10	.21	.15	2.02	399	.0442	2.2594	.1328	.41	35	7	-Ves doble?
AT	.03	-.83	.10	ME	-.02	-.92	.11	-.09	.15	-.63	419	.5290	.0127	.9103	-.03	35	8	-Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
AT	-.07	.56	.10	ME	.05	.34	.09	.22	.13	1.66	405	.0981	2.4294	.1191	.37	35	9	-Te parece que lees lento?
AT	.01	-.09	.08	ME	.00	-.09	.08	.00	.12	.00	409	1.0000	.0233	.8851	.04	35	10	-Te duelen los ojos?
AT	.14	-.17	.08	ME	-.10	.19	.09	-.36	.12	-3.03	420	.0026	5.7655	.0183	-.52	35	11	-Se te irritan los ojos?
AT	.02	.22	.09	ME	-.01	.27	.09	-.05	.12	-.40	414	.6930	.0295	.8826	-.04	35	12	-Tienes sensación de "vibrar" alrededor de los ojos?
AT	-.05	.09	.09	ME	.04	-.03	.08	.12	.12	1.01	406	.3148	1.5525	.2128	-.27	35	13	-Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
AT	-.32	.36	.09	ME	.24	-.44	.08	.80	.12	6.75	381	.0000	34.6324	.0000	1.37	35	14	-Te pierdes de línea al leer?
AT	.05	.04	.09	ME	-.03	.16	.08	-.12	.12	-1.00	414	.3156	2.6296	.1049	-.34	35	15	-Tienes que releer la misma línea de texto?
ME	-.10	-.45	.07	AT	.14	-.74	.08	.29	.11	2.65	409	.0083	12.0095	.0005	.76	35	1	-Notas tus ojos cansados?
ME	-.13	-.07	.08	AT	.18	-.36	.08	.42	.12	3.66	419	.0003	16.5793	.0000	.91	35	2	-Notas incomodidad en tus ojos?
ME	.01	-.49	.07	AT	-.01	-.49	.08	.00	.11	.00	405	1.0000	.3515	.5532	-.12	35	3	-Te duele la cabeza?
ME	.02	-.72	.07	AT	-.03	-.67	.08	-.05	.11	.50	401	.6182	1.4155	.2341	-.24	35	4	-Te entra sueño?
ME	.03	-.52	.07	AT	-.04	-.44	.08	-.08	.11	.78	404	.4384	.5338	.4650	-.14	35	5	-Pierdes la concentración?
ME	-.03	-.06	.08	AT	.04	-.15	.08	.09	.12	.77	412	.4407	.8692	.3512	-.19	35	6	-Te cuesta recordar lo que has leído?
ME	.06	.81	.10	AT	-.07	1.11	.11	-.21	.15	-2.02	399	.0442	2.2594	.1328	-.41	35	7	-Ves doble?
ME	-.02	.92	.11	AT	.03	.83	.10	.09	.15	.63	419	.5290	.0127	.9103	.03	35	8	-Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
ME	.05	.34	.09	AT	-.07	-.56	.10	-.22	.13	-1.66	405	.0981	2.4294	.1191	-.37	35	9	-Te parece que lees lento?
ME	.00	-.09	.08	AT	.01	-.09	.08	.00	.12	.00	409	1.0000	.0233	.8851	-.04	35	10	-Te duelen los ojos?
ME	-.10	.19	.09	AT	.14	-.17	.08	.36	.12	3.03	420	.0026	5.7655	.0183	.52	35	11	-Se te irritan los ojos?
ME	-.01	.27	.09	AT	.02	.22	.09	.05	.12	.40	414	.6930	.0295	.8826	.04	35	12	-Tienes sensación de "vibrar" alrededor de los ojos?
ME	.04	-.03	.08	AT	-.05	-.09	.09	-.12	.12	-1.01	406	.3148	1.5525	.2128	-.27	35	13	-Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
ME	.24	-.44	.08	AT	.24	-.44	.08	-.80	.12	-6.75	381	.0000	34.6324	.0000	1.37	35	14	-Te pierdes de línea al leer?
ME	-.03	.16	.08	AT	.05	.04	.09	.12	.12	1.00	414	.3156	2.6296	.1049	-.34	35	15	-Tienes que releer la misma línea de texto?

Width of Mantel slice: MHSLICE = .010 logits

FIGURA 5.52 – ANÁLISIS DEL DIF EN CUANTO A LA EDAD MEDIANTE EL MÉTODO DE MANTEL (WINSTEPS)

Finalmente se realizaron las pruebas de Kruskal-Wallis y la de Jonckheere-Terpstra, que una vez más nos confirmaron la diferencia significativa existente entre los resultados de las puntuaciones del grupo de niños de 9 a 17 y la obtenida por los adultos jóvenes de 18 a 30.



**FIGURA 5.5.53 – PRUEBAS NO PARAMÉTRICAS QUE CONFIRMAN LAS DIFERENCIAS ENTRE LAS PUNTUACIONES DE LOS SUJETOS SEGÚN EL RANGO DE EDAD AL QUE PERTENEZCAN (9-17 O 18-30 AÑOS)**

5.2.3.3.7 ANÁLISIS FACTORIAL

Se realizó un análisis factorial utilizando el análisis de las componentes principales (PCA) con una rotación Varimax con normalización de Kaiser sobre los datos recopilados en los 438 participantes que respondieron todos los ítems de CISSve.

Tanto la prueba de esfericidad de Bartlett como el examen de la medida de adecuación muestral de Kaiser-Meyer Olkin ( $KMO = .911$ ,  $p < 0.0001$ ) sugirieron que la muestra era perfectamente factorizable como se puede apreciar en la Figura 5.54.

**Prueba de KMO y Bartlett**

Medida Kaiser-Meyer-Olkin de adecuación de muestreo		,911
Prueba de esfericidad de Bartlett	Aprox. Chi-cuadrado	2635,091
	gl	105
	Sig.	,000

**FIGURA 5.54 – PRUEBA KMO Y PRUEBA DE ESFERICIDAD DE BARTLETT QUE CONFIRMAN LA VIABILIDAD PARA REALIZAR UN ANÁLISIS FACTORIAL**

Para realizar el análisis de las componentes principales se eligieron las componentes con autovalores mayores que 1. Con el objeto de maximizar el peso de los ítems más relevantes y minimizar el de los menos relevantes, se realizó una rotación Varimax.

Como se puede observar en la Figura 5.55, se obtuvo que el total de la varianza explicada por los 3 factores encontrados ascendía a un 58,2%, representando el primer factor un 24,70% de la misma, el segundo factor un 17,15% y el tercero un 16,36%.

**Varianza total explicada**

Componente	Autovalores iniciales			Sumas de cargas al cuadrado de la rotación		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	6,257	41,712	41,712	3,704	24,697	24,697
2	1,359	9,060	50,772	2,572	17,147	41,844
3	1,115	7,434	58,207	2,454	16,363	58,207
4	,915	6,101	64,308			
5	,750	5,003	69,311			
6	,652	4,345	73,656			
7	,611	4,077	77,732			
8	,553	3,685	81,417			
9	,519	3,459	84,876			
10	,488	3,252	88,128			
11	,417	2,782	90,910			
12	,383	2,551	93,462			
13	,366	2,438	95,900			
14	,314	2,095	97,995			
15	,301	2,005	100,000			

Método de extracción: análisis de componentes principales.

**FIGURA 5.55 – VARIANZA EXPLICADA POR LOS 3 FACTORES ENCONTRADOS EN EL ANÁLISIS FACTORIAL**

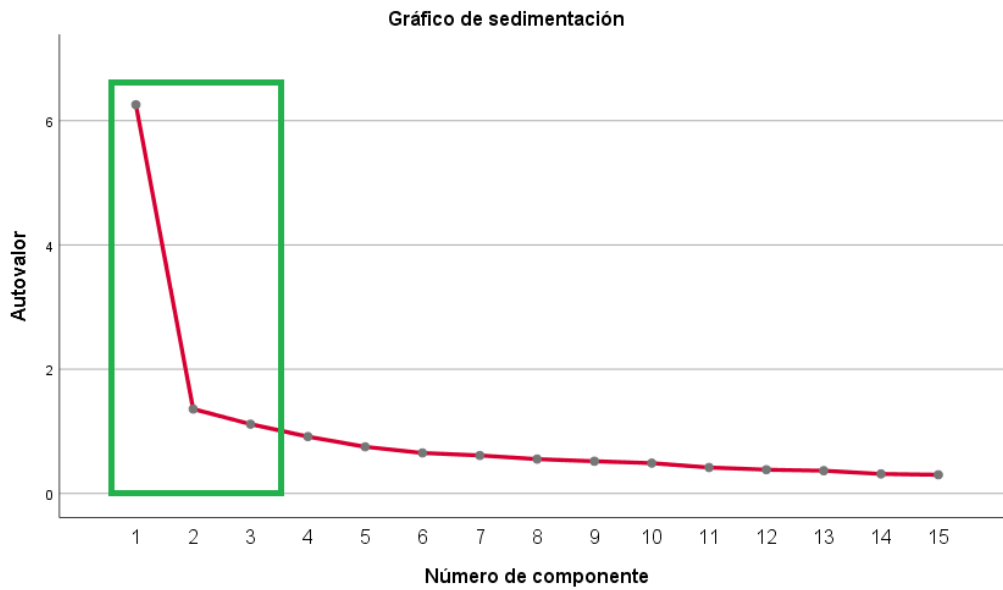
En la Figura 5.56 se puede apreciar el detalle sobre los factores obtenidos a través de la tabla que muestra la matriz de componentes. En ella, se puede visualizar la contribución de cada ítem a cada uno de los factores. A la derecha se muestra una interpretación que aporta valor semántico a cada factor. Observando la naturaleza del conjunto de ítems de cada grupo, se puede inducir cuál es la variable latente que tienen en común y que origina los síntomas de cada tipología.

Observamos como los ítems del factor 1 (ítem 1, 2, 3, 10, 11 y 12) comparten una expresión de síntomas que indican molestias asociadas con la fatiga visual derivada del trabajo prolongado de cerca. Los ítems del factor 2 (7, 8, 9, 13 y 14) muestran síntomas derivados de un fallo de la función visual, entendiéndose por ello que pueda ser un problema binocular, acomodativo o de motilidad ocular y por último, los ítems del factor 3 (4, 5, 6 y 15) denotan exteriorizaciones sintomáticas que indican pérdida de atención o dificultad para concentrarse en las tareas de lectura o trabajo de cerca.



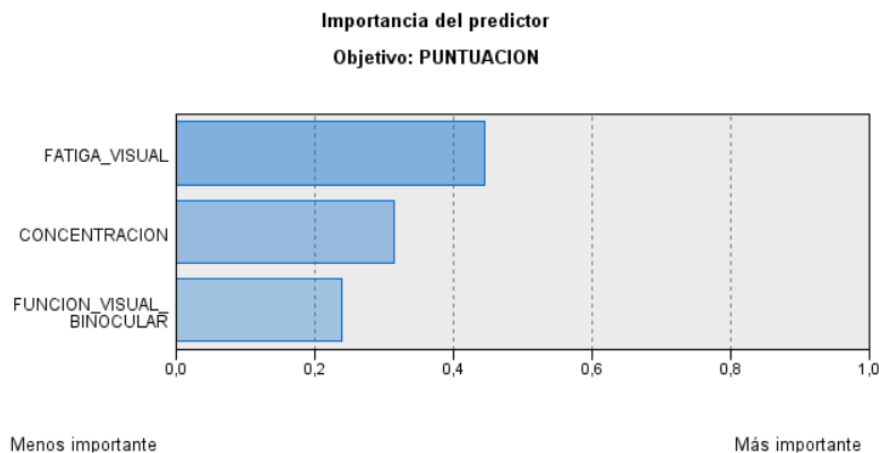
FIGURA 5.56 – FACTORES IDENTIFICADOS A TRAVÉS DEL ANÁLISIS FACTORIAL

En la Figura 5.57 se puede ver la representación del gráfico de sedimentación del análisis factorial donde se observan recuadrados en verde los 3 factores con autovalor > 1.



**FIGURA 5.57– DIAGRAMA DE SEDIMENTACIÓN**

Si realizamos una regresión múltiple considerando la puntuación total como variable dependiente, y los distintos factores como variables independientes para valorar el peso de cada uno de los factores sobre la puntuación total ( $\text{Puntuación} = \beta_0 + \text{Fatiga V.} \cdot \beta_1 + \text{Concentración} \cdot \beta_2 + \text{Función V.} \cdot \beta_3$ ), observamos, tal y como se puede observar en la tabla de la Figura 5.59, que los síntomas relacionados con las molestias asociadas a la fatiga visual en trabajos prolongados de cerca explican un 44,6% del total. Los síntomas asociados a fallos de la visión binocular, por su parte explican el 23.8% y, por último, los que se relacionan con la falta de concentración explican el 31,5% restante. En la Figura 5.58 se puede ver de manera gráfica el peso de cada uno de los factores.



**FIGURA 5.58 – REPRESENTACIÓN GRÁFICA DEL PESO DE CADA FACTOR**

## RESULTADOS

Por último, apoyándonos en el software estadístico SPSS v25.00 <sup>137</sup> realizamos una separación de la muestra por grupos de edad con el objeto de comparar si los síntomas de cada tipología mantienen el mismo peso en niños que en adultos.

El estudio del funcionamiento diferencial (DIF) por el método de Mantel- Haenszel (apartado 5.3.3.3.6), había marcado diferencias entre estos dos grupos en cuanto al funcionamiento de los ítems 1, 2, y 14, todos ellos pertenecientes al clúster que agrupa los síntomas derivados de la fatiga visual tras un trabajo prolongado de cerca. Al realizar las correspondientes regresiones, se confirmaron las diferencias entre ambos grupos de edad.

Observando la Figura 5.59, se puede ver cómo los adultos jóvenes con edades entre 18 y 30 años, presentaron un mayor porcentaje de la sintomatología percibida, explicada en base a síntomas relacionados con la fatiga tras trabajos prolongados de cerca, que los niños de entre 9 y 17 años. Por su parte, los niños explicaron un porcentaje mayor de sus síntomas a través del factor que relaciona los síntomas asociados a la dificultad de concentración para tareas asociadas con la lectura o el trabajo de cerca. En la Figura 5.60 se puede observar gráficamente este efecto.

FACTOR	NIÑOS	ADULTOS	TODOS
FATIGA VISUAL	39,20%	49,50%	44,60%
PROBLEMAS VISION BINOCULAR	23,20%	25,90%	23,80%
CAPACIDAD DE CONCENTRACION	37,70%	24,60%	31,50%

FIGURA 5.59 – CONTRIBUCIÓN DE CADA FACTOR AL TOTAL DE SÍNTOMAS PERCIBIDOS (POR GRUPOS DE EDAD)

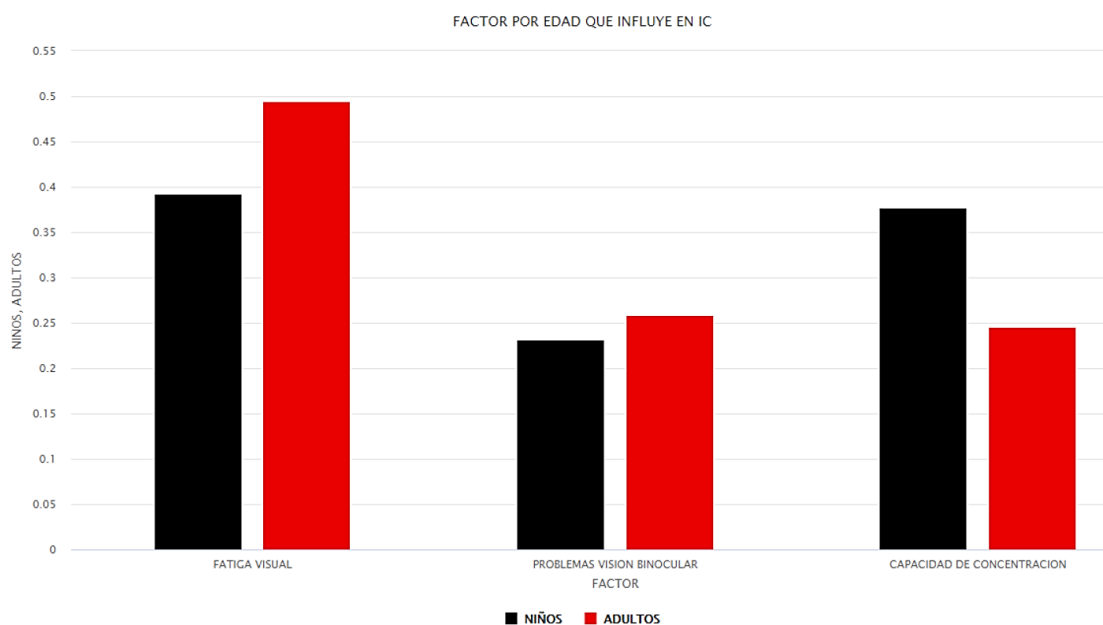


FIGURA 5.60 – REPRESENTACIÓN GRÁFICA DE LA CONTRIBUCIÓN DE CADA FACTOR AL TOTAL DE SÍNTOMAS PERCIBIDOS (POR GRUPOS DE EDAD)

5.2.3.3.8 VALIDEZ CONVERGENTE Y VALIDEZ DIVERGENTE

Se comenzó realizando la prueba de normalidad de Kolmogórov-Smirnov (K-S) para comprobar si las puntuaciones del CISSve, CVSS17 y WEMWBS se ajustaban a una distribución normal. En todos los casos se pudo comprobar que no fue así, con lo que, en lugar de utilizar la correlación de Pearson, hubo que recurrir a la estadística no paramétrica, utilizando el coeficiente “rho de Spearman” para evaluar la correlación. En la Figura 5.61 se pueden ver los resultados de esta prueba.

		PUNTUACIÓ N CVSS17 (17 ITEMS)	PUNTUACIÓ N CISSVE (15 ITEMS)	PUNTUACIÓ N WEMWBS (14 ITEMS)
N		252	252	252
Parámetros normales <sup>a,b</sup>	Media	3,045883	2,140211640	3,753968254
	Desv. Desviación	,3536235	,6321379291	,5775536524
Máximas diferencias extremas	Absoluto	,116	,077	,094
	Positivo	,088	,077	,035
	Negativo	-,116	-,038	-,094
Estadístico de prueba		,116	,077	,094
Sig. asintótica(bilateral)		,000 <sup>c</sup>	,001 <sup>c</sup>	,000 <sup>c</sup>
Significación exacta (bilateral)		,002	,098	,021
Probabilidad en el punto		,000	,000	,000

FIGURA 5.61 – PRUEBA DE NORMALIDAD DE K-S SOBRE LAS PUNTUACIONES DE LAS ESCALAS CISSVE, CVSS17 Y WEMWBS

Al realizar el test de Friedman (prueba no paramétrica alternativa al ANOVA de medidas repetidas)<sup>173</sup> para comparar las medias de las puntuaciones de las diferentes escalas se pudo observar que eran diferentes, tal y como constata el p-valor de 0.000 (con el truncamiento a 3 cifras decimales) mostrado en la Figura 5.62.

	Hipótesis nula	Prueba	Sig.	Decisión
1	Las distribuciones de PUNTUACIÓN CVSS17 (17 ITEMS), PUNTUACIÓN CISSVE (15 ITEMS) and PUNTUACIÓN WEMWBS (14 ITEMS) son las mismas.	Análisis de varianza de dos vías por rangos de Friedman para muestras relacionadas	,000	Rechazar la hipótesis nula.

Se muestran significaciones asintóticas. El nivel de significación es de ,05.

FIGURA 5.62– TEST DE “FRIEDMAN” PARA LA COMPARACIÓN DE LAS PUNTUACIONES DE LAS ESCALAS CISSVE, CVSS17 Y WEMWBS

Se analizó si cada una de las escalas per sé presentaba diferencias significativas entre los diferentes grupos muestrales que colaboraron completándolas. En todos los casos, la prueba “U de Mann-Whitney” desmintió esta hipótesis, encontrándose que los resultados, independientemente del lugar donde se hubieran completado, eran consistentes. Se pueden observar los resultados en la Figura 5.63.

**Resumen de prueba de hipótesis**

	Hipótesis nula	Prueba	Sig.	Decisión
1	La distribución de PUNTUACIÓN CVSS17 (16 ITEMS) es la misma entre las categorías de LUGAR.	Prueba U de Mann-Whitney para muestras independientes	,772	Retener la hipótesis nula.
2	La distribución de PUNTUACIÓN CISSVE (15 ITEMS) es la misma entre las categorías de LUGAR.	Prueba U de Mann-Whitney para muestras independientes	,152	Retener la hipótesis nula.
3	La distribución de PUNTUACIÓN WEMWBS (14 ITEMS) es la misma entre las categorías de LUGAR.	Prueba U de Mann-Whitney para muestras independientes	,573	Retener la hipótesis nula.

Se muestran significaciones asintóticas. El nivel de significación es de ,05.

**FIGURA 5.63 – PRUEBA “U DE MANN-WHITNEY” PARA LAS MUESTRAS INDEPENDIENTES DE LAS ESCALAS CISSVE, CVSS17 Y WEMWBS**

Por último, siguiendo con la estadística no paramétrica, mediante los coeficientes “rho de Spearman” y “Tau de Kendall” se analizaron las posibles relaciones entre la escala CISSve y la escala CVSS17 y entre la escala CISSve y la escala WEMWBS. El análisis condujo a la conclusión de que sí existía cierta correlación (“rho” =0.311 y “Tau” =0.220) entre los resultados de la puntuación de la escala CISSve y la escala CVSS17, lo que supone una prueba de validez convergente.

En el caso de la comparación entre las escalas CISSve y WEMWBS, se encontró que no existía correlación entre ambas (“rho” =-0,135 y “Tau” =-0.096), hallazgo lógico al hacer la comparación entre dos constructos que miden variables latentes de tan diferente naturaleza. Este resultado se estimó como una prueba de validez divergente.

En la Figura 5.64 se pueden observar los cálculos realizados mediante el software estadístico SPSS v25.00 <sup>137</sup>.

**Correlaciones<sup>a</sup>**

			PUNTUACIÓ N CVSS17	PUNTUACIO N CISSVE	PUNTUACIÓ N WEMWBS
<b>Tau_b de Kendall</b>	PUNTUACIÓN CVSS17	Coefficiente de correlación	1,000	,220**	,031
		Sig. (bilateral)	.	,000	,580
		N	162	162	162
	PUNTUACION CISSVE	Coefficiente de correlación	,220**	1,000	-,096
		Sig. (bilateral)	,000	.	,078
		N	162	162	162
	PUNTUACIÓN WEMWBS	Coefficiente de correlación	,031	-,096	1,000
		Sig. (bilateral)	,580	,078	.
		N	162	162	162
<b>Rho de Spearman</b>	PUNTUACIÓN CVSS17	Coefficiente de correlación	1,000	,311**	,045
		Sig. (bilateral)	.	,000	,567
		N	162	162	162
	PUNTUACION CISSVE	Coefficiente de correlación	,311**	1,000	-,135
		Sig. (bilateral)	,000	.	,086
		N	162	162	162
	PUNTUACIÓN WEMWBS	Coefficiente de correlación	,045	-,135	1,000
		Sig. (bilateral)	,567	,086	.
		N	162	162	162

\*\* La correlación es significativa en el nivel 0,01 (bilateral).

**FIGURA 5.64 – PRUEBAS DE CORRELACIÓN “RHO DE SPEARMAN” Y “TAU DE KENDALL” ENTRE LAS PUNTUACIONES DE LAS ESCALAS CISSVE, CVSS17 Y WEMWBS**

En la Figura 5.65 se puede observar gráficamente un diagrama de dispersión con la recta de regresión que relaciona las puntuaciones entre las escalas CISSve y CVSS17. Se observa cierta correlación positiva.

La Figura 5.66 muestra gráficamente un diagrama de dispersión con la recta de regresión que relaciona las puntuaciones entre las escalas CISSve y WEMWBS. Claramente se ve que no existe correlación, incluso esta es ligeramente negativa como se apreciaba en su expresión numérica.

RESULTADOS

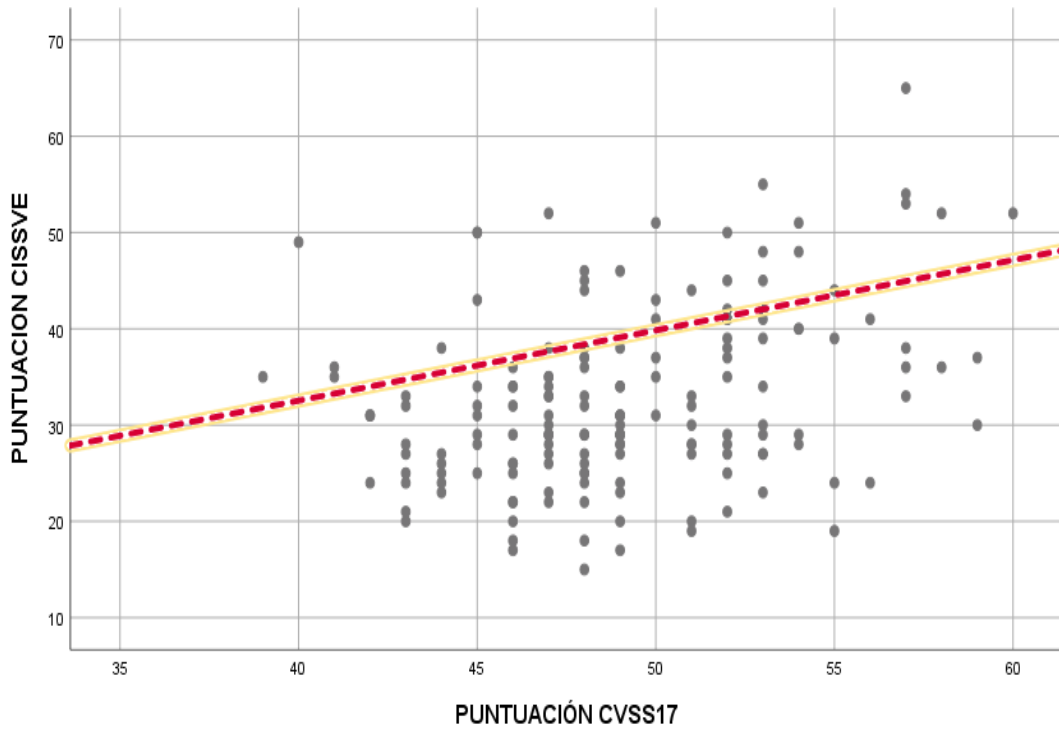


FIGURA 5.65 – DIAGRAMA DE DISPERSIÓN CON RECTA DE REGRESIÓN ENTRE LAS PUNTUACIONES DE LAS ESCALAS CISSVE Y CVSS17

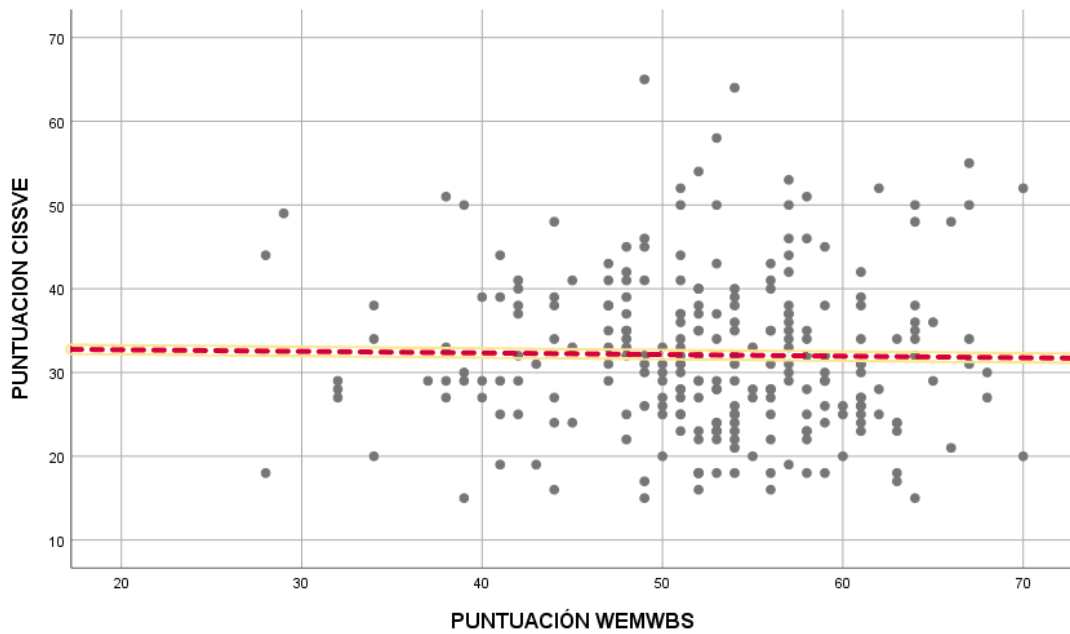


FIGURA 5.66 – DIAGRAMA DE DISPERSIÓN CON RECTA DE REGRESIÓN ENTRE LAS PUNTUACIONES DE LAS ESCALAS CISSVE Y WEMWBS

## 5.2.3.3.9 REPETIBILIDAD

Tras una primera administración de la escala CISSve, se volvió a pasar de nuevo la escala a todos los participantes excepto a los niños de educación primaria, dejando entre ambas sesiones un periodo de entre 7 y 15 días (Media 10.23. DT: 3.4 días).

Mediante el método de Bland-Altman, se calculó la media de la diferencia que fue -0.36 y los límites del intervalo de concordancia que incluye el 95% de las diferencias (1,96 veces la desviación típica de la diferencia) que fueron -10,27 y 9,55, obteniendo un COR de 9,91.

De acuerdo a Rouse et al. <sup>59,61</sup>, sería lícito descartar del análisis a aquellos sujetos con una diferencia entre las dos sesiones superior a dos desviaciones estándar, considerando a los sujetos “outliers” por expresar un comportamiento anómalo al presentar una desviación significativamente superior a la variabilidad normal esperada entre las puntuaciones de las dos sesiones. Descartando a 6 sujetos en los que la diferencia entre la puntuación de las dos sesiones era superior a 10 puntos, el COR mejoró a un valor de 9,14.

La Figura 5.67 muestra el diagrama de Bland-Altman. La línea roja marca la media de la diferencia entre ambas sesiones, y las dos líneas verdes los límites superior e inferior del intervalo de concordancia.

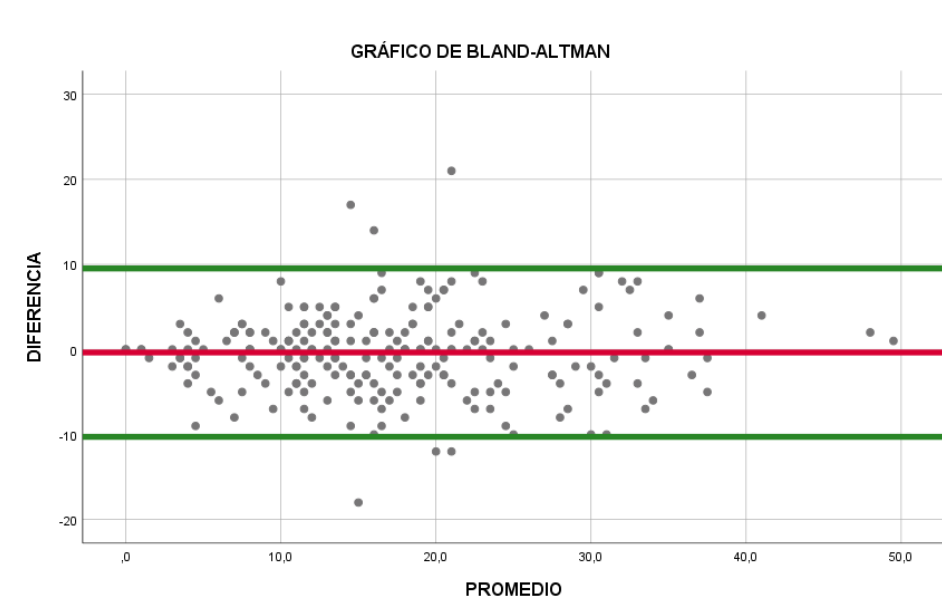
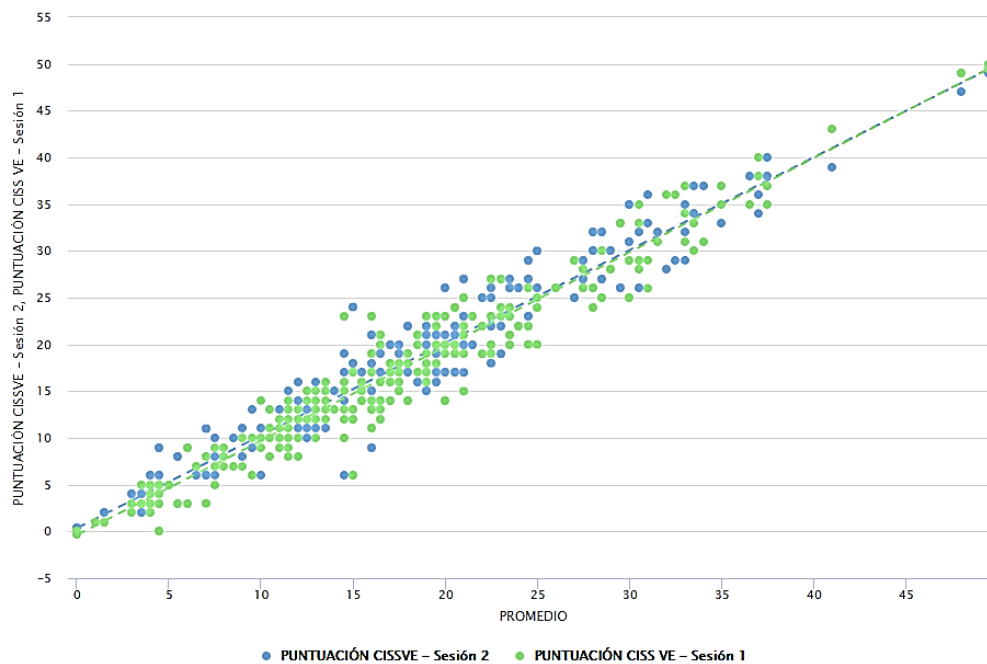


FIGURA 5.67 – DIAGRAMA DE BLAND-ALTMAN SOBRE LA REPETIBILIDAD DE LA ESCALA CISSVE

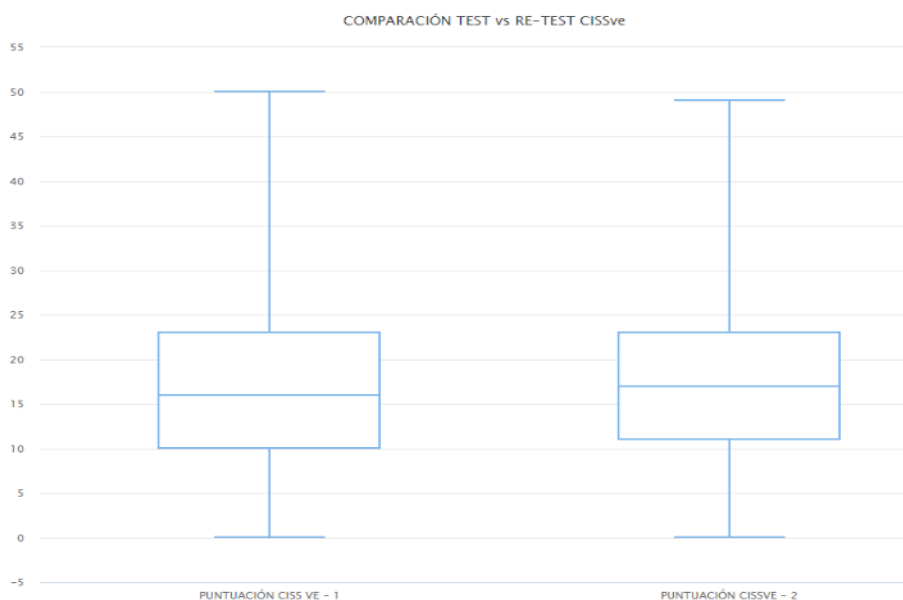
La Figura 5.68 muestra una representación del ajuste lineal entre la sesión 1 (test) y la sesión 2 (re-test), y da muestra gráfica de la alta correlación entre ambos eventos.

## RESULTADOS



**FIGURA 5.68– SESIÓN 1 VS SESIÓN 2 SOBRE LA MISMA MUESTRA - ESCALA CISSVE**

En la Figura 5.69, se puede ver a través de un diagrama de caja y bigote la representación gráfica de la igualdad entre las medias de ambas sesiones, que corrobora de nuevo la alta repetitividad de la prueba. Como se puede observar, el rango de las puntuaciones recogidas en ambas sesiones es muy similar, al igual que la media y la desviación típica de cada una.



**FIGURA 5.69 – COMPARACIÓN ENTRE LA SESIÓN 1 Y LA SESIÓN 2 - ESCALA CISSVE**

También se llevó a cabo el cálculo del coeficiente de correlación interclase. Para ello se utilizó el software estadístico SPSS v25.00 <sup>137</sup>, obteniendo un valor de 0.874, 95% de intervalo de confianza (0,840; 0,902), ratificando la consistencia entre las dos iteraciones.

La Figura 5.70 muestra el cálculo del coeficiente de correlación intraclase (ICC).

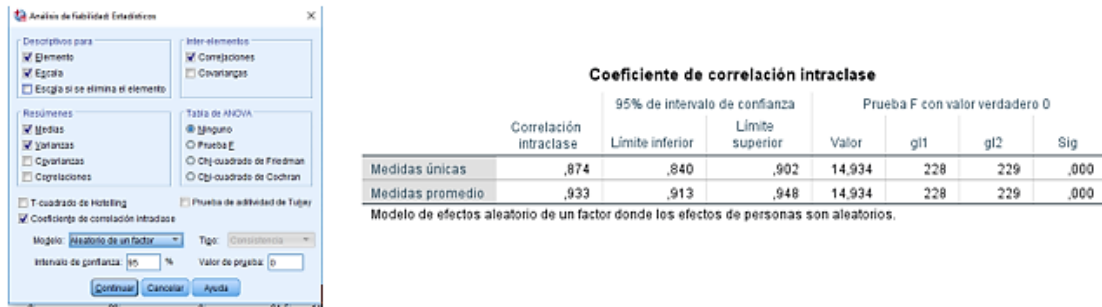


FIGURA 5.70 – COEFICIENTE DE CORRELACIÓN INTRACLASE - ESCALA CISSVe

### 5.2.3.4 APLICACIÓN DE LA ESCALA CISSVe A POBLACIÓN CLÍNICA

#### 5.2.3.4.1 COMPARACIÓN DE LAS RESPUESTAS AL CISSVe DE PACIENTES CON INSUFICIENCIA DE CONVERGENCIA DIAGNOSTICADOS CLÍNICAMENTE CON LOS DE UNA POBLACIÓN GENERAL

En este apartado se compararon las respuestas a la escala CISSVe en una población sintomática diagnosticada clínicamente (Figura 5.71) de acuerdo con los criterios descritos en el apartado de métodos (Apartado 4.2.3.4.1) frente a los de una muestra de población general escogida aleatoriamente.

RESULTADOS

Nº DE SUJETO	Foria VL (DP)	Foria VP (DP)	Rotura BT (DP)	Rotura PPC (cm)
1	8	25	30	6
2	0	12	14	6
3	2	10	19	13
4	6	18	10	7
5	2	8	14	6
6	6	16	30	7
7	0	16	6	6
8	4	14	8	SUPRIME
9	4	12	23	7
10	2	8	10	6
11	0	10	19	7
12	0	14	2	6
13	0	12	12	6
14	4	18	16	6
15	0	10	6	8
16	0	8	10	15
17	0	10	15	16
18	0	14	14	6
19	0	12	8	40
20	4	20	SUPRIME	SUPRIME
21	4	20	20	8
22	4	20	14	12
23	4	16	25	8
24	0	8	14	7
25	2	12	8	16
26	0	14	SUPRIME	17
27	0	16	30	8
28	8	25	12	6
29	0	10	15	6
30	0	8	8	8
31	0	8	14	7
32	0	14	26	6
33	4	18	20	7
34	8	14	SUPRIME	SUPRIME
35	4	18	16	6
36	2	10	12	6
37	4	20	14	6
38	0	12	6	7
39	4	18	12	6
40	0	18	20	12
41	0	8	10	8
42	2	12	20	6
43	2ENDO	10	18	8
44	0	6	12	6
45	4ENDO	20	12	12
46	0	10	19	7
47	4	10	19	6
48	4	20	18	7
49	0	14	6	14

VL = Visión Lejana
VP = Visión Próxima
DP = Dioptrías Prismáticas
BT = Base Temporal
PPC = Punto Próximo de Convergencia

FIGURA 5.71 – PACIENTES DIAGNOSTICADOS CLÍNICAMENTE CON IC

La muestra analizada, tal y como se puede observar en la Figura 5.72, contenía una submuestra sintomática (n=49, M=19,80, DT=9.13) y una no sintomática (n=52, M=14.02, DT=6.36)

Estadísticos descriptivos														
MUESTRA		N	Rango	Mínimo	Máximo	Suma	Media	Desv. Desviación	Varianza	Asimetría		Curtosis		
		Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Desv. Error	Estadístico	Desv. Error	
CLI	EDAD	49	21	9	30	905	18,47	,951	6,656	44,296	,372	,340	-1,146	,668
	Puntuación CISSve	49	39	8	47	970	19,80	1,305	9,138	83,499	1,378	,340	1,521	,668
	N válido (por lista)	49												
GEN	EDAD	52	21	9	30	895	17,21	,922	6,652	44,248	,611	,330	-,852	,650
	Puntuación CISSve	52	24	1	25	729	14,02	,882	6,363	40,490	-,133	,330	-,715	,650

FIGURA 5.72 – ESTADÍSTICOS DESCRIPTIVOS DE LA MUESTRA (SINTOMÁTICA/GENERAL)

Se analizaron las diferencias entre las puntuaciones medias de ambas submuestras para comprobar si eran estadísticamente significativas, realizando una prueba t de Student para muestras independientes. Antes, se realizó la prueba de Levene para comprobar la igualdad de las varianzas, resultando un p-valor de 0,095, mayor que 0,05 que permitió asumir que las varianzas eran iguales. La diferencia de medias entre las dos subpoblaciones (19.8-14.02) resultó ser 5,78. La prueba t, con un p-valor de 0,000 constató una diferencia estadísticamente significativa entre los dos grupos de la puntuación registrada en la escala CISSve y, por lo tanto, se pudo concluir que las diferencias observadas en las medias de las puntuaciones entre estos dos grupos no se debían al azar (véase Figura 5.73).

**Prueba T**

Estadísticas de grupo					
	MUESTRA	N	Media	Desv. Desviación	Desv. Error promedio
Puntuación CISSvelteración1	CLI	49	19,80	9,138	1,305
	GEN	52	14,02	6,363	,882

Prueba de muestras independientes										
	Prueba de Levene de igualdad de varianzas				prueba t para la igualdad de medias					
	F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia		
Puntuación CISSvelteración1	Se asumen varianzas iguales	2,844	,095	3,705	99	,000	5,777	1,559	2,683	8,871
		No se asumen varianzas iguales			3,666	85,154	,000	5,777	1,576	2,644

FIGURA 5.73 – PRUEBA DE LEVENE Y PRUEBA T PARA MUESTRAS INDEPENDIENTES

Como se muestra en la Figura 5.74, utilizando los factores encontrados durante el trabajo de validación de la escala CISSve, se analizó la influencia de cada uno de ellos en la puntuación total de la escala CISSve para cada una de las muestras.

**Estadísticas de grupo**

	MUESTRA	N	Media	Desv. Desviación	Desv. Error promedio
Puntuación CISSve	CLI	49	19,80	9,138	1,305
	GEN	52	14,02	6,363	,882
Trabajoprolongado	CLI	49	8,67	4,346	,621
	GEN	52	5,71	3,038	,421
Concentracion	CLI	49	4,43	4,243	,606
	GEN	52	4,29	2,484	,344
FuncionVisual	CLI	49	6,69	3,793	,542
	GEN	52	4,02	2,804	,389

FIGURA 5.74 - ESTADÍSTICOS DESCRIPTIVOS CISSVE – FACTORES POBLACIÓN CLÍNICA VS GENERAL

Por simple inspección ya se observaba cómo tanto los síntomas asociados a fallos en la función visual binocular como los referidos a la fatiga visual por trabajo prolongado de cerca se acentúan en la parte de la muestra diagnosticada clínicamente como sintomática, pero no es así en el caso de los asociados a falta de concentración, en los que se observan valores medios muy similares (4.43 vs 4.29)

Realizando la prueba de Levene para comprobar la homocedasticidad y la prueba t de igualdad de medias (véase Figura 5.75) para comprobar si las medias son iguales, corroboramos que efectivamente, existen diferencias significativas con respecto a los síntomas asociados a disfunción visual binocular y de la misma manera con los asociados a la fatiga visual causada por trabajo prolongado de cerca. Sin embargo, un p-valor de 0,841 constata de forma clara que no ocurre lo mismo con los síntomas asociados a la falta de concentración, los cuales no varían significativamente. Se muestran los resultados a continuación en la Figura 5.75.

		Prueba de Levene de igualdad de varianzas		prueba t para la igualdad de medias				95% de intervalo de confianza de la diferencia		
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	Inferior	Superior
Puntuación CISSve	Se asumen varianzas iguales	2,844	,095	3,705	99	,000	5,777	1,559	2,683	8,871
	No se asumen varianzas iguales			3,666	85,154	,000	5,777	1,576	2,644	8,909
Trabajoprolongado	Se asumen varianzas iguales	4,845	,030	3,988	99	,000	2,962	,743	1,488	4,435
	No se asumen varianzas iguales			3,947	85,343	,000	2,962	,750	1,470	4,454
Concentracion	Se asumen varianzas iguales	14,005	,000	,204	99	,839	,140	,687	-1,223	1,503
	No se asumen varianzas iguales			,201	76,504	,841	,140	,697	-1,248	1,528
FuncionVisual	Se asumen varianzas iguales	3,768	,055	4,046	99	,000	2,675	,661	1,363	3,986
	No se asumen varianzas iguales			4,011	88,167	,000	2,675	,667	1,349	4,000

FIGURA 5.75 - PRUEBA DE LEVENE Y T DE STUDENT - CLÍNICA VS GENERAL

Se estudió si existían diferencias significativas en cuanto al sexo y en cuanto a la edad en cada una de las muestras y se pudo comprobar que no era el caso, tal como se puede observar en la Figura 5.76 en la que se muestran los estadísticos descriptivos agrupados en cuanto al sexo y a continuación, en la Figura 5.77, la prueba de Levene y la prueba t que demuestra que no existe diferencia estadística entre las medias de los diferentes factores.

	Sexo	N	Media	Desv. Desviación	Desv. Error promedio
Puntuación CISSve	H	42	16,12	7,523	1,161
	M	59	17,32	8,866	1,154
Trabajoprolongado	H	42	6,45	4,145	,640
	M	59	7,64	3,850	,501
Concentracion	H	42	4,95	2,921	,451
	M	59	3,93	3,723	,485
FuncionVisual	H	42	4,71	3,359	,518
	M	59	5,75	3,675	,478

FIGURA 5.76 -ESTADÍSTICOS DESCRIPTIVOS CISSVE AGRUPADOS POR SEXO

		Prueba de Levene de igualdad de varianzas		prueba t para la igualdad de medias				95% de intervalo de confianza de la diferencia		
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	Inferior	Superior
Puntuación CISSve	Se asumen varianzas iguales	1,893	,172	-,715	99	,476	-1,203	1,683	-4,542	2,136
	No se asumen varianzas iguales			-,735	95,890	,464	-1,203	1,637	-4,452	2,046
Trabajoprolongado	Se asumen varianzas iguales	,217	,642	-1,485	99	,141	-1,192	,802	-2,784	,400
	No se asumen varianzas iguales			-1,467	84,332	,146	-1,192	,813	-2,807	,424
Concentracion	Se asumen varianzas iguales	1,884	,173	1,480	99	,142	1,020	,689	-,347	2,388
	No se asumen varianzas iguales			1,541	98,002	,126	1,020	,662	-,293	2,334
FuncionVisual	Se asumen varianzas iguales	1,271	,262	-1,440	99	,153	-1,031	,716	-2,452	,390
	No se asumen varianzas iguales			-1,462	92,937	,147	-1,031	,705	-2,432	,369

FIGURA 5.77 - PRUEBAS DE LEVENE Y T DE STUDENT – COMPARATIVA EN CUANTO AL SEXO

También se analizó si existía alguna diferencia entre los niños (9-17) y los adultos jóvenes (18-30), se pudieron observar los descriptivos que aparecen en la Figura 5.78 y a través de la prueba de Levene y prueba t (Figura 5.79) se pudo comprobar que no existían diferencias estadísticas significativas en cuanto a los síntomas referidos a disfunciones de la función visual y a la falta de concentración, pero sí se destacaba un mayor reporte de síntomas (3.05 puntos más en promedio) asociados a la fatiga derivada del trabajo prolongado de cerca en adultos jóvenes (18-30) que en niños (9-17).

## RESULTADOS

**Estadísticas de grupo**

	NIÑO0ADULTO1	N	Media	Desv. Desviación	Desv. Error promedio
Trabajoprolongado	0	63	6,00	3,750	,472
	1	38	9,05	3,698	,600
Concentracion	0	63	4,32	2,939	,370
	1	38	4,42	4,170	,676
FuncionVisual	0	63	5,05	3,391	,427
	1	38	5,76	3,845	,624

**FIGURA 5.78 - ESTADÍSTICOS DESCRIPTIVOS CISSVE AGRUPADOS POR GRUPOS DE EDAD**

**Prueba de muestras independientes**

		Prueba de Levene de igualdad de varianzas		prueba t para la igualdad de medias				95% de intervalo de confianza de la diferencia		
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	Inferior	Superior
Trabajoprolongado	Se asumen varianzas iguales	,065	,800	-3,984	99	,000	-3,053	,766	-4,573	-1,532
	No se asumen varianzas iguales			-3,998	79,007	,000	-3,053	,764	-4,573	-1,533
Concentracion	Se asumen varianzas iguales	8,344	,005	-,146	99	,884	-,104	,709	-1,510	1,303
	No se asumen varianzas iguales			-,134	59,324	,894	-,104	,771	-1,646	1,439
FuncionVisual	Se asumen varianzas iguales	,228	,634	-,977	99	,331	-,716	,733	-2,169	,738
	No se asumen varianzas iguales			-,947	70,593	,347	-,716	,756	-2,223	,792

**FIGURA 5.79 - PRUEBAS DE LEVENE Y T DE STUDENT – COMPARATIVA EN CUANTO A LA EDAD**

Comparando las medidas de cada ítem al aplicar el análisis de Rasch con el software Winsteps para cada tipo de población (sintomática diagnosticada clínicamente vs general), podemos ver que las mayores diferencias se observan en los ítems que se refieren a síntomas asociados a las disfunciones de la función visual, como ya habíamos comprobado anteriormente y como cabía esperar.

Véase la tabla de la Figura 5.80 que muestra ítem a ítem la diferencia en la medida para cada tipo de población, es decir, lo difícil que resulta cada ítem en cada tipo de población. Se puede observar cómo los ítems 7, 8 y 13, relacionados con síntomas asociados a la disfunción visual binocular se presentan con una mayor relevancia en pacientes sintomáticos, mientras que los sujetos de una población general en promedio dan más importancia al ítem 4 por ejemplo “¿Te entra sueño?”, o al 5, ¿Pierdes la concentración?, asociados a la falta de concentración por trabajo prolongado de cerca.

DESCRIPCIÓN	ITEM	MEDIDA SINTOMÁTICOS	MEDIDA GENERAL	DIFERENCIA	ABS(DIFERENCIA)	COMENTARIO
-Ves doble?	7	-0,19	1,04	-1,23	1,23	DISFUNCIÓN VISUAL - Mayor incidencia en P.Sintomática
-Te entra sueño?	4	0,37	-0,71	1,08	1,08	CONCENTRACIÓN - Mayor incidencia en P.General
-Te parece que las palabras se mueven, se mezclan o flotan?	8	0,16	0,93	-0,77	0,77	DISFUNCIÓN VISUAL - Mayor incidencia en P.Sintomática
-Notas que las palabras se ponen borrosas o que se enfocan?	13	-0,55	0,03	-0,58	0,58	DISFUNCIÓN VISUAL - Mayor incidencia en P.Sintomática
-Pierdes la concentración?	5	0,09	-0,49	0,58	0,58	CONCENTRACIÓN - Mayor incidencia en P.General
-Notas tus ojos cansados?	1	-0,3	-0,65	0,35	0,35	FATIGA - Mayor incidencia en P.General
-Te cuesta recordar lo que has leído?	6	0,11	-0,09	0,2	0,2	CONCENTRACIÓN - Mayor incidencia en P.General
-Te parece que lees lento?	9	0,64	0,48	0,16	0,16	CONCENTRACIÓN - Mayor incidencia en P.General
-Se te irritan los ojos?	11	-0,12	0,01	-0,13	0,13	FATIGA - Mayor incidencia en P.Sintomática
-Tienes que releer la misma línea de texto?	15	0,2	0,09	0,11	0,11	Valores similares
-Notas incomodidad en tus ojos?	2	-0,08	-0,18	0,1	0,1	Valores similares
-Tienes sensación de "tirantez" alrededor de los ojos?	12	0,36	0,27	0,09	0,09	Valores similares
-Te pierdes de línea al leer?	14	-0,01	-0,07	0,06	0,06	Valores similares
-Te duele la cabeza?	3	-0,56	-0,53	-0,03	0,03	Valores similares
-Te duelen los ojos?	10	-0,12	-0,14	0,02	0,02	Valores similares

FIGURA 5.80 - COMPARATIVA DE ÍTEMS - MEDIDA ANÁLISIS DE RASCH (CLÍNICA VS GENERAL)

5.2.3.4.2 APLICACIÓN DEL CISSve EN PACIENTES DIAGNOSTICADOS CLÍNICAMENTE DE INSUFICIENCIA DE CONVERGENCIA, ANTES Y DESPUÉS DE TERAPIA VISUAL

En este apartado se analizan las puntuaciones obtenidas en la escala CISSve por una población sintomática diagnosticada clínicamente (n=49, M=18.47, DT=6.65 años) de acuerdo a los criterios descritos en el apartado de métodos (Apartado 4.2.3.4.2), parte de la cual (n=27, M=16,22, DT=7,10 años) fue sometida posteriormente a terapia visual (Figura 5.81). De los 27 pacientes que se sometieron a terapia, 3 fueron excluidos por no completar el tratamiento, resultando finalmente un espacio muestral (n=24, M=17.29, DT=6.74 años).

Nº SUJETO	Foria VL (DP)	Foria VP (DP)	Rotura BT (DP)	Rotura PPC (cm)	Foria VL - f (DP)	Foria VP - f (DP)	Rotura BT - f (DP)	Rotura PPC - f (cm)
1	8	25	30	6	4	14	30	0
2	0	12	14	6	0	6	40	0
3	2	10	19	13	2	10	40	0
4	6	18	10	7	0	14	>40	0
5	6	16	30	7	2	12	>40	0
6	0	16	6	6	0	14	40	0
7	4	14	8	SUPRIME	4	12	>40	0
8	4	12	23	7	4	8	35	0
9	2	8	10	6	2	8	>40	0
10	0	10	19	7	0	8	35	0
11	0	14	2	6	0	14	30	0
12	0	12	12	6	0	12	35	0
13	4	18	16	6	4	16	>40	0
14	0	10	6	8	0	6	30	0
15	0	8	10	15	0	4	35	0
16	0	10	15	16	0	8	>40	0
17	0	14	14	6	0	10	30	0
18	0	12	8	40	0	6	40	4
19	4	20	SUPRIME	SUPRIME	0	18	40	0
20	4	20	20	8	0	16	40	0
21	4	20	14	12	4	14	40	3
22	4	16	25	8	4	14	35	6
23	2	12	8	16	2	12	40	0
24	0	16	30	8	0	12	>40	0

VL = Visión Lejana  
 VP = Visión Próxima  
 DP = Dioptías Prismáticas  
 BT = Base Temporal  
 PPC = Punto Próximo de Convergencia  
 f = Final tras tratamiento de Terapia Visual

FIGURA 5.81 – PACIENTES DIAGNOSTICADOS CLÍNICAMENTE DE IC – PRE Y POST TERAPIA

Presentamos a continuación en la Figura 5.82 y en la Figura 5.83, los estadísticos descriptivos de cada uno de los ítems y de la puntuación del CISSve en cada una de las dos iteraciones obtenidos con ayuda del paquete estadístico SPSS 25.0<sup>137</sup>

## RESULTADOS

Estadísticos descriptivos <sup>a</sup>									
	N	Rango	Media	Dev.	Varianza	Asimetría		Curtosis	
	Estadístico	Estadístico	Estadístico	Desviación Estadístico	Estadístico	Estadístico	Desv. Error	Estadístico	Desv. Error
1. ¿Notas tus ojos cansados?	24	3	1,67	1,007	1,014	-,640	,472	-,594	,918
2. ¿Notas incomodidad en tus ojos?	24	3	1,29	,999	,998	,491	,472	-,654	,918
3. ¿Te duele la cabeza?	24	4	1,63	1,377	1,897	-,012	,472	-1,603	,918
4. ¿Te entra sueño?	24	4	1,08	1,349	1,819	1,112	,472	-,003	,918
5. ¿Pierdes la concentración?	24	4	1,29	1,160	1,346	,647	,472	-,314	,918
6. ¿Te cuesta recordar lo que has leído?	24	4	1,42	1,442	2,080	,608	,472	-,963	,918
7. ¿Ves doble?	24	4	1,42	1,283	1,645	,337	,472	-1,162	,918
8. ¿Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?	24	4	1,17	1,239	1,536	,852	,472	-,390	,918
9. ¿Te parece que lees lento?	24	4	,75	1,073	1,152	1,697	,472	2,784	,918
10. ¿Te duelen los ojos?	24	4	1,50	1,103	1,217	,318	,472	-,368	,918
11. ¿Se te irritan los ojos?	24	4	1,54	1,285	1,650	,436	,472	-,826	,918
12. ¿Tienes sensación de "tirantez" alrededor de los ojos?	24	4	,96	1,268	1,607	1,062	,472	-,106	,918
13. ¿Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?	24	4	1,96	1,083	1,172	,089	,472	-,348	,918
14. ¿Te pierdes de línea al leer?	24	4	1,29	1,301	1,694	,437	,472	-1,128	,918
15. ¿Tienes que releer la misma línea de texto?	24	4	1,08	1,213	1,471	,785	,472	-,336	,918
PUNTUACIÓN CISSVe	24	41	20,04	9,928	98,563	1,041	,472	1,182	,918

FIGURA 5.82 - ESTADÍSTICOS DESCRIPTIVOS CISSVE – ANTES DE TERAPIA

Estadísticos descriptivos <sup>a</sup>									
	N	Rango	Media	Dev.	Varianza	Asimetría		Curtosis	
	Estadístico	Estadístico	Estadístico	Desviación Estadístico	Estadístico	Estadístico	Desv. Error	Estadístico	Desv. Error
1. ¿Notas tus ojos cansados?	24	1	,50	,511	,261	,000	,472	-2,190	,918
2. ¿Notas incomodidad en tus ojos?	23	2	,57	,590	,348	,454	,481	-,616	,935
3. ¿Te duele la cabeza?	24	2	,54	,721	,520	,981	,472	-,285	,918
4. ¿Te entra sueño?	24	2	,79	,884	,781	,443	,472	-1,615	,918
5. ¿Pierdes la concentración?	24	3	,88	,947	,897	,937	,472	,178	,918
6. ¿Te cuesta recordar lo que has leído?	24	3	1,04	1,233	1,520	,674	,472	-1,230	,918
7. ¿Ves doble?	24	2	,13	,448	,201	3,797	,472	14,650	,918
8. ¿Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?	24	4	,38	,924	,853	3,109	,472	10,563	,918
9. ¿Te parece que lees lento?	24	2	,42	,584	,341	1,067	,472	,295	,918
10. ¿Te duelen los ojos?	24	1	,29	,464	,216	,979	,472	-1,145	,918
11. ¿Se te irritan los ojos?	24	3	,38	,711	,505	2,463	,472	7,336	,918
12. ¿Tienes sensación de "tirantez" alrededor de los ojos?	24	3	,21	,658	,433	3,756	,472	15,108	,918
13. ¿Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?	24	1	,29	,464	,216	,979	,472	-1,145	,918
14. ¿Te pierdes de línea al leer?	24	4	,92	1,100	1,210	1,246	,472	1,288	,918
15. ¿Tienes que releer la misma línea de texto?	24	3	,83	1,007	1,014	1,197	,472	,585	,918
PUNTUACIÓN CISSVe	24	20	8,13	4,981	24,810	,640	,472	-,011	,918

FIGURA 5.83 - ESTADÍSTICOS DESCRIPTIVOS CISSVE – DESPUÉS DE TERAPIA

Resumimos la comparación entre las medias y desviaciones típicas de las puntuaciones obtenidas en las respuestas a cada ítem en cada una de las dos iteraciones (pre-post terapia),

en la Figura 5.84. Observamos cómo las mayores variaciones se producen en los ítems 7, 13, 10, 11 y 1.

ÍTEM	Media It 1	DT It 1	Media It 2	DT It 2	Dif. Medias	Dif. DTs	% Bajada Medias	% Bajada DT
7. ¿Ves doble?	1,3	1,265	0,08	0,392	-1,22	-0,873	-1525%	-69%
13. ¿Notas que las palabras se ponen borrosas o que se enfocan?	1,89	1,05	0,35	0,485	-1,54	-0,565	-440%	-54%
10. ¿Te duelen los ojos?	1,33	1,144	0,27	0,452	-1,06	-0,692	-393%	-60%
12. ¿Tienes sensación de "tirantez" alrededor de los ojos?	0,85	1,231	0,19	0,634	-0,66	-0,597	-347%	-48%
11. ¿Se te irritan los ojos?	1,44	1,281	0,42	0,703	-1,02	-0,578	-243%	-45%
1. ¿Notas tus ojos cansados?	1,59	1,01	0,5	0,51	-1,09	-0,5	-218%	-50%
8. ¿Te parece que las palabras se mueven, se mezclan o flotan sc	1,04	1,224	0,35	0,892	-0,69	-0,332	-197%	-27%
3. ¿Te duele la cabeza?	1,44	1,396	0,5	0,707	-0,94	-0,689	-188%	-49%
2. ¿Notas incomodidad en tus ojos?	1,33	0,961	0,58	0,578	-0,75	-0,383	-129%	-40%
9. ¿Te parece que lees lento?	0,96	1,255	0,62	0,752	-0,34	-0,503	-55%	-40%
5. ¿Pierdes la concentración?	1,44	1,219	0,96	0,958	-0,48	-0,261	-50%	-21%
4. ¿Te entra sueño?	1,07	1,299	0,77	0,863	-0,3	-0,436	-39%	-34%
14. ¿Te pierdes de línea al leer?	1,44	1,368	1,04	1,076	-0,4	-0,292	-38%	-21%
15. ¿Tienes que releer la misma línea de texto?	1,26	1,289	0,96	0,999	-0,3	-0,29	-31%	-22%
6. ¿Te cuesta recordar lo que has leído?	1,3	1,409	1	1,233	-0,3	-0,176	-30%	-12%

FIGURA 5.84 –COMPARACIÓN ESTADÍSTICOS DESCRIPTIVOS ANTES DE TERAPIA VS DESPUÉS DE TERAPIA

Realizando con el software SPSS v25 un análisis factorial para cada una de las dos iteraciones (pre y post terapia) con el objetivo de reducir el número de variables, encontramos los mismos factores que ya se describieron en el apartado 5.2.3.3.7. Por un lado, el que agrupaba los síntomas asociados a la fatiga por trabajo prolongado de cerca, por otro, el que describía las disfunciones visuales binoculares, y por último, el que se asociaba con los síntomas que reportaban problemas de concentración.

En la Figura 5.85, podemos observar los descriptivos de cada uno de los 3 factores en la primera iteración, antes de someterse a terapia, y en la segunda, después de someterse a terapia

ANTES DE TERAPIA

	N Estadístico	Rango Estadístico	Media Estadístico	Desv. Desviación Estadístico	Varianza Estadístico	Asimetría		Curtosis	
						Estadístico	Desv. Error	Estadístico	Desv. Error
FUNCIÓN VISUAL	24	11	4,54	3,107	9,650	,374	,472	-,891	,918
FATIGA	24	18	8,58	4,995	24,949	,108	,472	-,632	,918
CONCENTRACIÓN	24	18	6,92	5,733	32,862	,263	,472	-1,077	,918
N válido (por lista)	24								

DESPUÉS DE TERAPIA

	N Estadístico	Rango Estadístico	Media Estadístico	Desv. Desviación Estadístico	Varianza Estadístico	Asimetría		Curtosis	
						Estadístico	Desv. Error	Estadístico	Desv. Error
FUNCIÓN VISUAL	24	7	,79	1,560	2,433	3,077	,472	11,060	,918
FATIGA	24	6	2,46	2,085	4,346	,331	,472	-,978	,918
CONCENTRACIÓN	24	14	4,87	4,416	19,505	,596	,472	-,799	,918
N válido (por lista)	24								

FIGURA 5.85 – COMPARACIÓN ESTADÍSTICOS DESCRIPTIVOS DE CADA FACTOR, ANTES Y DESPUÉS DE TERAPIA

Por simple inspección se puede observar cómo los síntomas asociados a fallos en la función visual binocular disminuyen abruptamente entre la primera y la segunda aplicación del CISSve,

## RESULTADOS

después de la terapia. También se observa una bajada importante en los referidos a fatiga visual, pero no es así en el caso de los asociados a falta de concentración, en los que sólo se percibe una ligera disminución de estos.

Realizando la prueba de Levene para comprobar la homocedasticidad y la prueba t de igualdad de medias para comprobar si las medias son iguales corroboramos que efectivamente, existe una mejoría significativa con respecto a los síntomas asociados a disfunción visual binocular y de la misma manera con los asociados a la fatiga visual causada por trabajo prolongado de cerca. Sin embargo, un p-valor de 0,174 constata que no ocurre lo mismo con los síntomas asociados a la falta de concentración, los cuales no varían significativamente. Se muestran los resultados a continuación en la Figura 5.86

		Prueba de Levene de igualdad de varianzas		prueba t para la igualdad de medias					95% de intervalo de confianza de la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	Inferior	Superior
FUNCIÓN VISUAL	Se asumen varianzas iguales	16,581	,000	5,285	46	,000	3,750	,710	2,322	5,178
	No se asumen varianzas iguales			5,285	33,904	,000	3,750	,710	2,308	5,192
CONCENTRACIÓN	Se asumen varianzas iguales	1,651	,205	1,382	46	,174	2,042	1,477	-,932	5,015
	No se asumen varianzas iguales			1,382	43,190	,174	2,042	1,477	-,937	5,020
FATIGA	Se asumen varianzas iguales	13,192	,001	5,544	46	,000	6,125	1,105	3,901	8,349
	No se asumen varianzas iguales			5,544	30,777	,000	6,125	1,105	3,871	8,379

**FIGURA 5.86 - PRUEBA DE LEVENE Y T DE STUDENT FACTORES**

Podemos ver este resultado también de manera gráfica en la Figura 5.87

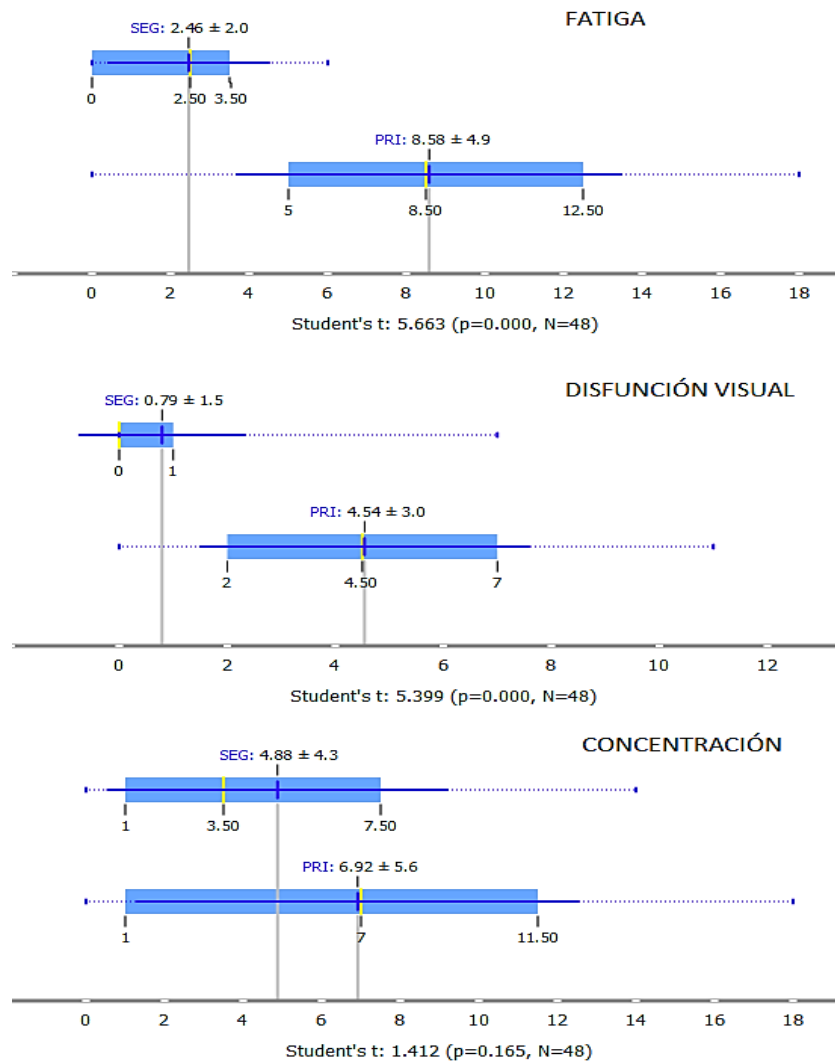


FIGURA 5.87 – DIAGRAMA DE CAJA Y BIGOTE – VARIACIÓN DE SINTOMATOLOGÍA PRE VS POST TERAPIA

Adicionalmente, tomando la muestra total de los sujetos diagnosticados como sintomáticos en la clínica y tras excluir a los 3 que no cumplían los criterios de edad se analizó si existía alguna diferencia de comportamiento del CISSve en cuanto al sexo con esa muestra (n=49, M=18.47, DT=6.65 años). Se pudo comprobar que no era el caso, como se puede observar en la Figura 5.88

RESULTADOS

		Prueba de Levene de igualdad de varianzas				prueba t para la igualdad de medias			95% de intervalo de confianza de la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	Inferior	Superior
FUNCIÓN VISUAL	Se asumen varianzas iguales	1,794	,186	-1,397	52	,168	-1,113	,797	-2,711	,486
	No se asumen varianzas iguales			-1,386	48,015	,172	-1,113	,803	-2,727	,502
CONCENTRACIÓN	Se asumen varianzas iguales	,066	,798	-,511	52	,611	-,731	1,429	-3,598	2,136
	No se asumen varianzas iguales			-,509	50,123	,613	-,731	1,435	-3,613	2,152
FATIGA	Se asumen varianzas iguales	,105	,747	-1,082	52	,284	-1,387	1,282	-3,961	1,186
	No se asumen varianzas iguales			-1,078	50,516	,286	-1,387	1,287	-3,972	1,197

FIGURA 5.88 – PRUEBA DE LEVENE Y T DE STUDENT - VARIACIÓN DE SÍNTOMAS EN CUANTO AL SEXO

Analizando si existía alguna diferencia entre los niños (9-17) y los adultos jóvenes (18-30), se pudieron observar los descriptivos que aparecen en la Figura 5.89

	EDAD	N	Media	Desv. Desviación	Desv. Error promedio
FUNCIÓN VISUAL	>= 18	14	3,00	3,637	,972
	< 18	40	2,33	2,702	,427
CONCENTRACIÓN	>= 18	14	5,93	6,673	1,784
	< 18	40	6,50	4,685	,741
FATIGA	>= 18	14	7,43	5,827	1,557
	< 18	40	4,43	4,063	,642

FIGURA 5.89 - DESCRIPTIVOS FACTORES PARA LOS 2 GRUPOS DE EDAD – NIÑOS (9-17) Y ADULTOS (18-30)

Pese a que en la tabla de la Figura 5.89 se reflejan ligeras diferencias entre los grupos de edad. Los niños reportan puntuaciones menores en cuanto a problemas de la función visual binocular y la fatiga y mayores puntuaciones en los ítems referidos a la falta de concentración, en la Figura 5.90 se puede observar cómo al realizar la prueba de Levene y la t de Student, cómo sí existen diferencias en cuanto a las desviaciones típicas en los síntomas asociados a fatiga y falta de concentración, pero que las medias no se pueden considerar estadísticamente diferentes para ninguno de los factores.

		Prueba de Levene de igualdad de varianzas				prueba t para la igualdad de medias			95% de intervalo de confianza de la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	Inferior	Superior
FUNCIÓN VISUAL	Se asumen varianzas iguales	2,922	,093	,733	52	,467	,675	,920	-1,172	2,522
	No se asumen varianzas iguales			,636	18,280	,533	,675	1,062	-1,553	2,903
CONCENTRACIÓN	Se asumen varianzas iguales	6,637	,013	-,350	52	,728	-,571	1,631	-3,845	2,702
	No se asumen varianzas iguales			-,296	17,696	,771	-,571	1,931	-4,634	3,491
FATIGA	Se asumen varianzas iguales	4,962	,030	2,117	52	,039	3,004	1,419	,157	5,850
	No se asumen varianzas iguales			1,783	17,630	,092	3,004	1,685	-,541	6,548

FIGURA 5.90 - PRUEBA DE LEVENE Y T DE STUDENT - VARIACIÓN DE SÍNTOMAS EN CUANTO A LA EDAD

Por último, para comprobar la eficacia de la terapia visual en los sujetos que se sometieron a la misma (n=24, M=17.29, DT=6.74 años), se realizó un análisis cluster con todas las respuestas al CISSve (incluyendo las pre y post terapia), y se pudo observar, tal y como se ve en la Figura 5.91 cómo se formaban 2 grupos. Uno (zona roja) que contenía las respuestas al CISSve que reportaban menos sintomatología, y otro en azul que agrupaba los cuestionarios con respuestas que constataban una mayor sintomatología. Se observa cómo en la zona roja aparecen prácticamente todas las respuestas en la segunda iteración, después de terapia, mientras que en la azul están la mayoría de los cuestionarios realizados antes de la terapia visual. Vemos como hay un paciente, el P51, que después de la terapia sigue contestando que nota sintomatología, y también se observan otros, como el P42 o el P02 que ya desde el primer momento no reportan sintomatología severa ni mejoran significativamente los síntomas percibidos después de la terapia.

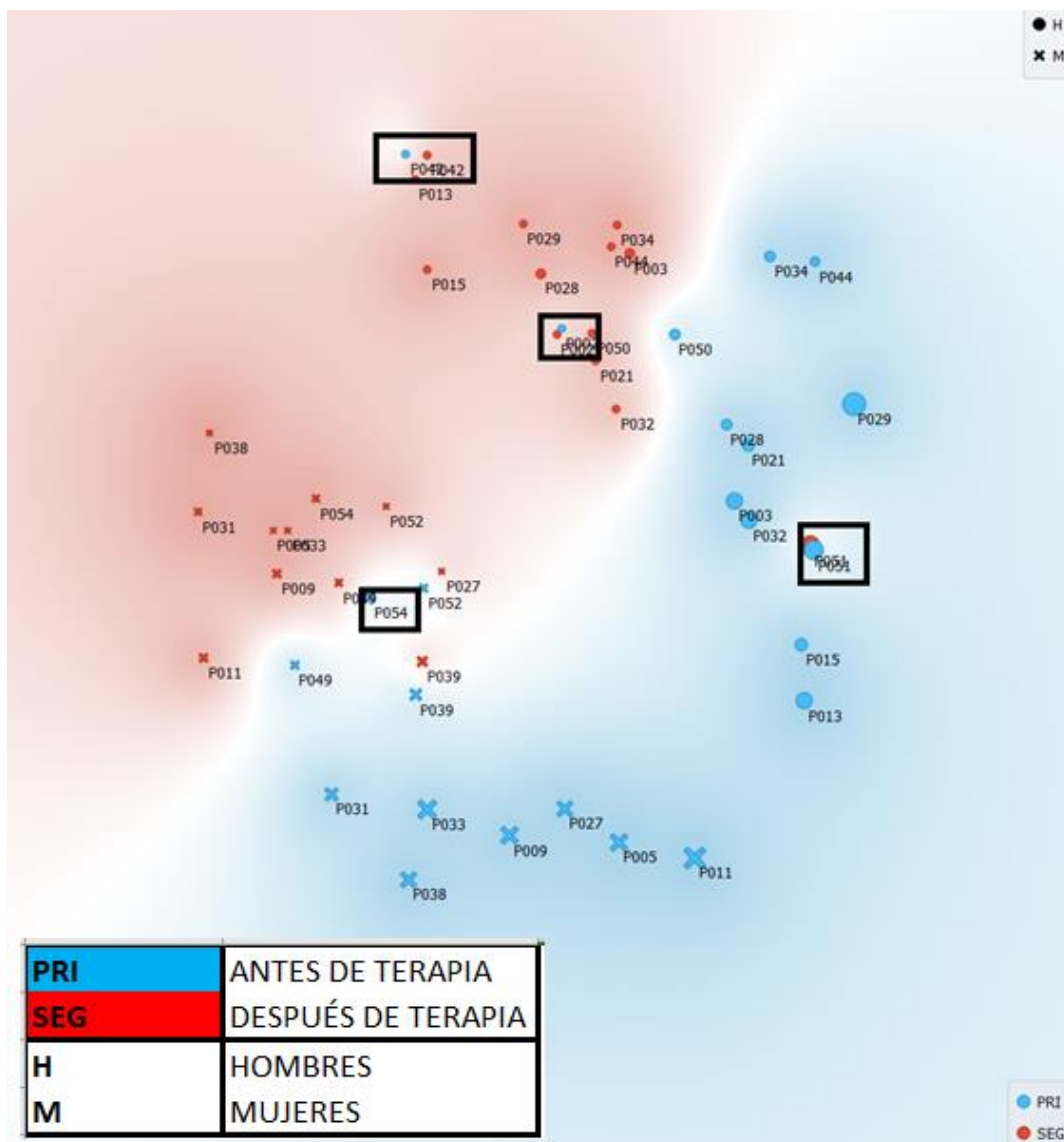


FIGURA 5.91 - ANÁLISIS CLUSTER – RESPUESTAS PRE Y POST TERAPIA

## RESULTADOS

### 5.2.3.4.3 COMPARACIÓN DE LA PREVALENCIA DE LAS DIFERENTES TIPOLOGÍAS DE SÍNTOMAS POR GRUPOS DE EDAD, INCLUYENDO UNA MUESTRA DE PACIENTES PRÉSBITAS

En este apartado se analizan las puntuaciones obtenidas en la escala CISSve de una pequeña muestra de sujetos presbíta (mayores de 40 años) que participaron en la investigación ( $n=29$ ,  $M=60,86$ ;  $DT=13,82$ ), y que no se habían considerado por estar fuera del rango de edad que cubre este estudio (9-30 años). Se observó el comportamiento de las puntuaciones de estos sujetos presbíta sobre la escala CISSve en cuanto a las diferentes tipologías de síntomas determinadas por los factores encontrados al realizar el análisis factorial (Apartado 5.2.3.3.7),

Agregando las puntuaciones obtenidas en todos los ítems correspondientes a cada factor para cada grupo de edad, y representando gráficamente los resultados en un diagrama de barras, se pudo observar cualitativamente la diferencia en prevalencia de cada tipo de síntomas para cada grupo de edad, (Figura 5.92).

En la Figura se aprecia cómo con la edad se acentúa aún más la polarización de los síntomas en los presbíta hacia los relacionados con la fatiga visual al realizar trabajo prolongado de cerca, y cómo por el contrario los síntomas asociados a la capacidad de concentración apenas adquieren relevancia en este grupo de edad. En cuanto a los síntomas asociados a los problemas en la visión binocular, no se observaron grandes variaciones entre los diferentes grupos.

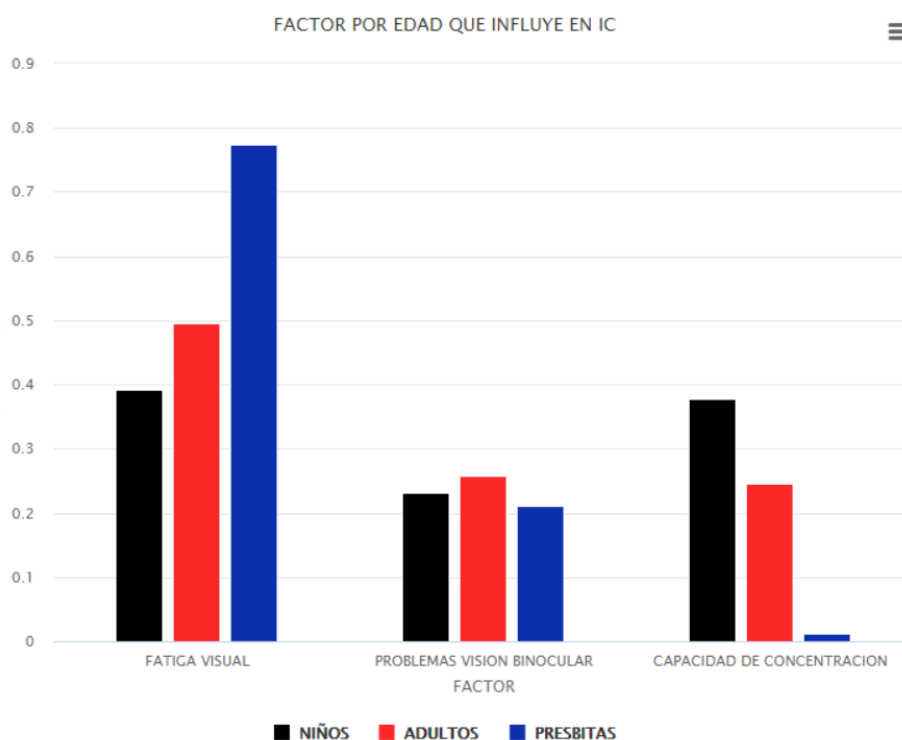


FIGURA 5.92 – TIPOLOGÍA DE SÍNTOMAS POR GRUPOS DE EDAD (9-17, 18-30, >40 AÑOS)

### 5.3 COMPARACIÓN ENTRE LAS PROPIEDADES DE CISSv15 Y CISSve

Siguiendo la secuencia explicada en el apartado 4.4 de la parte de métodos, en este apartado se realiza una comparación entre las propiedades psicométricas de la escala adaptada al español CISSve y las de la versión original del mismo CISSv15.

Para la evaluación de los resultados, presentamos la Figura 5.93 que resume el comportamiento de cada una de las escalas comparando el comportamiento de ambas frente a los criterios propuestos por Khadka et al.<sup>97</sup>, que se pueden encontrar en 11.1 (anexo I).

Como se puede observar en la Figura 5.93, las propiedades de ambas escalas son prácticamente idénticas, presentando un comportamiento muy similar ante muestras de poblaciones similares. En ambos casos, la población analizada es predominantemente asintomática y por lo tanto los sujetos están situados en los niveles más bajos de sintomatología dentro del rango que ofrece la escala. El valor del targeting da buena cuenta de la baja prevalencia de IC y de otras alteraciones acomodativas y de vergencias vinculadas a alta sintomatología que se registra en las muestras elegidas.

Las categorías en las dos escalas aparecen perfectamente ordenadas y equidistantes y el grado de precisión de ambas es aceptable de acuerdo con los criterios definidos por Khadka et al.<sup>97</sup>. Los ítems también se ajustan bien al modelo de acuerdo con los estadísticos (Infit y Outfit), propuestos por Rasch al encontrarse dentro de los límites del intervalo esperados.

En cuanto a la dimensionalidad, el comportamiento es parecido. Claramente la sintomatología asociada a la IC presenta varias componentes, lo cual no quiere decir que el constructo sea multidimensional, sino como se justificó en el apartado 5.2.3.3.3, es indicio de que existe correlación entre los diferentes factores que componen la variable latente que se está midiendo.

Como se vio en el análisis factorial, la descomposición de los síntomas en categorías relacionadas con las diferentes etiologías que los producen, da explicación y sentido a la existencia de más de una componente. De hecho, se comprueba que, dependiendo de la edad de los sujetos, los síntomas asociados a la IC se exteriorizan más acentuadamente asociados a fatiga visual en caso de los adultos y más relacionados con problemas de concentración para tareas de cerca en el caso de los niños. Circunstancia totalmente lógica, justificada por la diferente tipología de actividades que se realizan a una edad y a otra. En ambos rangos de edad, sin embargo, se comprueba que los síntomas puramente referidos a disfunción del sistema de visión binocular conservan una prevalencia parecida. Lo cual también es lógico, ya que cuando hablamos estrictamente de ese tipo de disfunciones, la edad no es relevante.

En resumen, las escalas presentan un comportamiento muy similar y tienen una capacidad de discriminación muy parecida, circunstancia que prueba de nuevo la alta calidad de la adaptación transcultural del instrumento y garantiza la fiabilidad en su uso con poblaciones hispanohablantes al mantener todas las propiedades de medida del instrumento original CISSv15 (véase Figura 5.93).

RESULTADOS

PROPIEDADES PSICOMÉTRICAS	CISSve	CISSv15	EXPECTATIVA MODELO RASCH	EVALUACIÓN CALIDAD
<b>Muestra</b>	<b>DESCRIPTIVOS</b> n=439 Edad: (16.2±5.68) Población femenina (%): 55.2% Puntuación: (15,09±10,09) Rango de puntuación: [1, 50] <b>NIVEL DE SÍNTOMAS</b> Nivel 1: 35.67% Nivel 2: 37.12% Nivel 3: 15.19% Nivel 4: 9.18% Nivel 5: 1.82% Nivel 6: 0.23%	<b>DESCRIPTIVOS</b> n=208 Edad: (15.86 ± 1.62) Población femenina (%): 38.03% Puntuación: (16.10 ± 9,50) Rango de puntuación: [1, 49] <b>NIVEL DE SÍNTOMAS</b> Nivel 1: 28.85% Nivel 2: 40.38% Nivel 3: 21.15% Nivel 4: 8.17% Nivel 5: 1.44% Nivel 6: 0%	-	-
<b>Número de ítems</b>	15	15	-	-
<b>Categorías de respuesta</b>	Ordenadas y Equidistantes	Ordenadas y Equidistantes	Ordenadas y Equidistantes	A
<b>Dimensionalidad</b>	PCA: % Varianza explicada por las medidas: 45,4% -> Sugiere multidimensionalidad PCA: Autovalor de la parte de varianza no explicada en el primer contraste: 2.2 -> Sugiere multidimensionalidad PCA: Coeficiente de correlación desatenuada entre el primer y segundo contraste: 0.88 -> Justifica la unidimensionalidad.	PCA: % Varianza explicada por las medidas: 44,9% -> Sugiere multidimensionalidad PCA: Autovalor de la parte de varianza no explicada en el primer contraste: 2.19 -> Sugiere multidimensionalidad PCA: Coeficiente de correlación desatenuada entre el primer y segundo contraste: 0.95 -> Justifica la unidimensionalidad	UNIDIMENSIONALIDAD	C
<b>Precisión</b>	PSI = 2.22 α = 0.85	PSI = 2.31 α = 0.84	PSI > 2.0 α > 0.80	B
<b>Estadísticas de ajuste de los ítems</b>	<b>INFIT</b> 0 ítems fuera del intervalo (0.5-1.5) <b>OUTFIT</b> 0 ítems fuera del intervalo (0.5-1.5)	<b>INFIT</b> 0 ítems fuera del intervalo (0.5-1.5) <b>OUTFIT</b> 1 ítem fuera del intervalo (0.5-1.5)	Infity Outfit entre 0.7 y 1.3	B
<b>Funcionamiento diferencial de los ítems (DIF)</b>	<b>EN CUANTO AL SEXO</b> Todos los ítems con DIF < 0.5 logit <b>EN CUANTO A LA EDAD</b> 3 ítems con DIF > 0.5 logit	<b>EN CUANTO AL SEXO</b> Todos los ítems con DIF < 0.5 logit <b>EN CUANTO A LA EDAD</b> 3 ítems con DIF > 0.5 logit	DIF < 0.5 logits	B
<b>“Targeting” (Acoplamiento Item-Sujeto)</b>	-1.44	-1.16	Diferencia ≤ 1 logit	B

FIGURA 5.93 - COMPARACIÓN ENTRE LAS PROPIEDADES PSICOMÉTRICAS DE LA ESCALA CISSVE Y LA ESCALA CISSv15

La Figura 5.94 ilustra una comparación de los ítems “1 a 1” de la puntuación media de cada ítem en las escalas CISSve y CISSv15. En todos los casos se observa que las diferencias que se encuentran en el nivel de dificultad nunca exceden los 0.5 logits, lo que vuelve a ser muestra de la similitud en cuanto al comportamiento de ambas escalas.

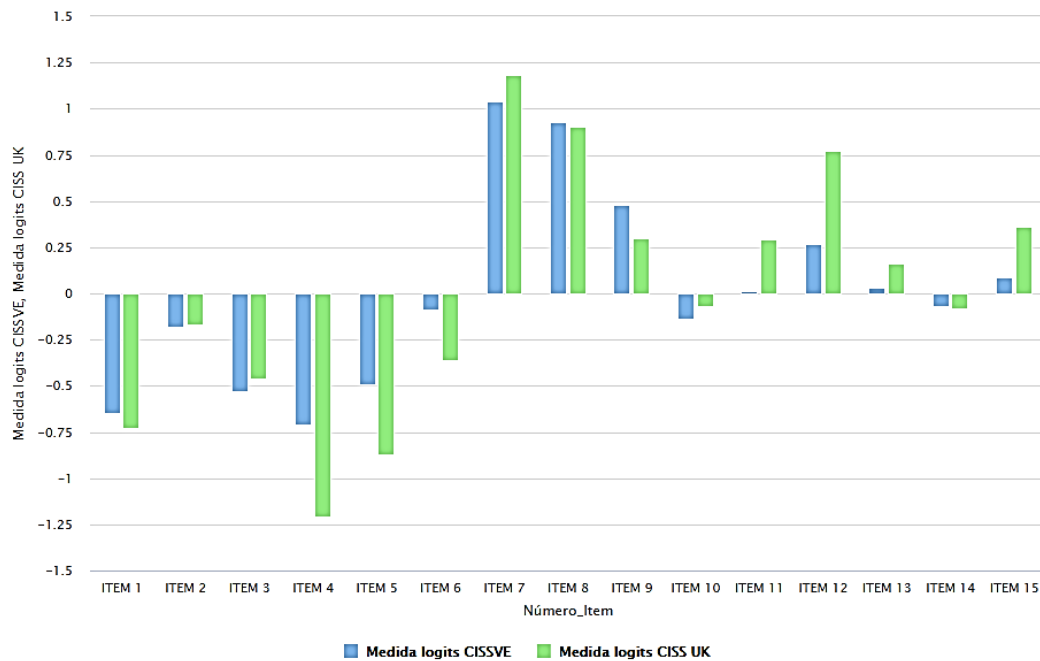


FIGURA 5.94 – COMPARACIÓN ENTRE LOS ÍTEMS DE LA ESCALA CISSVE Y LA ESCALA CISSV15

En la Figura 5.95 se muestra un mapa de calor que presenta qué ítems son los más difíciles (es decir, los ítems en los que las categorías que reflejan niveles de síntomas más altos se usan menos) y cuáles son los más fáciles (es decir aquellos en los que las categorías altas de síntomas son marcadas por un porcentaje más alto de sujetos). Otra vez se puede observar la correspondencia entre ambas escalas.

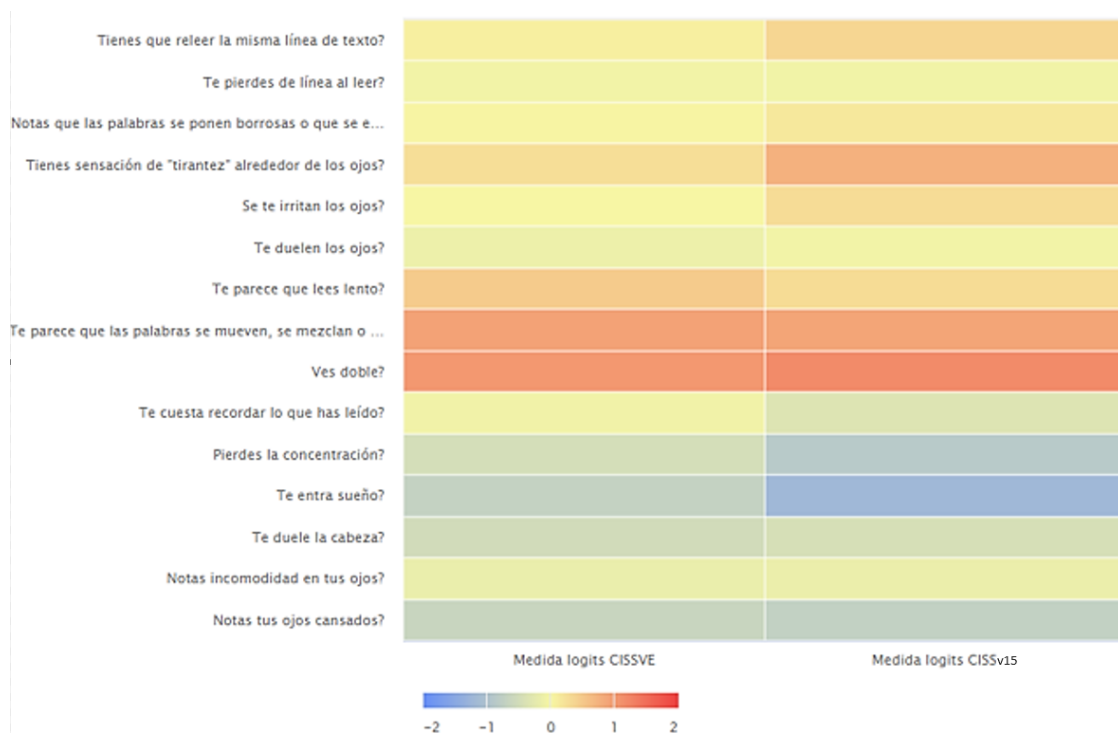


FIGURA 5.95 – NIVEL DE DIFICULTAD DE LOS ÍTEMS DE LA ESCALA CISSVE VS LA ESCALA CISSV15



# DISCUSIÓN



## 6 DISCUSIÓN

Los instrumentos PRO son muy útiles en la práctica clínica general, con fines diagnósticos, en la definición de políticas sanitarias y también en el diseño de nuevas líneas de investigación <sup>97</sup>. Por ello, es importante elegir instrumentos PRO de alta calidad para la práctica clínica y para la investigación científica.

Puesto que no existía ninguna herramienta de estas características en lengua española, capaz de medir síntomas de insuficiencia de convergencia, la presente Tesis Doctoral ha servido para aportar la escala CISSve y cubrir esa necesidad.

### 6.1 ADAPTACIÓN TRANSCULTURAL

La escala CISSve es una adaptación transcultural al español de la escala CISSv15 <sup>60</sup>, siguiendo las directrices planteadas por Gjersing, Caplehorn y Clausen <sup>122</sup>, que mediante el proceso de tres fases descrito en el apartado 4.2, asegura la calidad de la transcripción idiomática, cuidando minuciosamente los aspectos clínicos, lingüísticos, semánticos y culturales con el objeto de conseguir que el comportamiento de la escala adaptada sea exactamente el mismo que el de la original, consiguiendo expandir por tanto la usabilidad del instrumento a poblaciones con diferente lengua materna.

La nueva escala se ha validado siguiendo los procedimientos recomendados para la elaboración y validación de los instrumentos PRO <sup>97,174</sup>, que incluyen tanto métodos convencionales para la elaboración y selección de ítems, como el análisis de Rasch, para garantizar que aquellos proporcionan medidas válidas y fiables.

La validez de contenido indica hasta qué punto los ítems del instrumento reflejan en profundidad el concepto que pretenden medir, con lo que es algo que no puede ser evaluado de forma cuantitativa, se trata de un proceso puramente cualitativo. Sin embargo, existen métodos que permiten valorar de forma sistemática la calidad de los ítems que forman la escala y su propósito e idoneidad dentro del nuevo instrumento PRO.

Para realizar esta fase se realizaron revisiones exhaustivas con los pacientes, con el objeto de asegurar la claridad de los ítems y la correcta interpretación de los mismos de forma precisa <sup>97</sup>. Para ello, se llevaron a cabo tanto sesiones individuales con algunos de los sujetos participantes, como algunas discusiones grupales. Además de contar con la colaboración de los sujetos participantes, se realizó una completa revisión bibliográfica, tanto de estudios publicados sobre el tema de la IC, como sobre cuestionarios validados en el ámbito de la visión. Adicionalmente, se crearon foros de discusión y se realizaron consultas a profesionales del campo optométrico para capturar sus impresiones sobre el valor clínico de cada ítem y su idoneidad para el fin diagnóstico del instrumento.

Completada esta fase y tras realizar una prueba piloto del instrumento (pre-test), se concluyó que tanto la translación de los ítems de inglés a español, como su inteligibilidad y correcta

interpretación por parte de los pacientes era totalmente adecuada y cumplía los criterios de calidad establecidos<sup>94</sup>.

### 6.2 PROPIEDADES PSICOMÉTRICAS DE LA ESCALA CISSve

Una traducción de calidad es una parte esencial de la adaptación intercultural, pero no garantiza que la versión traducida conserve el comportamiento psicométrico de la herramienta original. Autores como Gjersing, Caplehorn y Clausen<sup>122</sup> o también Gandek y Ware<sup>175</sup>, recomiendan un proceso de tres pasos en el que justo después de la traducción y adaptación se realiza la validación de las propiedades psicométricas de la escala y el análisis de su validez u repetibilidad. En este estudio tras realizar la adaptación idiomática, se realizó el análisis por el método de Rasch con el que se confirmó que el rendimiento general del instrumento presentaba un buen comportamiento.

Otros autores como Bradley y Massof<sup>176</sup> recomiendan comparar directamente las propiedades psicométricas de la escala adaptada con las de la original, para comprobar si ambas funcionan de manera similar. En nuestro caso, también se realizó esa comparación, y analizando las propiedades psicométricas de la nueva escala CISSve frente a las del instrumento original en inglés, CISSv15, se pudo comprobar una gran similitud entre ambas como se detalla en el apartado 5.4.

Ambas escalas presentan una precisión prácticamente idéntica (PSI=2.31 en CISSv15 y PSI=2.22 en CISSve) y en el análisis de las componentes principales se reflejan también resultados prácticamente iguales. En las dos escalas las categorías de respuesta aparecen ordenadas, igualmente separadas y equidistantes; y cuando se aplican a muestras de carácter similar, “el targeting”, o el “acoplamiento ítem-persona” que es un parámetro que estima la distancia en logits entre el valor medio de dificultad de los ítems y de habilidad de los sujetos, muestra valores cercanos en ambas escalas (T=-1.16 en CISSv15 y T=-1.44 en CISSve).

En cuanto al ajuste de los ítems a los valores esperados por el modelo, en la escala CISSve se encontró que todos los valores de los residuos (Infit y Outfit) estaban dentro del intervalo considerado aceptable (0.5-1.5). En la versión inglesa, solamente el ítem 12, (“Do you feel a “pulling” feeling around your eyes when reading or doing close work?”) mostró una aparente falta de ajuste (Outfit=1.86). No podemos asegurar si esta falta de ajuste se ha observado en estudios anteriores ya que esta es la primera publicación que analiza un análisis de Rasch de este cuestionario. No obstante, según Khadka<sup>97</sup>, un ítem con este Outfit puede ser aceptable si es el único. Tras descartar las medidas de 6 sujetos en los que se detectaron valores especialmente anómalos con la ayuda de Winsteps, se pudo observar que el valor del Outfit se regularizaba, bajando hasta 0.9, dentro del rango considerado óptimo. Parece claro que se trataba de un problema de colaboración de una pequeña parte de la muestra inglesa a la hora de completar las respuestas al cuestionario. El ítem 1 de la versión en español, “¿Notas tus ojos cansados al leer o realizar trabajos de cerca?”, resultó tener un Infit de 0.61, que a pesar de estar dentro del intervalo de ajuste considerado como aceptable (0.5-1.5), está por debajo de 0.7 lo que indica, que es excesivamente predecible según los valores propuestos por Khadka et al.<sup>97</sup>

En una muestra con un tamaño tan grande, extraída de orígenes heterogéneos (diferentes lugares y entornos) y que barre un intervalo de edades tan grande (9-30 años), no resulta extraño encontrarnos algún sujeto desalineado con el modelo. De los 19 encontrados, en la mayoría de casos pertenecían a la muestra MVL, tomada en un colegio y representada en su mayoría por niños de 9 y 10 años. Esta muestra se observó que era estadísticamente diferente al resto y que presentaba puntuaciones medias bastante más bajas. Circunstancia que por otra parte es lógica<sup>56,177</sup>, ya que es extraño encontrar niños de esa edad con disfunciones visuales binoculares.

Recordemos que se eligió una población en el intervalo 9-30 años para ser coherentes con las muestras observadas por Borsting y Rouse<sup>60</sup>, donde se demostró la validez del CISS para la detección de pacientes sintomáticos en muestras clínicas.

Para confirmar la equivalencia entre ambas versiones, se realizó también un análisis DIF que confirmó que los ítems del CISS habían sido traducidos óptimamente al español (europeo).

Realizando un análisis de la escala mediante el método de Wright que analiza los niveles de síntomas que es capaz de medir una escala con independencia de la muestra estudiada, mucho más indicado para evaluar la calidad de la escala cuando las distribuciones de las muestras con las que contamos están algo sesgadas hacia uno de los lados o cuándo se analizan muestras distintas,<sup>127</sup> encontramos que en la escala CISSve se pueden distinguir hasta 6.3 niveles de síntomas de IC significativamente distintos, valor muy parecido al encontrado para CISSv15 (5.9 niveles). Como vimos, la separación entre categorías obtuvo un nivel B de acuerdo a los criterios de Khadka, lo que resulta coherente con el hecho de que las distancias entre umbrales de Andrich rondan la unidad, en lugar de estar entre 1,4 y 5 como recomienda Khadka para obtener un nivel A en su escala. Esto es consecuencia de que, al tratarse de una muestra predominantemente asintomática, como indican las puntuaciones obtenidas (M=15,1 puntos; DT=10 puntos sobre un máximo de 60 puntos), apenas se discrimina en los niveles superiores, ya que la mayoría de la muestra (el 92%), se encuentra comprendida en los 3 primeros niveles, y el 97,5% (3 desviaciones típicas), en 3.3 niveles de síntomas como nos indican nuestros cálculos. Esto, obviamente tiene sentido considerando que el rango de edad elegido comprende edades entre 9 y 30 años,

El análisis del DIF por sexo no encontró diferencias significativas en el comportamiento de ningún ítem en cuanto al sexo, ni a través el método de Rasch- Welch, ni con el de Mantel-Haentzel. Sin embargo, en el análisis del DIF por rangos de edad, realizado por el método de Mantel-Haentzel sí se vieron ciertas diferencias en cuanto al comportamiento para los ítems 1, (“¿Notas tus ojos cansados?”), 2 (“¿Notas incomodidad al leer?”) y 14 (“Te pierdes de línea al leer?”), encontrando significativo sólo este efecto en el ítem 14 según los criterios de Khadka et al.<sup>97</sup> aplicados sobre el método de Rasch-Welch. Es un hallazgo razonable considerando la gran diferencia existente entre la tipología de la actividad realizada a una edad y a otra. Según estos criterios, una diferencia de 1 logit entre ambos grupos se consideraría un DIF notable, mientras que diferencias entre 0.5 y 1 como los que se encontraron en el ítem 14, serían aceptables siendo considerado “DIF mínimo”. Valores por debajo implicarían ausencia de DIF.

Cuando se administraron las copias en papel de la escala CISSve, los participantes, (excepto la parte de la muestra representada por los niños de educación primaria), también recibieron copias del “Computer Vision Symptom Scale” (CVSS17)<sup>51,127</sup> y del “Warwick-Edimburg Mental Well-Being Scale” (WEBWBS)<sup>118,163</sup> con objeto de poder evaluar respectivamente la validez convergente y la validez divergente de la escala CISSve con esos constructos.

## DISCUSIÓN

De acuerdo con Khadka et al.<sup>97</sup>, un coeficiente de correlación entre las escalas superior a 0.3 se considera como una prueba de validez convergente, mientras un coeficiente de correlación menor, constataría la divergencia entre las escalas. Como se había predicho, se encontró una asociación significativa entre el CISSve y el CVSS17, observando un coeficiente de correlación superior a 0.3. Entre el CISSve y el WEMWBS, sin embargo, no se encontró correlación, por lo que podemos considerar este hecho como una evidencia de validez divergente, ya que según lo esperado no se encontró relación alguna con este instrumento que mide un constructo de una naturaleza totalmente diferente.

Mediante el método de Bland-Altman, se evaluó la repetibilidad obteniendo un COR=9.14. El hecho de que el promedio de las diferencias entre las puntuaciones registradas en cada una de las dos evaluaciones se encontrara cercano a 0 (-0.36) sugirió ausencia de sesgo sistemático, esto es, la no existencia de una tendencia en la que las mediciones fueran sistemáticamente mayores o menores en la segunda iteración en comparación con la primera. Por otra parte, un CoR=9.14 (1,96 veces la desviación típica de la diferencia) indica que en el 95% de los casos, la diferencia entre las puntuaciones registradas por un mismo sujeto en cada una de las dos evaluaciones, era inferior a 9.14 puntos, lo cual representa una buena repetibilidad, considerando una escala en la que la puntuación máxima posible es de 60 puntos. Adicionalmente, a la evaluación mediante el método de Bland-Altman, se calculó el ICC=0.874, con un intervalo de confianza al 95% (0,840; 0,902). Este valor sugiere que la variabilidad dentro de los grupos (es decir, las mediciones repetidas del mismo sujeto) es considerablemente menor que la variabilidad entre los grupos. El intervalo de 0,840 a 0,902 es relativamente estrecho y totalmente contenido en un rango que se considera de buena repetibilidad (generalmente ICC > 0.75), lo cual robustece la interpretación del ICC como un indicador fiable y por lo tanto ratifica la buena repetibilidad que indica su valor, 0.874 ratificando la consistencia alta entre ambas evaluaciones,

La puntuación media del CISSve fue (15,10 ± 10,13 puntos), comparable a la reportada en los estudios realizados en poblaciones similares, como por ejemplo el elaborado en adolescentes por Horan et al.<sup>178</sup> (16.3 ± 11.4) y en universitarios con la versión portuguesa del CISSv15, CISSVP, realizado por Tavares<sup>179</sup>, (CISSVP; 15.56 ± 8.86). En ambos se trataba de poblaciones generales como se ha hecho en este estudio. La puntuación media obtenida en este estudio, proveniente de una población general y separando por edades, en niños (9-17 años, M=11,79 años), media del CISSve (12,92 ± 0,60 puntos) y en adultos jóvenes (18-30 años, M=21,87 años), media del CISSve (18,03 ± 0,71 puntos), es decir, son en promedio 5 puntos superiores a las obtenidas por niños. La media de puntuación es superior a la reportada por Rouse et al.<sup>60</sup> en adultos y niños con visión binocular normal, donde se obtuvo (11.3 ± 8.1) para adultos y (10.4 ± 8.1) para niños. Como cabía esperar, en nuestra muestra, la puntuación media fue inferior a los valores obtenidos por Rouse et al. en adultos<sup>61</sup> y por Borsting et al. en niños<sup>165</sup>, con IC sintomática (37.3 ± 9.3) para adultos y (29.8 ± 8.1) para niños.

Una de las principales fortalezas de nuestra investigación ha sido el elegir un espectro tan amplio de edades en nuestra muestra de estudio, trabajando en un rango desde los niños de 9 años hasta adultos jóvenes de 30, cubriendo así todas las franjas de edad en las que se habían realizado estudios hasta la fecha. Incluso se han capturado algunos datos de población présbita con fines experimentales que se salen del alcance de este trabajo. Esta circunstancia, que a priori podría parecer carente de sentido, ha sido de suma ayuda para poder entender cómo influye la edad en la percepción de síntomas de IC.

Gracias a utilizar el análisis de Rasch como soporte estadístico para evaluar las propiedades psicométricas de la escala, hemos podido llevar a cabo ese propósito. Recordemos que una de las principales propiedades del modelo de Rasch, es la independencia entre la dificultad de los ítems y la habilidad de los sujetos. Es decir, que el hecho de tomar una población u otra, más o menos sintomática, no condiciona el comportamiento de los ítems, en cuanto a su dificultad, con lo que el análisis es válido para cualquier muestra.

Al realizar el análisis de las componentes principales (PCA), dentro del desarrollo del análisis de Rasch, tanto en el caso de la escala original CISSv15, como en la adaptación, se obtuvieron indicios de multidimensionalidad, al encontrar que la varianza explicada por las medidas era ligeramente inferior al 50% en ambos casos y existía un autovalor (eigenvalue) en el primer contraste ligeramente superior a 2.0<sup>97,114</sup> (Apartado 5.2.3.3.3 ).

Al analizar con detalle la naturaleza de los síntomas a los que apuntaban los ítems que componían la supuesta segunda dimensión, se observó que había relación entre ellos y que todos ellos se referían a la capacidad de concentración para las tareas asociadas con la lectura, (“¿Pierdes la concentración al leer...?”, “¿Tienes problemas para recordar lo que has leído?”, “¿Tienes que releer la misma línea?”, “¿Sientes que lees lento?”, “¿Te entra el sueño al leer?”) y no directamente a sintomatología directamente reacionada con una disfunción visual. Sin embargo, debido a la altísima relación que sabemos que existe entre este tipo de sintomatología y el resto de sintomatología de IC<sup>180</sup> a que apunta la dimensión principal del análisis de Rasch, observando el coeficiente de correlación desatenuado entre el primer y segundo contraste obtenido mediante el Análisis de PCA se encontró un valor de 0.88. El valor de la correlación desatenuada, corresponde aproximadamente al valor de la correlación entre los dos contrastes sin error de medición. Cuando está cerca de 1, implica que las dos dimensiones miden realmente atributos de la misma variable latente. Según Linacre<sup>107</sup>, el valor 0.82 debería ser usado como límite para considerar que los dos contrastes miden la misma variable, ya que indicaría que son entre sí, el doble de dependientes que de independientes. Mientras que el valor 0.87, indicaría que el factor de dependencia sería triple al de independencia. En nuestro caso encontramos que incluso el valor encontrado supera ese umbral (0,87), con lo que concluimos que en realidad, las dos dimensiones del análisis PCA son atributos que contribuyen a medir síntomas de IC dentro del mismo constucto.

Entrando un poco más al detalle con este interesante hallazgo, realizamos un análisis factorial (Apartado 5.2.3.3.7), extrayendo las componentes principales otra vez por medio de un análisis PCA, realizando una rotación Varimax para evitar ruido. Tomando los autovalores mayores que 1, encontramos 3 factores (ver Figura 5.56). Profundizando un poco más en la etiología que subyace detrás de los síntomas de cada factor intentando entender mejor la tipología de síntomas que comparten los ítems que los componen, observamos como los ítems del factor 1 (ítem 1, 2, 3, 10, 11 y 12) comparten expresión de síntomas que indican molestias asociadas con la fatiga visual derivada del trabajo prolongado de cerca. Los ítems del factor 2 (7, 8, 9, 13 y 14) muestran síntomas derivados de un en la función visual binocular y/o acomodativa y por último, los ítems del factor 3 (4, 5, 6 y 15) denotan exteriorizaciones sintomáticas que indican pérdida de atención o dificultad para concentrarse en las tareas de lectura o trabajo de cerca. Revisando la literatura científica, encontramos un reciente artículo de la Universidad de Beiría<sup>181</sup> en el que se realiza un análisis factorial a la escala CISSVP (versión adaptada al portugués de la escala CISSv15 ) para una muestra de 183 adultos jóvenes universitarios. Este artículo sólo se concentra en el grupo de edad de 18-30, pero es curioso observar que los factores identificados coinciden con los hallados en este estudio.

## DISCUSIÓN

Analizando con más detalle la incidencia de cada tipología de síntomas en nuestros grupos de estudio (Figura 5.60), observamos cómo los adultos jóvenes con edades entre 18 y 30 años presentan mayor porcentaje de la sintomatología percibida, explicada en base a síntomas relacionados con la fatiga tras trabajos prolongados de cerca que los niños de entre 9 y 17 años, mientras que los niños, explican un porcentaje mayor de sus síntomas a través del factor que relaciona los asociados a la dificultad de concentración para tareas relacionadas con la lectura o el trabajo de cerca. Como también es lógico, la sintomatología asociada a disfunción visual por problemas de visión binocular y/o acomodativos se mantiene aproximadamente constante con la edad.

Cuando analizamos las medidas tomadas durante la investigación a una muestra pequeña de sujetos presbíta que participaron ( $n=29$ ,  $M=60,86$ ;  $DT=13,82$ ), y que no se habían considerado por estar fuera del rango de edad que cubre este estudio (9-30 años), pudimos observar cómo se comportaban sus puntuaciones sobre la escala en cuanto a las diferentes tipologías de síntomas determinadas por los factores encontrados (Figura 5.89) en la sección de resultados (Apartado 5.2.3.4.3). Se pudo apreciar cómo con la edad resultaba acentuarse aún más la polarización de los síntomas en los presbíta hacia los relacionados con la fatiga visual al realizar trabajo prolongado de cerca. Es algo esperable al tratarse de un grupo en el que clínicamente es frecuente encontrar sintomatologías de ojo seco y refracciones mal ajustadas, aunque debido al pequeño tamaño de la muestra, es algo que se debe estudiar en más detalle en próximos trabajos.

Este hallazgo referido al comportamiento diferente en cuanto a la percepción de los síntomas por parte de menores y adultos jóvenes es consistente con los resultados que se pudieron observar tras el estudio del funcionamiento diferencial de los ítems (DIF) que se realizó aplicando el método de Mantel-Haenszel, y a través del cual, se podía comprobar que el funcionamiento de los ítems 1, 2, y 14, (todos ellos pertenecientes al factor que agrupa los síntomas derivados de la fatiga visual tras un trabajo prolongado de cerca) era diferente cuando observábamos el comportamiento de los niños y los adultos jóvenes (Apartado 5.2.3.3.6). Conviene comentar que las puntuaciones son mayores (más síntomas) en sujetos más mayores pero que las propiedades psicométricas del cuestionario son similares en ambos grupos de edad.

En resumen, como se ha demostrado el CISSv15 y el CISSve son dos escalas equivalentes capaces de medir síntomas de IC en todos los rangos de edad y en poblaciones con incidencia diversa. Este hecho proporciona a la escala una alta versatilidad como punto fuerte, pero a su vez le resta especificidad como punto débil. Esa falta de especificidad es algo sobre lo que se podrá seguir trabajando y profundizando en estudios venideros. Quizá lo aprendido con este trabajo nos sirva para plantear nuevas vías como por ejemplo escalas más cortas parametrizadas según las características de cada franja de edad que permitan ganar especificidad, que sean más rápidas de administrar y mejoren la precisión en la zona de medida.

Cuando calculamos por el método de Wright (que es independiente de la muestra), el número de diferentes niveles de síntomas que es capaz de medir el instrumento (ver Figura 5.37 y Figura 5.38 del capítulo de resultados) se evidencia su alto poder de discriminación (hasta 6.3 niveles significativamente distintos), hecho otorga al CISSve un espectro de medición amplísimo. El porcentaje de sujetos encontrados en nuestra muestra en los niveles más altos (Figura 5.39), es coherente con la prevalencia de problemas binoculares y acomodativos descrita por Scheiman y Wick <sup>11</sup> (aproximadamente el 10%).

Como sabemos, en el modelo de Rasch, la medida en un sujeto es más precisa en los ítems que parean su nivel de dificultad con la habilidad de ese sujeto. Para la muestra incluida en nuestro estudio, una población general predominantemente asintomática, se observa que prácticamente sólo entran en funcionamiento los primeros niveles de síntomas de la escala, existiendo sin embargo un amplio margen de medición para pacientes con sintomatologías más severas. (ver Figura 5.37, Figura 5.38 y Figura 5.39). Esto quiere decir, que los ítems de la escala que más precisión aportan para medir a la muestra elegida son los de la parte baja del mapa de Wright, o en otras palabras, los más fáciles, los ítems en los que los sujetos, a pesar de no tener sintomatologías elevadas, otorgan puntuaciones en categorías superiores con más frecuencia que en los más difíciles. Por ello, estos ítems, son más aptos para gradar las diferencias entre sujetos situados en los niveles más bajos de síntomas, haciendo distinción entre los que son ligeramente más sintomáticos que otros que pasan desapercibidos en ítems más difíciles. El indicador del modelo de Rasch, “targeting” que indica la adaptación “item-sujeto”, presentando un valor de -1.44, es coherente con este hecho y refuerza la misma idea. Los ítems de la escala con dificultad más baja serán los más aptos para discriminar entre diferentes estratos de síntomas de los sujetos de la muestra elegida para el análisis.

La Figura 5.43 del capítulo de resultados muestra el mapa de Wright en el que se observa el descentramiento existente entre los ítems de la escala y la muestra elegida para el análisis en nuestro estudio. La interpretación es bastante intuitiva. Si se observa la tabla de la Figura 5.44, se puede observar que un sujeto de habilidad media, es decir que se encuentra 1.44 logits debajo del “0 de escala”, como nos dice el valor del “targeting”, (ya que el “0 de escala” marca la dificultad media de los ítems), tendrá menos de un 20% de probabilidad respecto a un ítem situado en el “0 de escala” de detectar síntomas a ese sujeto. Un descentramiento tan grande implica, que la precisión que pueden aportar la mayoría de los ítems para discriminar síntomas en una gran proporción de la muestra, es muy baja, por lo que se deduce que para una muestra de estas características, una población general predominantemente asintomática, probablemente una escala mucho más reducida que prescindiera de los ítems en la zona más alta de síntomas, nos proporcionaría una fiabilidad similar.

En cuanto al valor diagnóstico de la escala CISSve, tras comparar las puntuaciones del grupo de sujetos diagnosticados clínicamente ( $n=49$ ,  $M=19.80$ ,  $DT=9.13$  años), con los de otro grupo de población general ( $n=52$ ,  $M=14.02$ ,  $DT=6.36$  años), se constató la utilidad de la escala CISSve como instrumento para distinguir entre estos dos grupos. No obstante, el hecho de que la puntuación media de los sujetos diagnosticados arrojara valores en promedio por debajo de la mitad de la puntuación máxima posible en el CISSve, sugiere que quizá una escala con menos categorías podría ser suficiente para este mismo propósito.

De la misma forma que se pudo constatar en el experimento en el que se sometió el grupo de pacientes clínicamente diagnosticados a terapia visual, los síntomas que más diferenciaron a los sujetos sintomáticos de la población general fueron los relacionados con problemas de función visual binocular y fatiga visual tras un trabajo prolongado de cerca. No se observaron variaciones significativas en los síntomas asociados a la falta de concentración entre los pacientes diagnosticados clínicamente y la población general, circunstancia que refuerza la idea de que los síntomas asociados a la pérdida de concentración, pueden no ser tan relevantes de cara a la detección de IC. Adicionalmente, tampoco se encontraron diferencias significativas en los síntomas de IC entre sexos, pero sí se observó sin embargo que los adultos jóvenes refieren en promedio una mayor sintomatología asociada a la fatiga derivada del trabajo prolongado de cerca que los niños de entre 9 y 17 años. Esta situación parece

## DISCUSIÓN

consecuente con el hecho de que en la franja de edad de los 18 a 30 años habitualmente se emplea un lapso temporal mayor en tareas que requieren trabajo prolongado de cerca en comparación con los niños de 9 a 17, sobre todo en algunos tipos de trabajo; y por lo tanto esta circunstancia podría influir en que los síntomas relacionados con actividades prolongadas de trabajo de cerca resulten más relevantes para adultos jóvenes entre 18 y 30 años que para los niños.

En relación con los beneficios potenciales de la terapia visual aplicada a población diagnosticada clínicamente, a pesar de haber utilizado una muestra relativamente pequeña ( $n=24$ ,  $M=17.29$ ,  $DT=6.74$  años), nuestros resultados demuestran una mejora considerable tras el tratamiento en pacientes con IC. Para los 24 sujetos que participaron en el estudio antes de la terapia, la puntuación media en el cuestionario CISSve era de 20.04 (con una desviación estándar de 9.92). Este valor es más alto que el de la población general, donde la puntuación media fue de 15.08 ( $DT=10.07$ ) en una muestra total de 458 sujetos. Esto era esperable, dado que todos los pacientes en la muestra inicial eran sintomáticos. Tras la terapia visual, la puntuación media de los pacientes disminuyó a 8.13 ( $DT=4.98$ ), incluso por debajo de la puntuación media de la población general. Esta reducción tan significativa en la puntuación demuestra la alta eficacia de la terapia visual.

Encontramos que los síntomas más reducidos fueron los asociados con problemas de función visual binocular y los asociados a la fatiga visual tras un trabajo prolongado de cerca, mientras que sin embargo, no se observó una variación significativa en los síntomas relacionados con la falta de concentración antes y después de la terapia, lo cuál sugiere que los síntomas de falta de concentración pueden no ser tan relevantes o tan sensibles para la detección de IC como otros síntomas, como los asociados a la función visual binocular y la fatiga visual tras un trabajo prolongado de cerca. En otras palabras, la insuficiencia de convergencia puede no afectar de manera significativa la capacidad de concentración de un individuo o, al menos, no en la medida que afecta a otras áreas de la función visual o puede que la terapia visual no sea tan efectiva para abordar ese tipo de síntomas de falta de concentración como lo es para tratar otros síntomas de IC.

### 6.3 LIMITACIONES DEL ESTUDIO

El presente trabajo, centrado en la adaptación transcultural al español de la escala CISSv15 (Convergence Insufficiency Symptom Survey), aporta contribuciones relevantes al ámbito de la oftalmología y la psicometría. No obstante, es necesario reconocer una serie de limitaciones que podrían afectar la interpretación y la generalización de los resultados obtenidos.

Por una parte, a pesar de contar con una muestra compuesta por 458 sujetos, suficiente a nivel estadístico para sacar resultados concluyentes (apartado 4.2.3.2), ésta está circunscrita exclusivamente a la población española. Esta delimitación geográfica y lingüística introduce limitaciones en cuanto a la aplicabilidad de la escala CISSve en contextos hispanohablantes distintos al español de España, dadas las variaciones idiomáticas y culturales potenciales.

El rango de edad de los participantes, entre 9 y 30 años, seleccionado para adecuarse a las mismas franjas de edad utilizada en los estudios de validación de la escala original CISSv15,

acota la aplicabilidad de los resultados, excluyendo a la población infantil y a los adultos mayores. Asimismo, las submuestras clínicas evaluadas (sintomática y no sintomática) poseen tamaños modestos, lo cual limita la generalización de las inferencias.

En cuanto a la selección de centros para la administración de la escala, aun habiendo sido cuidadosamente seleccionados, cabría la posibilidad de introducir pequeños sesgos en los resultados, relacionados con características idiosincráticas, como podría ser su ubicación geográfica (urbana /rural) o el tipo de institución (pública/privada).

Por último, el estudio no contempla variables como el nivel socioeconómico de los participantes, su nivel educativo o la presencia de otros problemas de salud de tipo general en algunos sujetos que de alguna manera pudieran influir en la forma de responder de los sujetos.

Es fundamental considerar estas limitaciones al interpretar y generalizar los hallazgos del estudio. Aunque no menoscaban la importancia de la investigación, sí ofrecen un marco crítico que contextualiza su relevancia y delimita su aplicabilidad.

## 6.4 PROPUESTA PARA POSIBLES NUEVAS LÍNEAS DE INVESTIGACIÓN

Aprovechando toda la información extraída durante el estudio que nos ha llevado a comprender pormenorizadamente todas las propiedades psicométricas y el comportamiento de la escala CISSve, gracias a la gran cantidad de datos recogidos y al potencial del análisis de Rasch, estamos en posición de hacer propuestas de nuevas líneas de investigación para conseguir reingenierías del instrumento que mejoren su especificidad y usabilidad.

Una de las iniciativas sería la obtención de una versión simplificada de la escala CISSve con el fin de obtener una escala mucho más corta y simple que posibilite realizar “screenings”<sup>182</sup> de forma muy rápida en poblaciones de carácter general y así proporcionar a los profesionales del cuidado de la visión una herramienta capaz de marcar sujetos susceptibles de ser examinados posteriormente de forma más pormenorizada en un gabinete clínico. En este trabajo se ha modelizado una versión reducida de CISSve. Para analizar su rendimiento, se ha evaluado su comportamiento frente al de la versión de 15 ítems, comparando las medidas de ambas escalas de acuerdo a los criterios de discriminación definidos en cada caso. Sería interesante en futuros trabajos, analizar el comportamiento clínico de esta herramienta comparándolo con el resultado diagnóstico después de realizar pruebas optométricas en gabinete a los mismos sujetos que completen el cuestionario.

Otra propuesta sería la de tratar de segmentar la escala CISSve en base a los factores encontrados, creando varias subescalas parametrizadas para cada rango de edad. Aprovechando el conocimiento adquirido sobre el comportamiento de la muestra respecto a cada tipo de síntomas, se pueden ajustar los ítems para cada grupo de interés, mejorando la precisión en la zona de medida y consiguiendo más especificidad, a la vez que una escala más corta y por lo tanto, más rápida de administrar.

## DISCUSIÓN

También sería interesante intentar buscar nuevos ítems que mejoraran la precisión en la zona de medida para muestras poco sintomáticas. Para ello se necesitaría encontrar algunos ítems adicionales que se acercaran más a la media de una muestra poco sintomática, de forma que se mejorara la precisión en esa zona, es decir, “ítems fáciles”. La inclusión de ítems de baja dificultad mejoraría el “targeting” y por lo tanto, la precisión de las medidas para ese tipo de poblaciones.

Durante la fase de toma de medidas, se administraron algunos ítems adicionales a los mismos sujetos que colaboraron en el estudio. No obstante, sería interesante seguir avanzando en esa línea para crear un amplio banco de ítems, y aprovechando la potencia del análisis de Rasch, que nos posibilita ordenar por dificultad (o entendido de otra forma, susceptibilidad a los síntomas) los ítems, poder adaptar escalas a poblaciones según la incidencia de la IC, es decir, quizá sería apropiado utilizar ítems diferentes en poblaciones clínicas sintomáticas que en poblaciones generales asintomáticas.

También en relación con este punto, sería interesante explorar una posible consolidación de ítems por grupos según su etiología, creando otros nuevos capaces de cubrir unívocamente los factores (dimensiones) encontrados durante el estudio, por ejemplo: “¿Notas que te pierdes de línea al leer, se te mezclan las líneas o tienes algún tipo de sensación de que las palabras se muevan o incluso alguna vez ves doble?” recogería los síntomas asociados a la “disfunción visual binocular”, “¿notas que no recuerdas lo que lees y tienes que releer o que pierdes la concentración mientras lees a trabajas de cerca?” recogería los síntomas asociados a la falta de concentración en la lectura y “¿notas sensación de cansancio o molestias en los ojos al final del día (tirantez, rojez, irritación, incomodidad)?” recogería la dimensión asociada a la fatiga visual tras un trabajo prolongado de cerca. Una posible línea continuista, sería el tratar de crear una escala más corta que CISSve que, cubriendo todas las componentes de la IC, permitiera contrastar los hallazgos encontrados para CISSve y estudiar, desde otro prisma y quizá con el soporte de pruebas clínicas, la influencia de la edad en la forma que se percibe la sintomatología de la IC.

Por último, aprovechando el potencial de la “minería de datos” y explotando el potente análisis predictivo que facilita, un punto de evolución consistiría en la creación de tests adaptativos mediante árboles de decisión dinámicos que consiguieran adaptar las preguntas de la escala a los sujetos en función de las respuestas a las anteriores. Recordemos que el modelo de Rasch sostiene que la dificultad de un ítem no depende de la muestra medida ni de los demás ítems y en base a estas premisas, un sujeto que contestara de determinada forma a un ítem podría ser dirigido hacia el siguiente en función de su respuesta, minimizando el camino crítico hasta conseguir clasificarlo. En este trabajo se presenta de forma sucinta en el apartado de “Aportaciones adicionales”, una introducción a este campo, pero sería interesante seguir profundizarlo en ese ámbito con el fin de darle aplicación clínica. Por ejemplo, el desarrollo de una aplicación para móviles o tablets capaz de administrar tests adaptativos basados en algoritmos de aprendizaje, podría significar una gran aportación para una investigación futura.

En resumen, la escala CISSve es susceptible de varias reingenierías:

- Con objeto de mejorar el comportamiento de la escala en la medición a muestras poco sintomáticas se sugiere la inclusión de nuevos ítems en la zona de medida de esas poblaciones (ítems fáciles), que aumenten la fiabilidad de la escala y calidad de las medidas en esa zona.

- Con objeto de una administración más rápida, a modo “screening” a grandes poblaciones y la clasificación de los sujetos para su posterior análisis pormenorizado en gabinete, se sugiere la reducción de la escala CISSve, recortando el número de ítems, dejando los que realmente sirven para discriminar en la zona correspondiente a poblaciones generales.
- Con objeto de hacer más precisa en sus medidas a la escala CISSve se sugiere la segmentación de la escala en varias subescalas, que sean aplicables según el rango de edad y que estén parametrizadas para medir en la zona de interés.
- Con objeto de clasificar rápidamente a los sujetos con el número mínimo de preguntas, se sugiere trabajar en test adaptativos que permitan ajustar la dificultad de los siguientes ítems en función de las primeras respuestas, de forma que con el mínimo número de preguntas se consiga medir con exactitud el nivel de síntomas de un sujeto



# CONCLUSIONES



## 7 CONCLUSIONES

Presentamos a continuación las conclusiones obtenidas de la adaptación transcultural de la escala CISSv15 a español (CISSve) y su posterior análisis mediante el método de Rasch, así como las obtenidas al aplicar la escala CISSve a población diagnosticada clínicamente.

### 7.1 CONCLUSIONES DERIVADAS DE LA ADAPTACIÓN TRASCULTURAL DE LA ESCALA CISSv15

- 1) Mediante la adaptación transcultural de la escala “Convergence Insufficiency Symptom Survey” (CISSv15), se ha creado la escala CISSve, que es el primer instrumento PRO (Patient Reported Outcomes) validado y en lengua española capaz de medir síntomas de insuficiencia de convergencia.
- 2) El análisis de Rasch ha confirmado que la escala CISSve posee propiedades psicométricas similares a la escala original CISSv15.
- 3) La escala CISSve, a pesar de medir un único constructo (la IC), está compuesto por varios factores que guardan correlación, un primer factor asociado con los síntomas relacionados con la fatiga visual tras actividades prolongadas relacionadas con la lectura o el trabajo de cerca. Un segundo factor vinculado a los problemas de concentración para la lectura o actividades en el ámbito del trabajo de cerca y por último, un tercer factor asociado con síntomas asociados a la disfunción de la visión binocular.
- 4) Los diferentes grupos de edad, reportan distinta proporción de síntomas asociada a cada factor. A medida que aumenta la edad, crece la percepción de síntomas asociados a la fatiga visual tras actividades prolongadas relacionadas con la lectura o el trabajo de cerca en detrimento de los síntomas relacionados con los problemas de concentración.
- 5) Las puntuaciones medias obtenidas en la escala por adultos jóvenes (18-30 años),  $M=18,03$  puntos, son en promedio 5 puntos superiores a las obtenidas por niños (9-17 años),  $M=12,94$  puntos. Este punto es concordante con la diferencia en el punto de corte para clasificar a los sujetos como sintomáticos en esos mismos rangos de edad, Puntuación $>16$  para niños y Puntuación $>21$  para adultos.
- 6) La escala presenta seis niveles de sintomatología estadísticamente distintos, lo que simplifica la interpretación clínica de la puntuación obtenida en el CISSve, al permitir establecer el grado de severidad de los síntomas en muestras sintomáticas.

## CONCLUSIONES

- 7) Los resultados de los análisis sobre validez convergente con el CVSS17 y validez divergente con el WEMWBS son una prueba más de que esta nueva escala proporciona datos válidos y mide adecuadamente el constructo que pretende medir.
- 8) Las medidas obtenidas con la escala CISSve, presentan una buena repetibilidad.
- 9) La escala CISSve está libre de DIF en cuanto al sexo. Sin embargo, sí hay diferencia por grupos de edad en los ítems, 1, 2, y 14, asociados con la concentración en la lectura.

## 7.2 CONCLUSIONES DERIVADAS DE LA APLICACIÓN DE LA ESCALA CISSve A POBLACIÓN DIAGNOSTICADA CLÍNICAMENTE

- 1) La puntuación de los sujetos diagnosticados clínicamente (n=49, M=19,80, DT=9.13 puntos) es en promedio superior a los de una población (n=52, M=14.02, DT=6.36 puntos) en la que conviven sujetos sintomáticos con un grupo mayoritario de sujetos no sintomáticos.
- 2) La puntuación de los sujetos sintomáticos diagnosticados clínicamente (n=49, M=19,80, DT=9.13 puntos), es inferior a la mitad de la máxima puntuación alcanzable en el CISSve. Esto indica que probablemente una escala con un número menor de categorías aportaría una información similar a esta escala.
- 3) La eficacia del tratamiento con terapia visual demostró ser altísima, reduciendo significativamente los síntomas de la mayoría de los pacientes y la escala CISSve fue capaz de detectar la mejoría, encontrándose una diferencia media entre la puntuación pre y post >10 puntos.
- 4) Los síntomas que más disminuyeron fueron los asociados a los problemas de la función visual binocular, seguidos de los asociados a la fatiga visual tras un trabajo prolongado de cerca
- 5) Los síntomas asociados a la falta de concentración no sufrieron variaciones significativas entre las dos repeticiones del cuestionario CISSve, antes y después de la terapia visual. Esta circunstancia es relevante, ya que se constata que los ítems asociados a localizar ese tipo de síntomas posiblemente nos son tan relevantes en la detección de IC, como los asociados a detectar disfunción visual o fatiga.
- 6) Los síntomas que en los que más diferencia mostró la población sintomática diagnosticada clínicamente respecto a la población general fueron los asociados a los problemas de la función visual binocular, seguidos de los asociados a la fatiga visual tras un trabajo prolongado de cerca.
- 7) Los síntomas asociados a la falta de concentración como se pudo comprobar a través de la prueba t, no mostraron variaciones significativas entre los pacientes

diagnosticados clínicamente sintomáticos y los pertenecientes a una población general. Esta circunstancia es relevante, ya que implica que los ítems asociados a localizar ese tipo de síntomas posiblemente no son tan relevantes en la detección de IC, como los asociados a detectar disfunción visual o fatiga, circunstancia que por otra parte ya se observó al comparar las respuestas de los pacientes sintomáticos antes y después de la terapia visual.

- 8) No se observaron diferencias significativas en cuanto al sexo en cuanto a la referencia de síntomas relacionados con la insuficiencia de convergencia por parte de los sujetos que participaron en el estudio.
- 9) Se observa que los adultos jóvenes refieren en promedio una mayor sintomatología asociada a la fatiga derivada del trabajo prolongado de cerca que los niños entre 9 y 17 años. Aunque también se observan pequeñas diferencias en cuanto a la falta de capacidad de concentración y a los síntomas asociados a disfunciones de la función visual binocular, estas resultaron ser estadísticamente no significativas.

## CONCLUSIONES

**APORTACIONES  
ADICIONALES  
DERIVADAS  
DE LA TESIS**



## 8 APORTACIONES ADICIONALES DERIVADAS DE LA TESIS

Este trabajo ha servido para crear la escala CISSve, que es el primer instrumento PRO (Patient Reported Outcomes) validado y en lengua española capaz de medir síntomas de insuficiencia de convergencia. La escala CISSve es una adaptación transcultural de la escala “Convergence Insufficiency Symptom Survey” (CISSv15)<sup>59</sup> y sus propiedades son equivalentes a las del instrumento original.

Adicionalmente, aprovechando la información extraída durante el estudio se proponen algunas cuestiones adicionales.

### 8.1 REDUCCIÓN DE ITEMS PARA OBTENCIÓN DE ESCALA CISSve ABREVIADA PARA “SCREENINGS”

#### 8.1.1 OBJETIVOS

Con esta aportación se pretende desarrollar una versión abreviada de la escala CISSve para su uso en “screenings” rápidos en poblaciones generales de edades comprendidas entre 9 y 30 años. El objetivo secundario consiste en evaluar la validez y el “targeting” de la versión reducida del cuestionario.

#### 8.1.2 JUSTIFICACIÓN

La necesidad de una versión abreviada de la escala CISSve se fundamenta en habilitar la capacidad de poder realizar evaluaciones psicométricas de manera más eficiente, tanto en tiempo como en recursos. El instrumento completo, aunque exhaustivo, puede ser demasiado largo para aplicaciones de screening masivo. Una versión reducida, que mantenga la validez y el “targeting” del cuestionario original, facilitaría la identificación de casos que requieren una atención más detallada, optimizando así la asignación de recursos en estudios a gran escala.

#### 8.1.3 HIPÓTESIS

La hipótesis principal postula que es posible elaborar una versión reducida de la escala CISSve que conserve un nivel aceptable de validez y mantenga un “targeting” adecuado. La hipótesis secundaria sugiere que la eliminación de ítems con bajas correlaciones biserial puntuales y/o con mal ajuste a las métricas de Infit y Outfit no comprometerá las propiedades psicométricas del cuestionario reducido.

#### 8.1.4 MÉTODOS

Se revisaron diversas guías sobre diseño de cuestionarios <sup>94,114,122,126</sup> y se implementó una metodología iterativa fundamentada en las recomendaciones de de Pesudovs, K. et al. <sup>94,128,183</sup> para la selección y descarte de ítems. Los criterios utilizados abarcaron la correlación biserial puntual, Infit y Outfit, dificultad relativa de los ítems, porcentaje de respuestas nulas, porcentaje de respuestas en la categoría de mayor dificultad, desviación estándar y normalidad en el patrón de respuestas. La eficacia de cada ítem se evaluó con base en estos indicadores.

Para realizar las evaluaciones estadísticas y el análisis factorial se utilizó el paquete de IBM SPSS Statistics versión 25.0 <sup>137</sup>, y para el análisis de las propiedades psicométricas de la escala, se aplicó el método de Rasch, apoyando los cálculos en el paquete de software WINSTEPS (Versión 4.0.1) <sup>107</sup>.

#### 8.1.5 RESULTADOS

Los criterios de reducción aplicaron de forma iterativa en el siguiente orden de acuerdo a las recomendaciones de Pesudovs, K. et al. <sup>94,128,183</sup>

- 1) Se inspeccionaron los 15 ítems con el fin de descartar los que tuvieran un coeficiente de correlación biserial puntual (Pt. Bis Correlation) inferior a 0.4, lo que implicaría un bajo poder de discriminación, pero se encontró que no había ninguno por debajo del mínimo exigido. Es más, los valores que se encontraron mostraban una discriminación excelente como se puede observar en la Figura 8.1. La correlación biserial puntual es una correlación producto-momento de Pearson entre una variable categórica (puntuación en el ítem concreto) y una variable continua (puntuación total en la escala CISSve). Este coeficiente indica cuánto contribuye cada ítem al constructo total medido por la escala. Valores más altos indican que el ítem está bien alineado con el constructo que se está midiendo, y viceversa. En general, los ítems con correlaciones biserial puntuales bajas podrían revisarse o eliminarse para mejorar la calidad de la prueba.

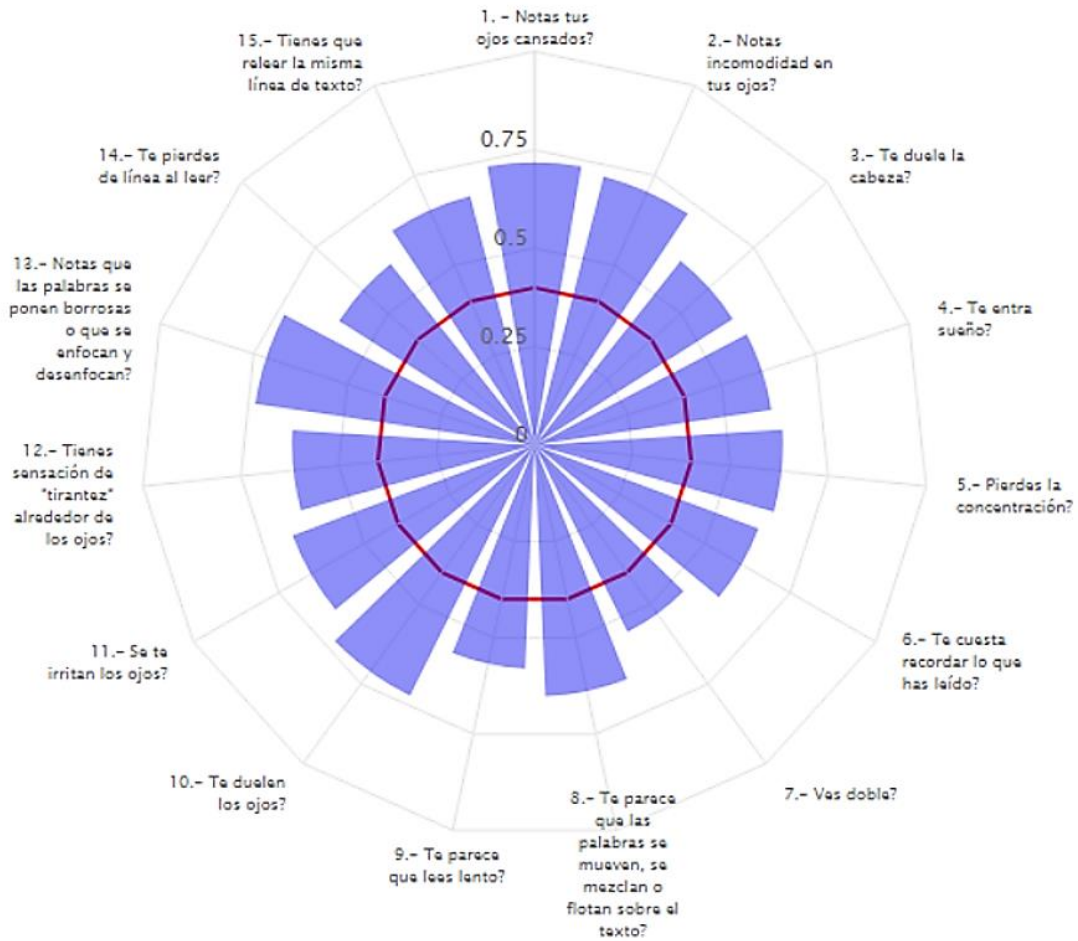


FIGURA 8.1 – CORRELACIÓN BISERIAL PUNTUAL DE LOS ÍTEMS DE LA ESCALA CISSVE

- 2) Se examinó el Infit y Outfit de los 15 ítems para descartar los que tuvieran los valores más elevados, es decir, los que, peor se ajustaban al modelo. No se consideró descartar los que tenían un ligero sobreajuste con el modelo por coincidir con los que más cerca se encontraban en dificultad de la habilidad media de los sujetos y por ser por esa razón los más precisos para la muestra considerada. Como ya se revisó en la sección de resultados, todos los ítems, estaban dentro del intervalo mínimo recomendado según los criterios de Khadka et al.<sup>114</sup>, (0.5,1.5), pero de cara a ser más exigentes y descartar los que a pesar de estar dentro de los límites mínimos aún estaban lejos del valor ideal, se situó el umbral en 1.2, de forma que se descartaron los ítems 3 y 4 como se observa en la Figura 8.2.

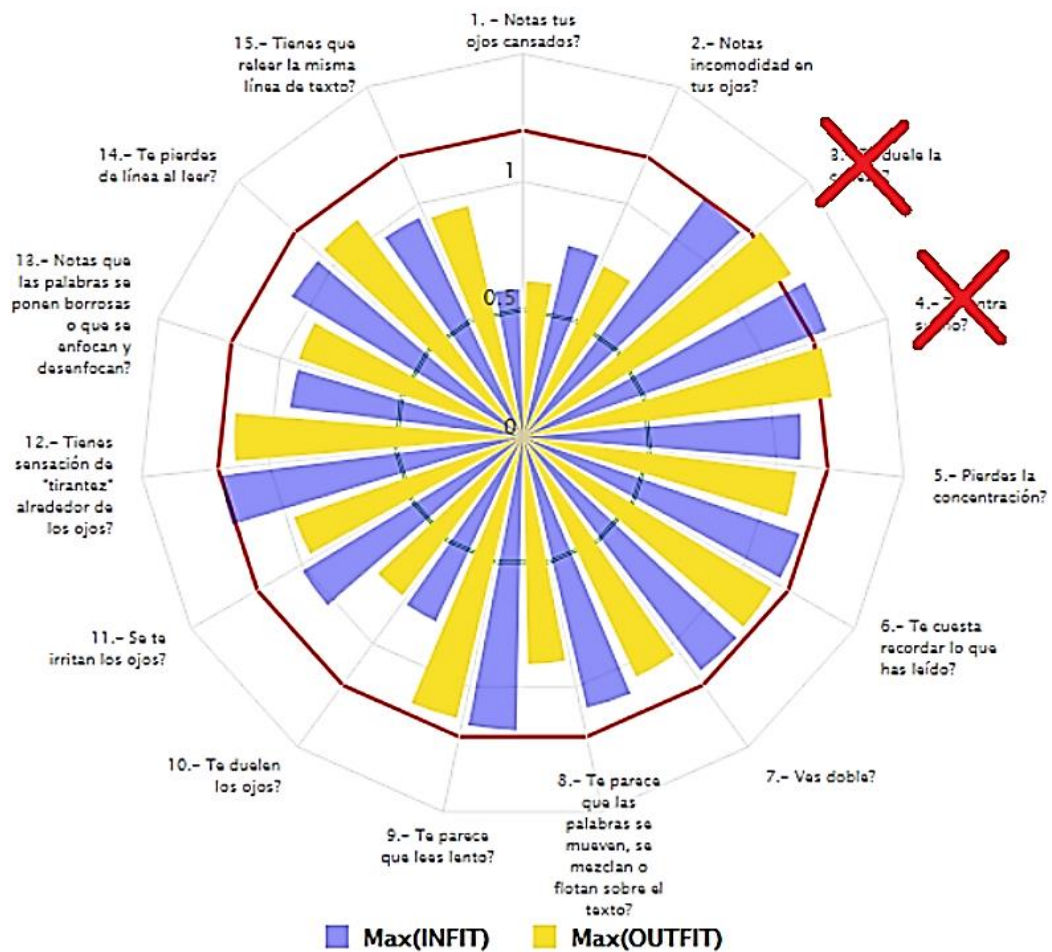


FIGURA 8.2 – INFIT Y OUTFIT DE LOS ÍTEMS DE LA ESCALA CISSVE

- 3) A continuación, se descartaron los ítems con una dificultad muy diferente ( $\pm 1,5$  logits) a la capacidad media de la muestra en estudio ( $-1.44$  logits). En la Figura 8.3 se muestran los ítems descartados con una cruz roja y los que ya estaban descartados de apartados anteriores con una cruz amarilla. Este criterio descartó a los ítems 7, 8, 9, 10, 11, 12, 13, 14, 15.



FIGURA 8.3 – DIFICULTAD (SUSCEPTIBILIDAD A LA SINTOMATOLOGÍA) DE LOS ÍTEMS DE LA ESCALA CISSVE

- 4) A continuación, se examinaron los ítems con más de un 50% de datos erróneos o no completados. En este caso, no se encontró ninguno, tal y como se puede apreciar en la Figura 8.4, ya que, en todos, el porcentaje de respuestas nulas es inferior al 2%.

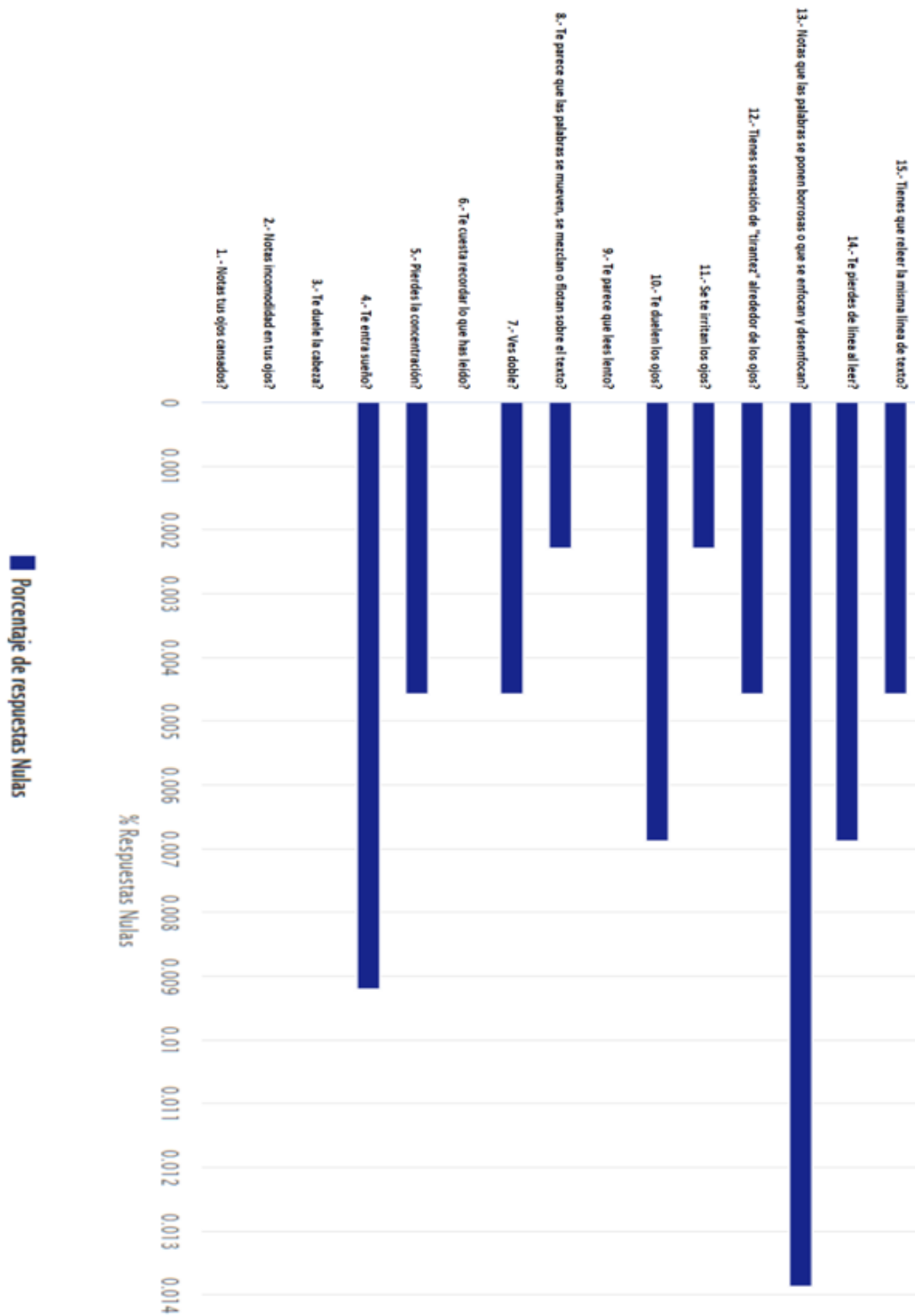


FIGURA 8.4 – RESPUESTAS NULAS EN LOS ÍTEMS DE LA ESCALA CISSVE

5) Seguidamente se examinó si había algún ítem con una elevada proporción de respuestas (>50%) en la categoría que implicaba una mayor dificultad (categoría 4 – “siempre”). Como se puede comprobar en la Figura 8.5, la proporción en la categoría más sintomática no alcanzaba esa cifra en ninguno de los 15 ítems, siendo el ítem 4 el que más respuestas acumula en esa categoría, llegando sólo a un 7%.

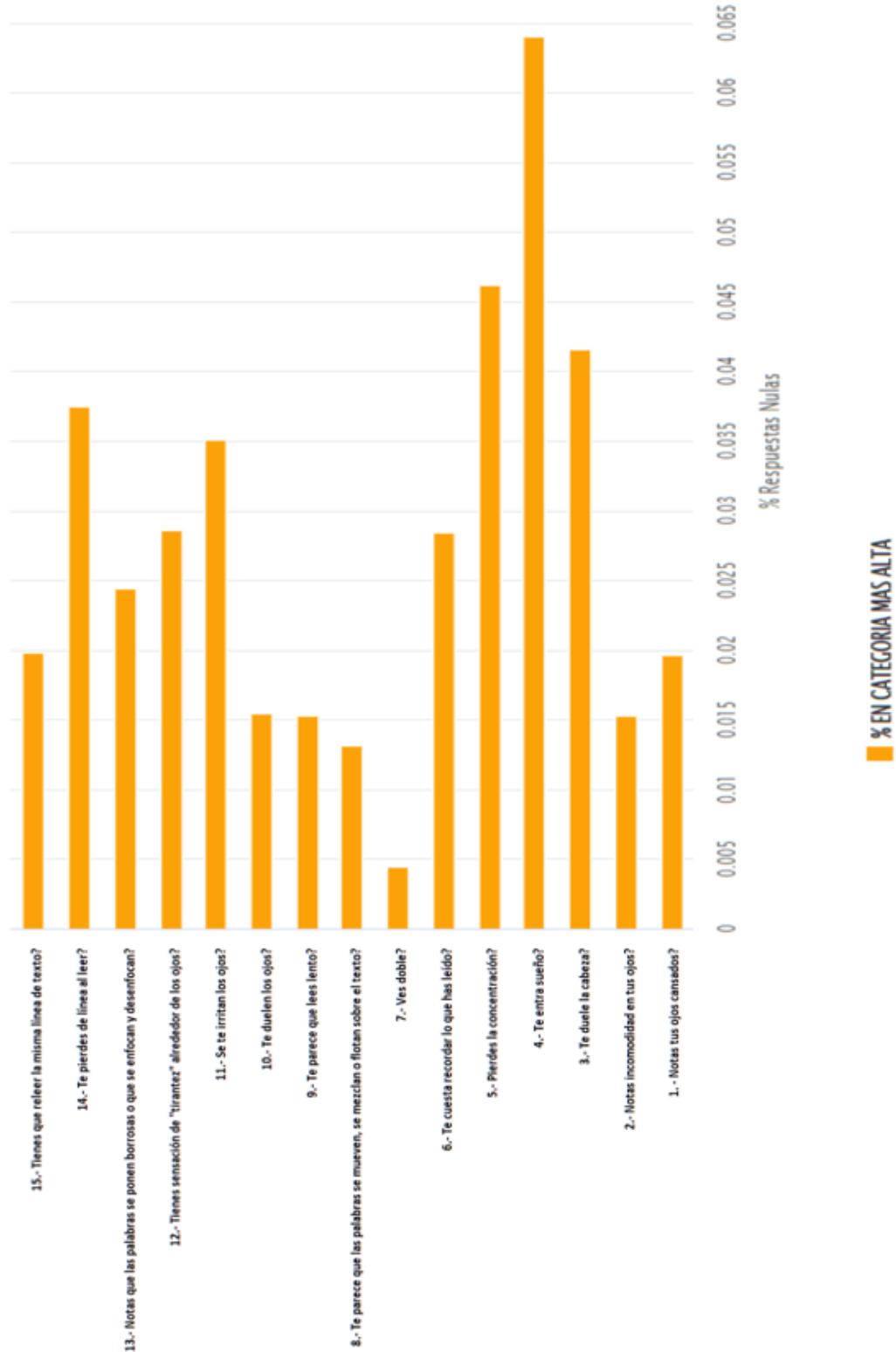


FIGURA 8.5 – PORCENTAJE DE ÍTEMS DE LA ESCALA CISSVe UBICADOS EN LA CATEGORÍA MÁS ALTA

6) A continuación, se observó si existían ítems con una desviación estándar de la puntuación media muy diferente a la de otros ítems para eliminarlos en el caso de que esta situación se diera. Como se observa en la Figura 8.6, sólo los ítems 7 y 4 quedarían excluidos por este criterio, con lo que, al haber sido excluidos ya anteriormente, este criterio no aporta ninguna selección adicional.

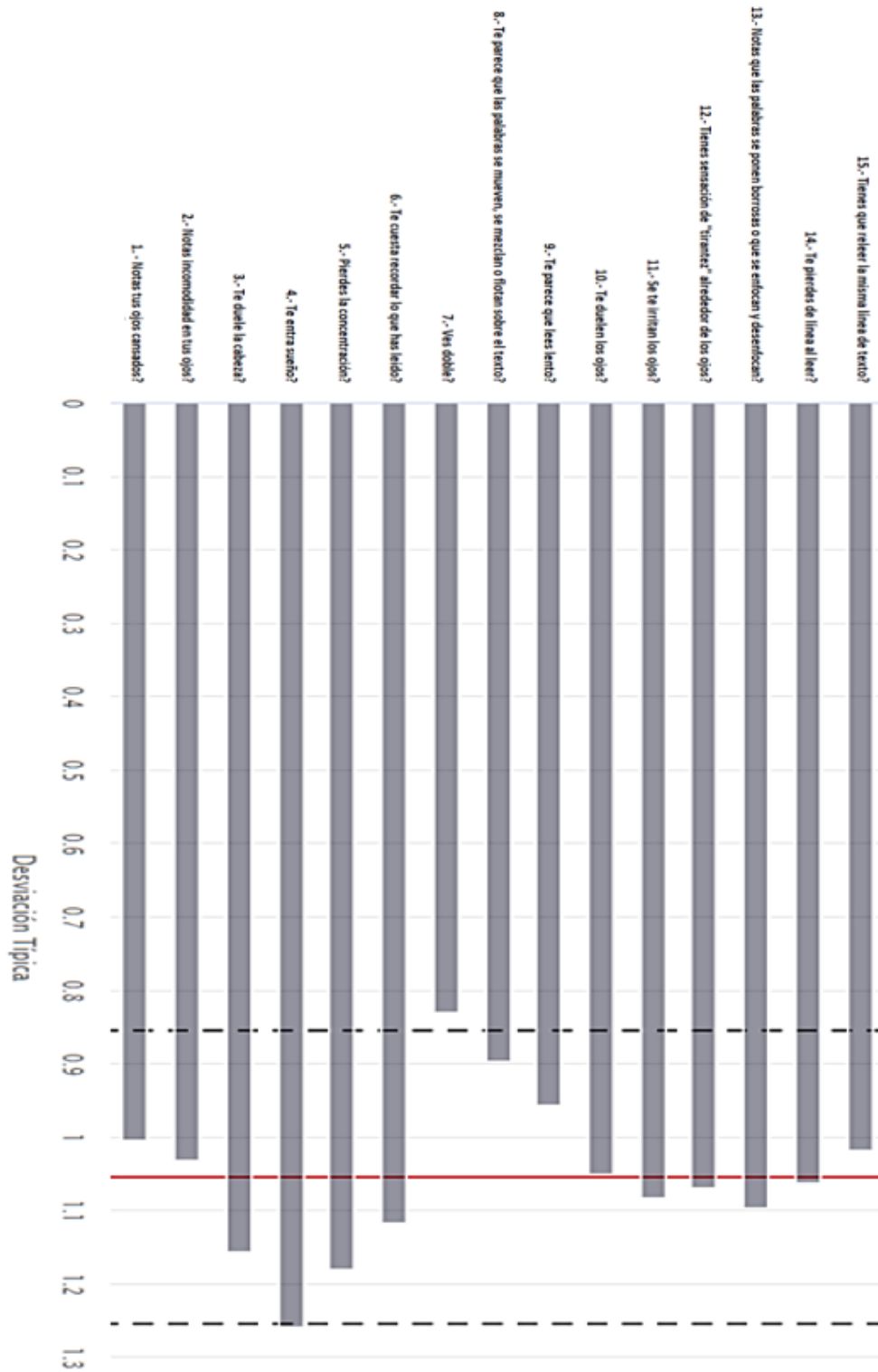


FIGURA 8.6 – DESVIACIÓN ESTÁNDAR DE LOS ÍTEMES DE LA ESCALA CISSVE

7) Por último, se examinó qué ítems presentaban un patrón de respuesta lejos de la normalidad. Se consideran para ser eliminados por este criterio los ítems con coeficientes de asimetría y curtosis fuera del intervalo (-2, +2). Como se puede observar en el gráfico de la Figura 8.7, solamente los valores de curtosis de los ítems 7 y 8 se salen de los márgenes establecidos, con lo que como ya habían sido eliminados, este criterio no nos aportó descartes adicionales.

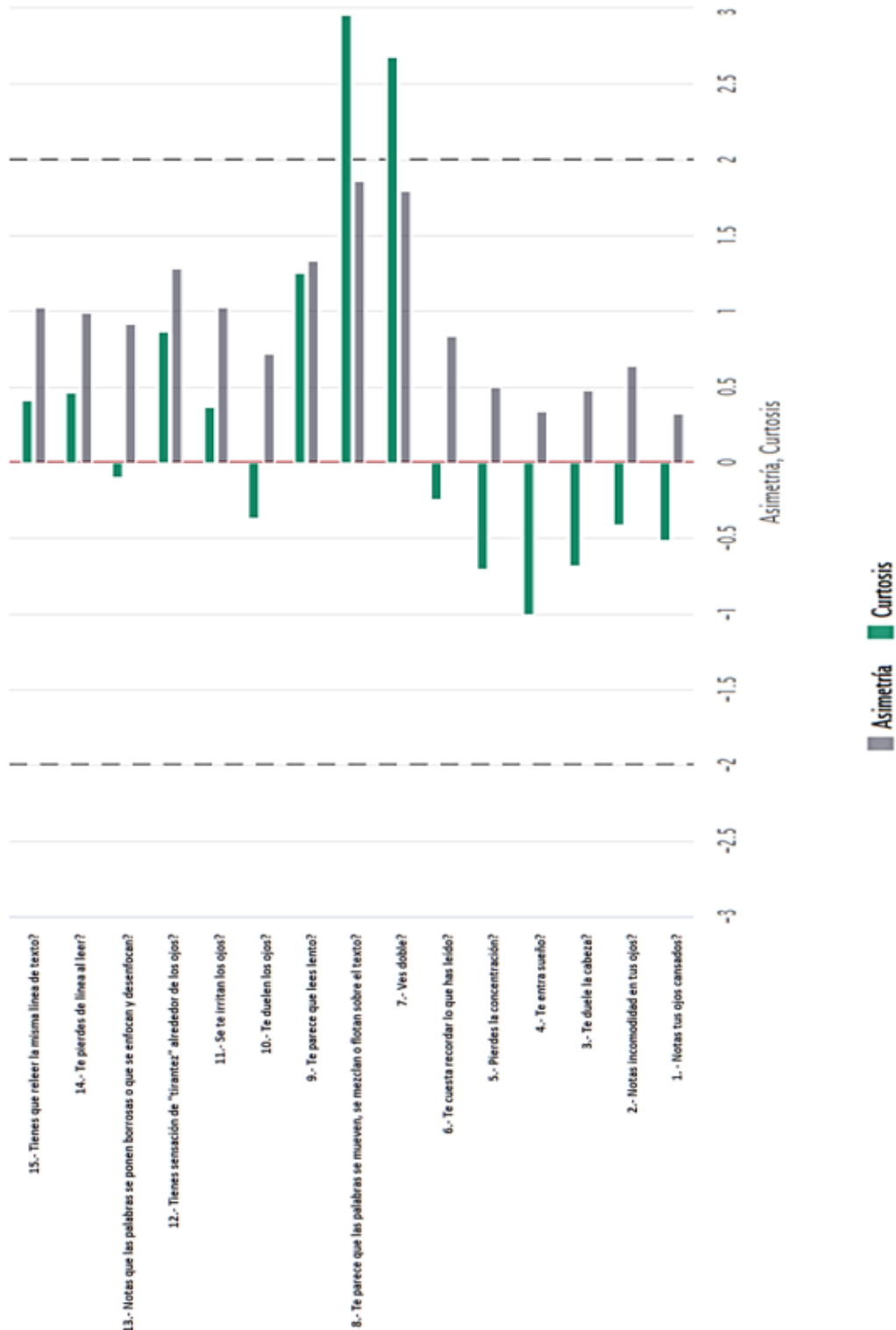


FIGURA 8.7 – ASIMETRÍA Y CURTOSIS DE LOS ÍTEMES DE LA ESCALA CISSVE

Llegados a este punto, y tras haber realizado las exclusiones pertinentes de acuerdo con los criterios explicados, se examinó cómo se comportaba la escala, aplicando con el soporte de Winsteps, el análisis de Rasch a la versión reducida, CISSve-R5 ya solamente compuesta por los ítems, 1,2,5,6 y 10.

Rápidamente se pudieron comprobar las características de la escala CISSve-R5, frente a las de la original. En la Figura 8.8, observamos cómo la escala resultante conserva las propiedades de la original en cuanto al targeting, aunque a nivel de fiabilidad pasa de un 83% a un 72%, lo cual representa un resultado lógico, al haber reducido los ítems a la tercera parte (de 15 a 5).

Person	458	INPUT	439	MEASURED		INFIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE		IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	15.1	14.9	-1.44	.42		1.00	.0	.99	.0
P.SD	10.0	.3	1.18	.25		.45	1.3	.47	1.21
REAL RMSE	.48	TRUE SD	1.07	SEPARATION	2.22	Person	RELIABILITY	.83	

Item	23	INPUT	15	MEASURED		INFIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE		IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	443.2	437.3	.00	.06		1.02	.1	.99	-.2
P.SD	130.7	1.7	.50	.01		.18	2.9	.18	2.5
REAL RMSE	.06	TRUE SD	.50	SEPARATION	7.69	Item	RELIABILITY	.98	

Person	458	INPUT	439	MEASURED		INFIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE		IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	6.0	5.0	-1.50	.77		.98	-.1	.98	-.1
P.SD	4.0	.1	1.58	.33		.68	1.2	.69	1.21
REAL RMSE	.84	TRUE SD	1.34	SEPARATION	1.59	Person	RELIABILITY	.72	

Item	23	INPUT	5	MEASURED		INFIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE		IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	523.2	438.0	.00	.07		1.00	-.3	.98	-.5
P.SD	68.7	1.3	.30	.00		.23	3.5	.21	3.1
REAL RMSE	.07	TRUE SD	.29	SEPARATION	4.16	Item	RELIABILITY	.95	

FIGURA 8.8 – “TARGETING” Y “FIABILIDAD” DE LA ESCALA CISSVE FRENTE A LA ESCALA CISSVE-R5

Dando un paso más para entender las propiedades de esta nueva escala, extrajimos una tabla para mostrar el ajuste al modelo de los ítems componentes de la nueva escala, y su situación en el mapa de Wright. Observando la Figura 8.9, podemos ver cómo los ítems 5 y 6 presentan valores de Infit y Outfit por encima de 1.2, el límite que habíamos considerado en el segundo criterio para realizar los descartes en base al ajuste al modelo. Además, el ítem 6 tiene una medida sólo 0.07 logits por encima del ítem 10, con lo que su aporte en la escala es muy reducido.

A continuación, examinamos la dimensionalidad para ver si la versión reducida de 5 ítems sigue conservando las propiedades que tenía CISSve o el hecho de haber eliminado 10 ítems introduce algún cambio en este aspecto. Podemos observar que el comportamiento sigue siendo similar al de la escala original CISSve

Person: REAL SEP.: 1.59 REL.: .72 ... Item: REAL SEP.: 4.16 REL.: .95

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		PTMEASUR-CORR.	AL-EXP.	EXACT MATCH		Item	
					MNSQ	ZSTD	MNSQ	ZSTD			OBS%	EXP%		
5	577	437	-.24	.07	1.27	3.6	1.23	3.2	A	.70	.74	42.8	49.5	-Pierdes la concentración?
6	456	439	.30	.07	1.23	3.1	1.20	2.5	B	.67	.70	48.5	52.3	-Te cuesta recordar lo que has leído?
10	468	436	.23	.07	1.00	.1	.98	-.3	C	.71	.71	50.1	51.9	-Te duelen los ojos?
2	484	439	.17	.07	.83	-2.5	.82	-2.7	b	.75	.71	54.6	49.9	-Notas incomodidad en tus ojos?
1	631	439	-.46	.06	.66	-5.7	.69	-5.1	a	.79	.75	57.1	47.6	-Notas tus ojos cansados?
MEAN	523.2	438.0	.00	.07	1.00	-.3	.98	-.5				50.6	50.2	
P.SD	68.7	1.3	.30	.00	.23	3.5	.21	3.1				5.0	1.7	

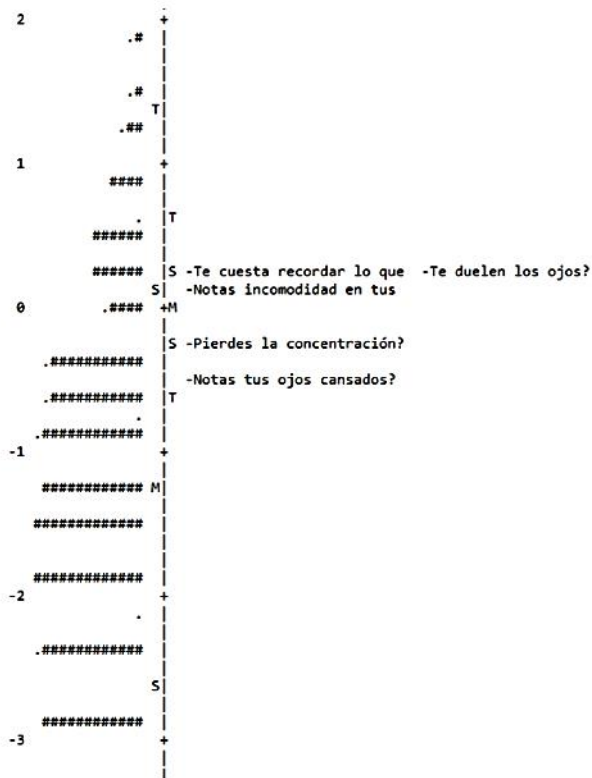


FIGURA 8.9 – ANÁLISIS DE LOS ÍTEMS DE LA ESCALA CISSVE-R5

Observando el resultado del análisis de las componentes principales que nos proporciona Winsteps, observamos que la varianza explicada por las medidas en la versión de 5 ítems (52%) es similar a la versión de 15 ítems (45,4%), y que sigue apareciendo la segunda dimensión, al igual que ocurría en CISSVe. Estos resultados se observan en la Figura 8.10

TABLE 23.0 AN-LISIS DEL CISS ZOU876WS.TXT May 13 2020 15:10  
 INPUT: 458 Person 23 Item REPORTED: 439 Person 15 Item 5 CATS WINSTEPS 4.1.0

---

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

	Eigenvalue	Observed	Expected
Total raw variance in observations =	27.4956	100.0%	100.0%
Raw variance explained by measures =	12.4956	45.4%	46.1%
Raw variance explained by persons =	5.0550	18.4%	18.7%
Raw Variance explained by items =	7.4406	27.1%	27.5%
Raw unexplained variance (total) =	15.0000	54.6%	100.0%
Unexplned variance in 1st contrast =	2.2082	8.0%	14.7%
Unexplned variance in 2nd contrast =	1.6982	6.2%	11.3%
Unexplned variance in 3rd contrast =	1.4748	5.4%	9.8%
Unexplned variance in 4th contrast =	1.2205	4.4%	8.1%
Unexplned variance in 5th contrast =	1.1059	4.0%	7.4%

TABLE 23.0 AN-LISIS DEL CISS ZOU415WS.TXT Jun 7 2020 22:24  
 INPUT: 458 Person 23 Item REPORTED: 439 Person 5 Item 5 CATS WINSTEPS 4.1.0

---

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

	Eigenvalue	Observed	Expected
Total raw variance in observations =	10.5129	100.0%	100.0%
Raw variance explained by measures =	5.5129	52.4%	52.3%
Raw variance explained by persons =	3.3169	31.6%	31.5%
Raw Variance explained by items =	2.1961	20.9%	20.8%
Raw unexplained variance (total) =	5.0000	47.6%	100.0%
Unexplned variance in 1st contrast =	2.0887	19.9%	41.8%
Unexplned variance in 2nd contrast =	1.0230	9.7%	20.5%
Unexplned variance in 3rd contrast =	.9860	9.4%	19.7%
Unexplned variance in 4th contrast =	.9019	8.6%	18.0%
Unexplned variance in 5th contrast =	.0036	.0%	.1%

FIGURA 8.10 – DIMENSIONALIDAD DE LA ESCALA CISSVE-R5

A continuación, revisitamos el análisis factorial, que realizamos en el apartado 6.3.3.3.7 y en el que encontrábamos que efectivamente tanto el ítem 5 cómo el ítem 6 pertenecían a la segunda dimensión, la relacionada con la concentración en la lectura. En la Figura 8.11 tenemos el detalle de los 3 factores se encontraron para CISSve y a su lado los que se encuentran en CISSve-R5, que como se ve son totalmente concordantes pues los ítems 1, 10 y 2 en CISSve-R5 siguen perteneciendo al primer factor y los ítems 6 y 5 al segundo.

	Componente		
	1	2	3
ITEM1	,773		
ITEM2	,753		
ITEM10	,714		
ITEM11	,691		
ITEM12	,586		
ITEM3	,567		
ITEM5		,763	
ITEM6		,718	
ITEM4		,642	
ITEM15		,570	,512
ITEM8			,672
ITEM7			,657
ITEM14			,612
ITEM13	,545		,557
ITEM9			,546

	Componente	
	1	2
ITEM1	,844	
ITEM10	,839	
ITEM2	,833	
ITEM6		,859
ITEM5		,851

FIGURA 8.11 – FACTORES DE LA ESCALA CISSVE (IZQUIERDA) Y CISSVE-R5 (DERECHA)

Por lo tanto, y llegados a este punto, el análisis nos sugiere que debemos descartar los ítems 5 y 6, pero optamos por hacerlo secuencialmente en vez de hacerlo a la vez para observar los cambios en dos pasos, de forma que podamos observar cómo afecta la desaparición de cada uno de los ítems a las propiedades de la escala. Tiene sentido realizar esta eliminación pues recordemos que el objetivo primario de la escala reducida era el construir un instrumento capaz de realizar exámenes rápidos a modo de “screening” a una población general y así poder derivarlos a la clínica, y para estos ítems, los valores de infit y outfit observados infieren que la discrepancia entre las respuestas reales de los participantes a esos ítems y las expectativas del modelo en función de las habilidades de los participantes son mayores de lo que cabría esperar.

En primer lugar, iterando sobre el proceso anterior, descartamos el ítem 6, ya que como se puede observar en la Figura 8.9, es el que peor se ajusta al modelo, y adicionalmente sabemos que pertenece a la segunda dimensión y que su dificultad es muy similar a la del ítem 10, siendo entre todos los ítems la más des posicionada con respecto a la media de la muestra, y por lo tanto es el ítem que aporta un valor más impreciso a la hora de medir la muestra seleccionada.

En la Figura 8.12 - “Targeting” y “Fiabilidad” de la Escala CISSve-R5 (arriba) frente a la escala CISSve-R4 podemos observar cómo cambia la escala en cuanto al “targeting” y a la fiabilidad, y como se puede comprobar, se cumplen las predicciones que habíamos realizado, encontrando que en la versión de 4 ítems, CISSve-R4, las propiedades se conservan intactas.

Person		458 INPUT		439 MEASURED		INFIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	6.0	5.0	-1.50	.77	.98	-.1	.98	-.1	
P.SD	4.0	.1	1.50	.33	.68	1.2	.69	1.21	
REAL RMSE	.84	TRUE SD	1.34	SEPARATION	1.59	Person RELIABILITY		.72	

Item		23 INPUT		5 MEASURED		INFIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	523.2	438.0	.00	.07	1.00	-.3	.98	-.5	
P.SD	68.7	1.3	.30	.00	.23	3.5	.21	3.1	
REAL RMSE	.07	TRUE SD	.29	SEPARATION	4.16	Item RELIABILITY	.95		

Person		458 INPUT		439 MEASURED		INFIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	4.9	4.0	-1.59	.90	.98	-.1	.98	-.1	
P.SD	3.3	.1	1.78	.34	.77	1.2	.77	1.21	
REAL RMSE	.96	TRUE SD	1.50	SEPARATION	1.55	Person RELIABILITY		.71	

Item		23 INPUT		4 MEASURED		INFIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	540.0	437.8	.00	.08	.99	-.7	.98	-.8	
P.SD	67.0	1.3	.33	.01	.37	5.1	.35	4.8	
REAL RMSE	.08	TRUE SD	.32	SEPARATION	4.25	Item RELIABILITY	.95		

FIGURA 8.12 - "TARGETING" Y "FIABILIDAD" DE LA ESCALA CISSVE-R5 (ARRIBA) FRENTE A LA ESCALA CISSVE-R4 (ABAJO)

En la Figura 8.13, podemos observar cómo al redistribuirse respecto a la muestra, los ítems que permanecen en la escala CISSve-R4, se mueve ligeramente el "0 de escala", que está situado en el punto medio de dificultad de los ítems, de forma que se acentúan las observaciones que habíamos visto anteriormente.

Los valores de ajuste del ítem 5 están fuera del intervalo recomendado, ya que muestran valores de Infit y Outfit por encima de 1.5. Algo que sí tiene de positivo este ítem, es el valor de su dificultad, que al ser uno de los más cercanos a la media de los sujetos, hace que contribuya positivamente a mejorar el valor de targeting de la escala. No obstante, aunque el eliminarlo presumiblemente nos empeorará ligeramente el valor del targeting, consideramos apropiado hacerlo, al haber comprobado que se refiere un factor distinto al principal que compone nuestro constructo actual que permite hacer "screenings" rápidos en poblaciones generales (síntomas asociados a la fatiga visual), y aún más habiendo observado que su ajuste al modelo no es aceptable de acuerdo con el criterio 2 que vimos anteriormente.

APORTACIONES ADICIONALES DERIVADAS DE LA TESIS

Person: REAL SEP.: 1.55 REL.: .71 ... Item: REAL SEP.: 4.25 REL.: .95

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFINIT MNSQ	INFINIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
5	577	437	<b>-.19</b>	.07	<b>1.60</b>	<b>7.4</b>	<b>1.56</b>	<b>7.0</b>	A .67	.77	38.6	51.1	-Pierdes la concentración?
10	468	436	.33	.07	.96	-1.6	.93	-1.9	B .77	.74	54.5	53.6	-Te duelen los ojos?
2	484	439	.29	.07	.81	-2.9	.80	-3.0	b .79	.75	60.8	53.6	-Notas incomodidad en tus ojos
1	631	439	-.45	.07	.62	-6.5	.64	-6.1	a .84	.78	61.3	50.7	-Notas tus ojos cansados?
MEAN	540.0	437.8	.00	.07	.99	-.7	.98	-.8			53.8	52.3	
P.SD	67.0	1.3	.33	.00	.37	5.1	.35	4.8			9.2	1.4	

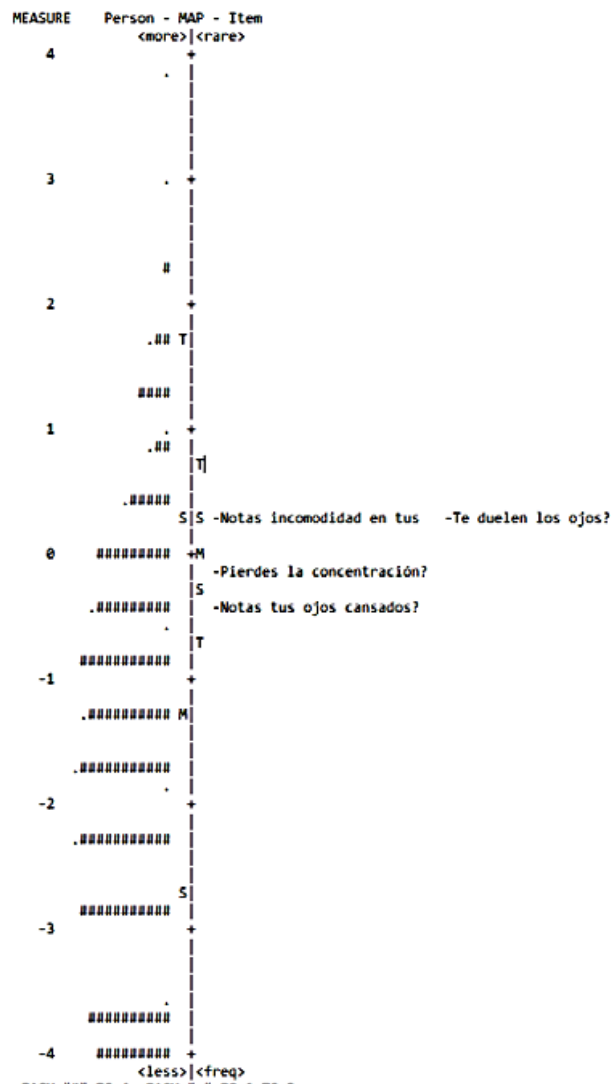


FIGURA 8.13 – ANÁLISIS DE LOS ÍTEMS DE LA ESCALA CISSVE-R4

Antes de realizar el descarte, se comprueba cómo se comporta la escala CISSve-R4 en cuanto a la dimensionalidad, y en la Figura 8.14 podemos comprobar cómo ahora ya no aparece la segunda dimensión, circunstancia que es lógica, al haber aislado a un solo ítem los componentes de esa dimensión. No obstante, es de esperar que al descartar el ítem 5, la



Person		458	INPUT	439	MEASURED	INFIT		OUTFIT	
	TOTAL		COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	4.9		4.0	-1.50	.98	.98	-1	.98	-1
P.SD	3.3		.1	1.70	.34	.77	1.2	.77	1.2
REAL RMSE	.96	TRUE SD	1.50	SEPARATION	1.55	Person	RELIABILITY	.71	

Item		23	INPUT	4	MEASURED	INFIT		OUTFIT	
	TOTAL		COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	540.0		437.8	.00	.08	.99	-.7	.98	-.8
P.SD	67.0		1.3	.33	.01	.37	5.1	.35	4.8
REAL RMSE	.08	TRUE SD	.32	SEPARATION	4.25	Item	RELIABILITY	.95	

Person		458	INPUT	439	MEASURED	INFIT		OUTFIT	
	TOTAL		COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	3.6		3.0	-2.71	1.29	.98	-1	.98	-1
P.SD	2.7		.1	2.92	.37	1.02	1.1	1.02	1.2
REAL RMSE	1.35	TRUE SD	2.59	SEPARATION	1.92	Person	RELIABILITY	.79	

Item		23	INPUT	3	MEASURED	INFIT		OUTFIT	
	TOTAL		COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	527.7		438.0	.00	.09	1.00	-1	.98	-.3
P.SD	73.4		1.4	.60	.00	.14	2.0	.11	1.7
REAL RMSE	.10	TRUE SD	.59	SEPARATION	6.26	Item	RELIABILITY	.98	

FIGURA 8.15 - "TARGETING" Y "FIABILIDAD" DE LA ESCALA CISSVE-R4 (ARRIBA) FRENTE A LA ESCALA CISSVE-R3 (ABAJO)

De forma más detallada, en la Figura 8.16 podemos observar los parámetros de ajuste al modelo de la escala CISSve-R3 y su posición en el mapa de Wright. Los resultados son coherentes con lo esperado y refuerzan la idea de haber logrado una escala razonablemente buena.

Haciendo un análisis pormenorizado, podemos comprobar que el ítem 10 está muy próximo al ítem 2 en nivel de dificultad, encontrándose ligeramente por encima y por tanto más alejado de la media de los sujetos.

Por ello, y si tenemos en cuenta, además que sus valores de Infit y Outfit, aun siendo aceptables, están lejos de ser los ideales, concluimos que no está aportando demasiado valor en la escala, con lo que nos planteamos que quizá el eliminarlo, no sólo no se dañarían las propiedades de la escala CISSve-R3, sino que probablemente mejorarían.

APORTACIONES ADICIONALES DERIVADAS DE LA TESIS

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
10	468	436	.47	.09	1.15	2.1	1.11	1.5	.85	.86	61.0	62.1	-Te duelen los ojos?
2	484	439	.37	.09	1.02	-.4	1.00	-.0	.86	.86	63.4	62.6	-Notas incomodidad en tus ojos
1	631	439	-.85	.09	.82	-2.8	.83	-2.5	.88	.87	68.4	60.9	-Notas tus ojos cansados?
MEAN	527.7	438.0	.00	.09	1.00	-.1	.98	-.3			64.3	61.8	
P.SD	73.4	1.4	.60	.00	.14	2.0	.11	1.7			3.1	.7	

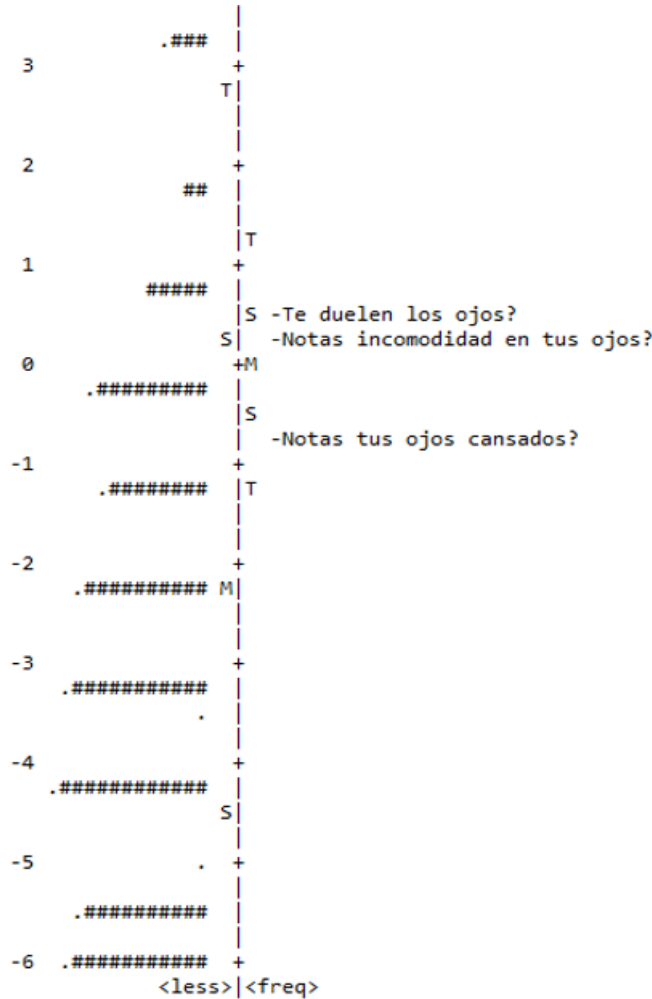


FIGURA 8.16 – ANÁLISIS DE LOS ÍTEMS DE LA ESCALA CISSVE-R3

Antes de hacer ese ejercicio y descartar el ítem 10, observamos en la Figura 8.17 el análisis de dimensionalidad de la escala de 3 ítems a través del análisis de las componentes principales que proporciona Winsteps. Ahora sí como era de esperar, tras descartar el ítem 5, la varianza explicada por las medidas ha aumentado considerablemente, y por lo tanto el porcentaje de varianza no explicada que se encuentra en el primer contraste es ya muy pequeña, encontrándose una única dimensión que explica más del 67% de la varianza del constructo.

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units			
	Eigenvalue	Observed	Expected
Total raw variance in observations =	9.0936	100.0%	100.0%
Raw variance explained by measures =	5.0936	56.0%	55.9%
Raw variance explained by persons =	3.3678	37.0%	37.0%
Raw Variance explained by items =	1.7258	19.0%	18.9%
Raw unexplained variance (total) =	4.0000	44.0%	44.1%
Unexplained variance in 1st contrast =	1.6855	18.5%	42.1%
Unexplained variance in 2nd contrast =	1.1977	13.2%	29.9%
Unexplained variance in 3rd contrast =	1.1139	12.2%	27.8%
Unexplained variance in 4th contrast =	.0020	.0%	.0%
Unexplained variance in 5th contrast =	.0009	.0%	.0%

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units			
	Eigenvalue	Observed	Expected
Total raw variance in observations =	9.3009	100.0%	100.0%
Raw variance explained by measures =	6.3009	67.7%	67.6%
Raw variance explained by persons =	4.9521	53.2%	53.1%
Raw Variance explained by items =	1.3488	14.5%	14.5%
Raw unexplained variance (total) =	3.0000	32.3%	32.4%
Unexplained variance in 1st contrast =	1.5878	17.1%	52.9%
Unexplained variance in 2nd contrast =	1.4075	15.1%	46.9%
Unexplained variance in 3rd contrast =	.0037	.0%	.1%
Unexplained variance in 4th contrast =	.0008	.0%	.0%
Unexplained variance in 5th contrast =	.0002	.0%	.0%

FIGURA 8.17 – DIMENSIONALIDAD DE LA ESCALA CISSVE-R3

Por último, al descartar el ítem 10, obtenemos una escala reducida a 2 únicos ítems, CISSve-R2. Como se había anticipado y se puede observar con detalle en la Figura 8.18, al aplicar el análisis de Rasch a CISSve-R2, y comparar sus propiedades con las de CISSve-R3, se observa una mejora de la fiabilidad respecto a la del instrumento de 3 ítems, quedando sólo 3 puntos por debajo de la escala de 15 ítems, que era de un 83%, presentándose dentro del mismo nivel de calidad que la escala original según los criterios de Khadka et al.<sup>114</sup> El targeting como es lógico, baja ligeramente, al redistribuirse de nuevo el “0 de escala” al punto medio de dificultad de los 2 ítems restantes, manteniéndose bastante cerca del modelo de 3 ítems.

En resumen, un instrumento de dos únicos ítems, CISSve-R2, sorprendentemente, consigue una fiabilidad similar a la del original, CISSve, en poblaciones con características similares a las estudiadas, circunstancia que provee un considerable valor práctico a esta escala, para utilizarla con fines de “screening” con el objetivo de identificar sujetos susceptibles de ser examinados con más rigor en un gabinete clínico.

Person									
	458	INPUT	439	MEASURED		INFINIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	3.6	3.0	-2.71	1.29	.98	-.1	.98	-.1	
P.SD	2.7	1.4	4.42	.37	1.02	1.1	1.02	1.1	
REAL RMSE	1.35	TRUE SD	2.59	SEPARATION	1.92	Person RELIABILITY	.79		

Item									
	23	INPUT	3	MEASURED		INFINIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	527.7	438.0	.00	.09	1.00	-.1	.98	-.3	
P.SD	73.4	1.4	.60	.00	.14	2.0	.11	1.7	
REAL RMSE	.10	TRUE SD	.59	SEPARATION	6.26	Item RELIABILITY	.98		

Person									
	458	INPUT	439	MEASURED		INFINIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	2.5	2.0	-3.66	1.90	.98	-.2	.98	-.2	
P.SD	1.8	1.0	4.42	.61	1.53	1.1	1.53	1.1	
REAL RMSE	2.00	TRUE SD	3.94	SEPARATION	1.98	Person RELIABILITY	.80		

Item									
	23	INPUT	2	MEASURED		INFINIT		OUTFIT	
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	557.5	439.0	.00	.11	.97	-.3	.98	-.2	
P.SD	73.5	.0	.94	.00	.01	.1	.00	.0	
REAL RMSE	.11	TRUE SD	.93	SEPARATION	8.13	Item RELIABILITY	.99		

FIGURA 8.18 - “TARGETING” Y “FIABILIDAD” DE LA ESCALA CISSVE-R3 (ARRIBA) FRENTE A LA ESCALA CISSVE-R2 (ABAJO)

Repasemos la distribución de los ítems y su ubicación frente a la muestra analizada para entender la razón por la que el targeting muestra un valor tan distante entre la media de los ítems y la media de los sujetos, y sobre todo, por qué se percibe una separación mucho mayor que en la escala de partida, CISSve, de 15 elementos, cuando en el instrumento de 2 ítems, CISSve-R2, solamente utilizamos los ítems más cercanos a la muestra en estudio y por lo tanto, en términos absolutos, la media de dificultad de los ítems, ahora está más cerca de la media de la habilidad de los sujetos.

En la Figura 8.19, podemos observar en la parte derecha las propiedades de los dos ítems, y en la parte izquierda su posición en el mapa de Wright. La letra “M” designa a la media y la letra “S” (standard deviation), a la desviación típica.

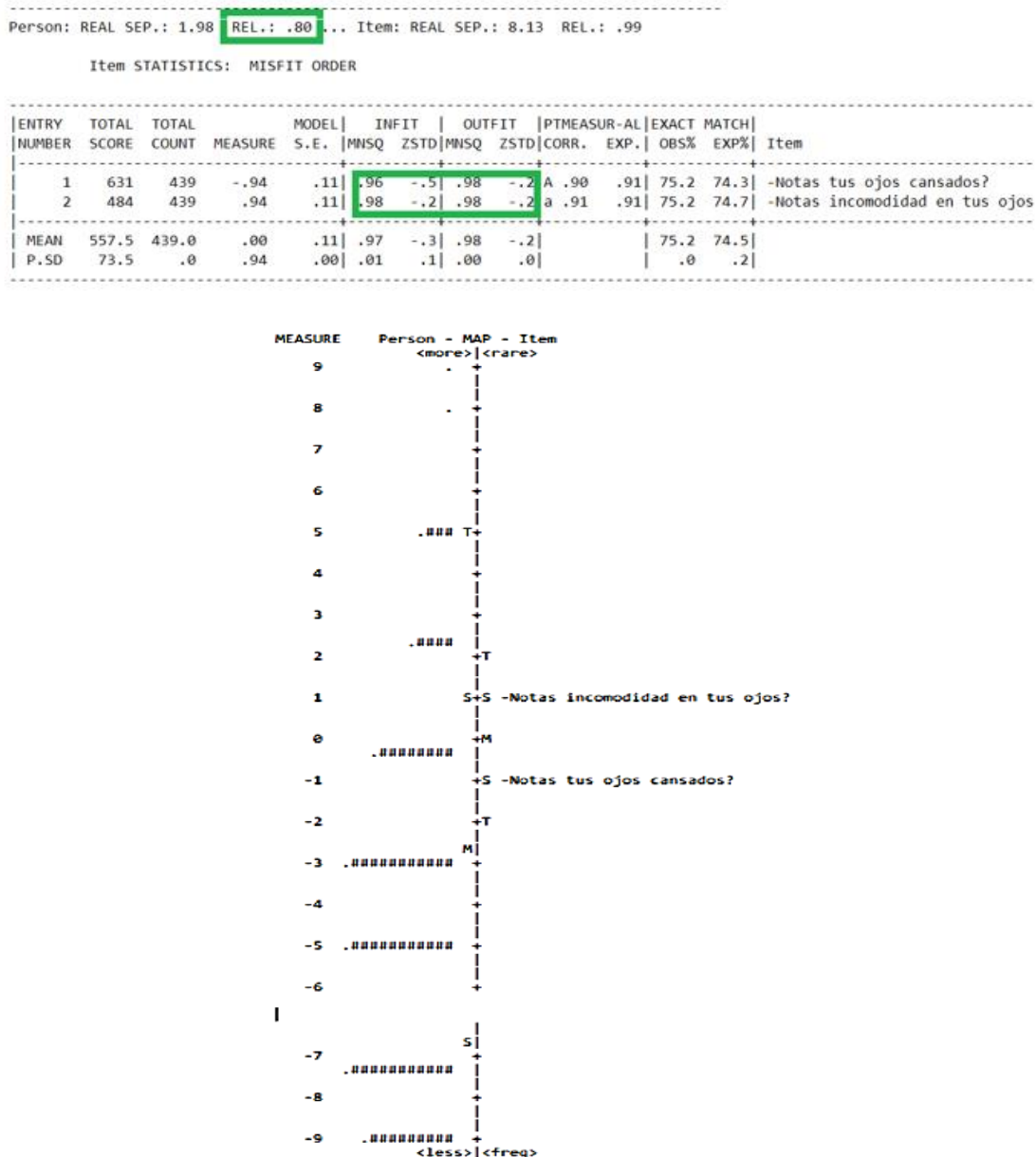


FIGURA 8.19 – ANÁLISIS DE LOS ÍTEMES DE LA ESCALA CISSVE-R2

Como se puede apreciar, según hemos ido realizando las iteraciones, hemos mantenido la muestra de estudio constante, y por lo tanto, sus valores de media y desviación típica en cuanto a parámetros de habilidad, en términos absolutos, han permanecido también constantes. Sin embargo, al hacer la transformación de las medidas a escala logit, y conservar el criterio en Winsteps en cuanto a los parámetros introducidos para hacer el re-escalado, (hacer coincidir el “0 de escala” con la dificultad media de los ítems y conservar la correspondencia entre 1 logit y el valor de la desviación típica de los ítems), cambiamos la unidad con la que se mide la distancia entre ítems y sujetos, es decir, 1 logit en el análisis de la escala de 15 ítems no es igual a un logit en la escala de 2 ítems, ya que la desviación típica de los ítems de la escala, CISSve-R2, es menor que la de la escala de 15 ítems, CISSve.

Realizando la adaptación de los parámetros en Winsteps para la escala CISSve-R2, con el fin de comparar en base a la misma unidad las características de las dos escalas podemos observar los nuevos valores de CISSve-R2, re-escalados a las unidades de la modelización de CISSve.

Como es lógico y se puede observar en la Figura 8.20, el valor de fiabilidad se mantiene en un 80%, pero, sin embargo, el parámetro del de “targeting” que mide “distancias ítem-sujeto” ha mejorado sustancialmente, y aunque como sucedía en la escala CISSve, se encuentra desplazado hacia la parte no sintomática de la muestra, al haber dejado solamente los ítems con dificultad media más cercana a la dificultad media de los sujetos, está más adaptado a la muestra.

Person									
	458 INPUT		439 MEASURED		INFIT		OUTFIT		
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	15.1	14.9	-1.44	.42	1.00	.0	.99	.0	
P. SD	10.0	.3	1.18	.25	.45	1.3	.47	1.21	
REAL RMSE	.48	TRUE SD	1.07	SEPARATION	2.22	Person RELIABILITY .83			

Item									
	23 INPUT		15 MEASURED		INFIT		OUTFIT		
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	443.2	437.3	.00	.06	1.02	.1	.99	-.2	
P. SD	130.7	1.7	.50	.01	.18	2.9	.18	2.5	
REAL RMSE	.06	TRUE SD	.50	SEPARATION	7.69	Item RELIABILITY .98			

Person									
	458 INPUT		439 MEASURED		INFIT		OUTFIT		
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	2.5	2.0	-.95	.49	.98	-.2	.98	-.2	
P. SD	1.8	.0	1.15	.16	1.53	1.1	1.53	1.1	
REAL RMSE	.52	TRUE SD	1.03	SEPARATION	1.98	Person RELIABILITY .80			

Item									
	23 INPUT		2 MEASURED		INFIT		OUTFIT		
	TOTAL	COUNT	MEASURE	REALSE	IMNSQ	ZSTD	OMNSQ	ZSTD	
MEAN	557.5	439.0	.00	.03	.97	-.3	.98	-.2	
P. SD	73.5	.0	.24	.00	.01	.1	.00	.0	
REAL RMSE	.03	TRUE SD	.24	SEPARATION	8.13	Item RELIABILITY .99			

FIGURA 8.20 – EXPLICACIÓN SOBRE EL “TARGETING” DE LA ESCALA CISSVE-R2 FRENTE A LA ESCALA CISSVE

En cuanto al comportamiento de los ítems y su situación en el mapa de Wright, podemos observar en la Figura 8.21, cómo el re-escalado simplemente, ha acercado los ítems a la misma distancia en logits que guardaban entre sí cuando se aplicó el modelo de Rasch a la escala CISSve con los 15 ítems. Es una circunstancia lógica y está de acuerdo con las propiedades de variable de intervalo que otorga el modelo de Rasch a las puntuaciones de la escala y que se explicaron en el apartado 1.2.5 de la introducción.

APORTACIONES ADICIONALES DERIVADAS DE LA TESIS

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S. E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
1	631	439	-.24	.03	.96	-.5	.98	-.2	.90	.91	75.2	74.3	-Notas tus ojos cansados?
2	484	439	.24	.03	.98	-.2	.98	-.2	.91	.91	75.2	74.7	-Notas incomodidad en tus ojos
MEAN	557.5	439.0	.00	.03	.97	-.3	.98	-.2			75.2	74.5	
P. SD	73.5	.0	.24	.00	.01	.1	.00	.0			.0	.2	

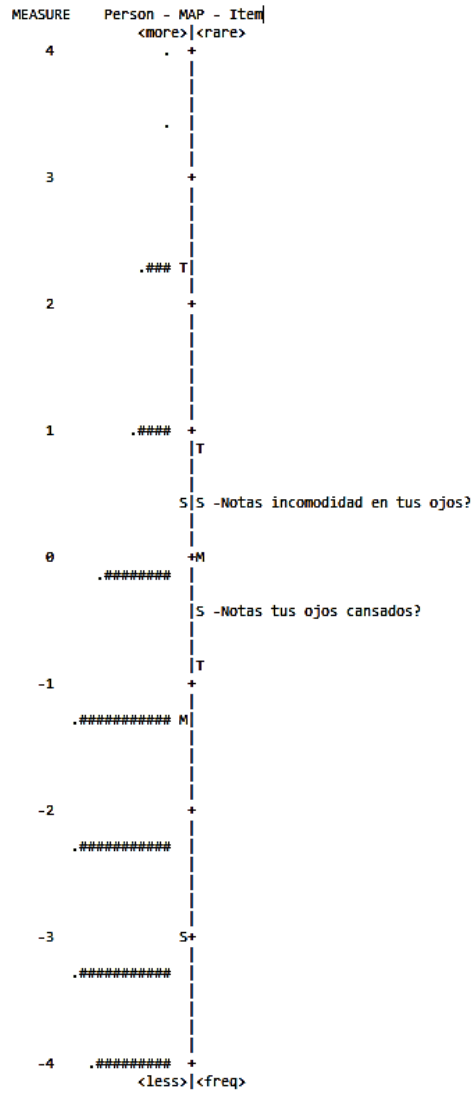


FIGURA 8.21 – ANÁLISIS DE LOS ÍTEMS DE LA ESCALA CISSVE-R2 EN LA ESCALA DE CISSVE

En la Figura 8.22, vemos cómo efectivamente la distancia entre ítems es la misma en el modelo de 15 ítems, CISSve, que en el de 2, CISSve-R2. Es decir, con independencia del resto de los ítems y de la muestra medida, la diferencia en la dificultad entre ítems, se mantiene constante, tal y como enuncia el modelo de Rasch.

ENTRY	TOTAL	TOTAL	MODEL	
NUMBER	SCORE	COUNT	MEASURE	S.E.
4	650	435	-.71	.05
3	593	439	-.53	.06
12	359	437	.27	.06
6	456	439	-.09	.06
7	202	437	1.04	.08
9	311	439	.48	.07
5	577	437	-.49	.06
8	219	438	.93	.08
14	447	436	-.07	.06
11	428	438	.01	.06
15	406	437	.09	.06
13	417	433	.03	.06
10	468	436	-.14	.06
2	484	439	-.18	.06
1	631	439	-.65	.05
MEAN	443.2	437.3	.00	.06
P.SD	130.7	1.7	.50	.01

ENTRY	TOTAL	TOTAL	MODEL	
NUMBER	SCORE	COUNT	MEASURE	S.E.
1	631	439	-.24	.03
2	484	439	.24	.03
MEAN	557.5	439.0	.00	.03
P.SD	73.5	.0	.24	.00

FIGURA 8.22 – INVARIABILIDAD DE DISTANCIA ENTRE LOS ÍTEMS CISSVE-R2 EN LA ESCALA DE CISSVE

En cuanto al estudio de la dimensionalidad, obviamente podemos comprobar que CISSve-R2, es unidimensional y explica un 77% de la varianza total con sus dos ítems tal y como se puede observar en la Figura 8.23.

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units				
	Eigenvalue	Observed	Expected	
Total raw variance in observations =	8.7058	100.0%	100.0%	
Raw variance explained by measures =	6.7058	77.0%	76.4%	
Raw variance explained by persons =	5.6955	65.4%	64.9%	
Raw Variance explained by items =	1.0102	11.6%	11.5%	
Raw unexplained variance (total) =	2.0000	23.0%	100.0%	23.6%
Unexplned variance in 1st contrast =	.0012	.0%	.1%	
Unexplned variance in 2nd contrast =	.0000	.0%	.0%	

FIGURA 8.23 – DIMENSIONALIDAD DE LA ESCALA CISSVE-R2

Para analizar el comportamiento de CISSve frente a CISSve-R2 con los datos de la muestra de estudio con los que contamos, lo primero que hacemos es revisar la puntuación de corte, es decir, la puntuación que se corresponde con el “0 de escala”. A partir de ahí, consideramos sujetos susceptibles de tener sintomatología, a los que excedan ese umbral.

Como se observa en la Figura 8.24, 4 puntos es la puntuación (“score”), que representa ese valor.

TABLE OF SAMPLE NORMS (500/100) AND FREQUENCIES CORRESPONDING TO COMPLETE TEST

SCORE	MEASURE	S.E.	NORMED S.E.	FREQUENCY	%	CUM.FREQ.	%	PERCENTILE	
0	-2.42E	.53	372	47	65	14.8	65	14.8	7
1	-1.91	.39	417	34	83	18.9	148	33.7	24
2	-1.33	.39	467	34	83	18.9	231	52.6	43
3	-.76	.38	516	33	80	18.2	311	70.8	62
4	-.08	.46	575	40	60	13.7	371	84.5	78
5	.61	.40	636	35	34	7.7	405	92.3	88
6	1.33	.49	699	43	25	5.7	430	97.9	95
7	2.06	.40	762	35	6	1.4	436	99.3	99
8	2.59E	.54	808	47	3	.7	439	100.0	99

FIGURA 8.24 – CORRESPONDENCIA ENTRE LOS PUNTOS OBTENIDOS EN LA ESCALA CISSVE-R2 Y SU VALOR EN LOGITS UNA VEZ SE HA APLICADO EL MODELO DE RASCH CON WINSTEPS

Organizando los datos de la muestra en estudio con ayuda de una hoja Excel, que se puede encontrar en el 11.4 (anexo IV), y que etiqueta a los sintomáticos según el criterio de la escala CISSve-R2 (los sujetos que exceden la puntuación de 4 puntos), y según el criterio de la escala CISSve determinado por los autores de la versión original de la escala CISSv15<sup>59,60,74</sup> (los niños que exceden 16 puntos y los adultos que exceden 21 puntos), llevamos a cabo la comparación entre los que se consideran sintomáticos para CISSve de 15 ítems y los que marca la escala reducida, CISSve-R2, de 2 ítems.

En la Figura 8.25 se puede ver gráficamente el resultado. En la ventana de arriba del diagrama se muestran en verde las coincidencias, mientras que, en las dos ventanas de abajo, se pueden observar en azul los “sintomáticos” de acuerdo a cada escala.

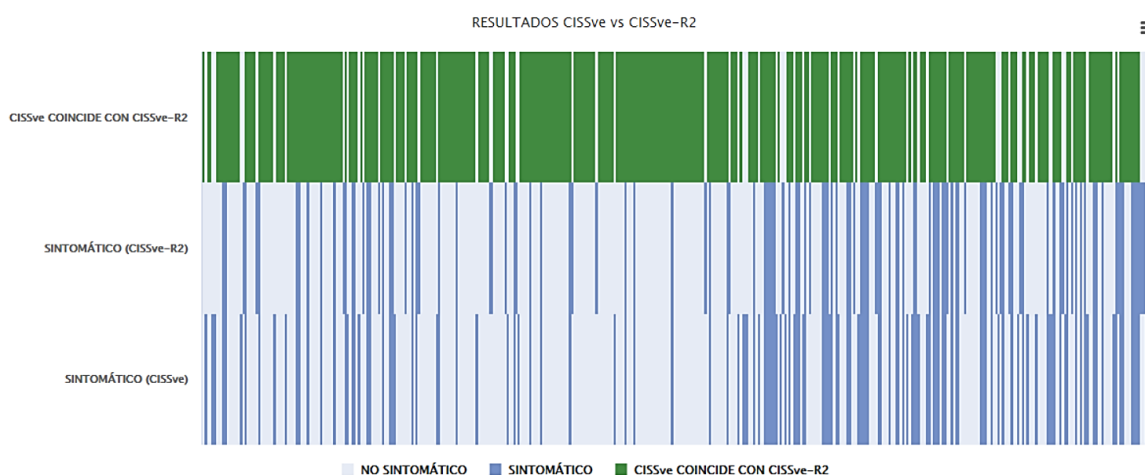


FIGURA 8.25 – SUJETOS SINTOMÁTICOS SEGÚN LA ESCALA CISSVE (ABAJO) Y CISSVE-R2 (CENTRO) Y COMPARACIÓN DIFERENCIAL ENTRE AMBOS RESULTADOS (ARRIBA)

En la Figura 8.26 se muestra de forma cuantitativa una tabla resumen con la matriz de confusión. En ella se significan, los falsos positivos, los falsos negativos y las coincidencias. Se

descartan para el análisis los sujetos que dejaron ítems sin contestar (“en blanco”), resultando un porcentaje de coincidencias de un 83.1%, sobre el total de 431 sujetos que se comparan.

		SINTOMÁTICO_2ITEMS		
SINTOMÁTICO (>16/21)	FALSO	VERDADERO	Total general	
FALSO	263	31	294	
VERDADERO	42	95	137	
<b>Total general</b>	<b>305</b>	<b>126</b>	<b>431</b>	

FIGURA 8.26 – TABLA CRUZADA QUE RELACIONA SINTOMÁTICOS Y NO SINTOMÁTICOS SEGÚN LAS ESCALAS CISSVE Y CISSVE-R2

Por último, analizamos los niveles de síntomas que es capaz de detectar CISSve-R2.

Winsteps, a través de la tabla 20.1, nos indica que este instrumento es capaz de diferenciar 3.4 niveles diferentes de síntomas de acuerdo con el método de Wright, independiente de la muestra que se considere. En la Figura 8.27 podemos ver el detalle de los cálculos de esos niveles.

Puntuación	Medida	SE	Normalizado	SE	Frecuencia	%	Frecuencia Acumulada	% Acumulado	Percentil	Mínima medida para siguiente Nivel (Indep Muestra) WRIGHT	Razón	Nivel
0	-2	0	372	47	65	2	65	14,8	7	-2,10	SE FIJA Porque	NIVEL 1
1	-2	0	417	34	83	2	148	33,7	24	-0,96	-0.96 > -1.66	NIVEL 1
2	-1	0	467	34	83	3	231	52,6	43	-0,96	R:-0.96>-1.16	NIVEL 1
3	-1	0	516	33	80	5	311	70,8	62	-0,97	R:-0.97<-0.66	NIVEL 2
4	-0	0	575	40	60	2	371	84,5	78	0,38	R:0.38>-0.07	NIVEL 2
5	0,5	0	636	35	34	4	405	92,3	88	0,30	R:0.30<0.53	NIVEL 3
6	1,2	0	699	43	25	2	430	97,9	95	1,64	R:1.64>1.16	NIVEL 3
7	1,8	0	762	35	6	5	436	99,3	99	1,52	R:1.52<1.8	NIVEL 4
8	2,3	0	808	47	3	5	439	100	27	2,97	R: 2.97>2,25	NIVEL4

FIGURA 8.27 NIVELES DE SÍNTOMAS DE LA ESCALA CISSVE-R2 ESTADÍSTICAMENTE DIFERENTES CON INDEPENDENCIA DE LA MUESTRA DE ACUERDO AL MÉTODO DE WRIGHT

En el caso de nuestra muestra de análisis, una población general, predominantemente asintomática, en la que vemos que el valor de separación entre umbrales es de 1.98 (Figura 8.20), el número de niveles de síntomas que distinguiría la escala CISSve-R2 sería  $(4*1,98)+1/3 = 2.97$ , que es mayor que 2, con lo que por lo tanto, se certificaría el hecho de que es capaz de distinguir los sujetos sintomáticos de los no sintomáticos.

### 8.1.6 DISCUSIÓN

La metodología iterativa aplicada permitió la exclusión eficiente de ítems preservando la fiabilidad del cuestionario original y manteniendo un “targeting” adecuado en el nuevo instrumento. Es importante tener en cuenta que la eliminación de ciertos ítems se realizó incluso si estos cumplían con ciertos criterios de aceptabilidad, teniendo en cuenta la consideración de que el objetivo principal era la obtención de una nueva escala apta para realizar “screenings” de forma eficiente en tiempo y coste.

### 8.1.7 CONCLUSIONES

Se ha logrado desarrollar con éxito una versión reducida de la escala CISSve que cumple con los requisitos necesarios en cuanto a fiabilidad y “targeting” para considerar al nuevo instrumento como apto para realizar “screenings”.

Este nuevo instrumento tiene el potencial de agilizar los exámenes de poblaciones generales sin comprometer significativamente la calidad psicométrica del cuestionario, ofreciendo así una herramienta más eficiente para la identificación de casos que requieren intervenciones específicas.

## 8.2 CREACIÓN DE ESCALAS PARA MEDIDA DE SÍNTOMAS ASOCIADOS A INSUFICIENCIA DE CONVERGENCIA POR GRUPOS DE EDAD

Durante el desarrollo de este trabajo, hemos podido entender cómo la sintomatología asociada a IC, puede ser disgregada en grupos de síntomas que entre sí guardan cierta asociación. El análisis factorial nos reveló que los ítems de la escala CISSve miden tres factores que aglutinan síntomas relacionados con la función visual, la fatiga asociada a trabajos prolongados de cerca y la falta de concentración. Separando la muestra de estudio por grupos de edad, niños (9-17) y adultos jóvenes (18-30), se pudo observar como la edad es un factor que influye en cómo se perciben los síntomas de IC.

Por lo tanto, puede resultar interesante que futuros trabajos de investigación se orienten a profundizar en los detalles asociados a este hallazgo. No obstante, y de cara a iniciar un enfoque hacia ese ámbito de trabajo se ha realizado una comparativa para observar cómo funcionan en niños y en adultos diferentes versiones reducidas de CISSve.

### 8.2.1 OBJETIVOS

El objetivo central de esta segunda aportación adicional es explorar la parametrización de versiones reducidas de la escala CISSve en función de la edad, con el fin de optimizar la detección de sintomatología asociada a insuficiencia de convergencia en distintos grupos etarios.

### 8.2.2 JUSTIFICACIÓN

La relevancia de esta aportación radica en la necesidad de contar con instrumentos de diagnóstico y evaluación más eficientes, especialmente adaptados a grupos de edad específicos. Esto permitiría no solo una mejor identificación de sujetos sintomáticos, sino también una optimización del tiempo de administración y del prediagnóstico.

### 8.2.3 HIPÓTESIS

La hipótesis principal postula que es factible desarrollar subescalas de CISSve parametrizadas por grupos de edad que mantengan niveles óptimos de fiabilidad y validez. La inclusión o exclusión de ciertos ítems de la escala CISSve en estas versiones reducidas varía en función del grupo etario, y esta variabilidad tiene un impacto directo en la fiabilidad del instrumento.

### 8.2.4 MÉTODOS

Se aplicó un enfoque analítico, utilizando el análisis de Rasch, para evaluar diferentes subconjuntos de ítems en versiones reducidas de la escala CISSve. Las reducciones de ítems se realizaron siguiendo un procedimiento idéntico al explicado en el apartado 8.1.4, de forma que el ajuste al modelo y la fiabilidad quedaban maximizados para cada tamaño de escala (número de ítems de la subescala). El análisis se realizó por separado para dos grupos de edad: niños (9-17) y adultos jóvenes (18-30) y los cálculos se apoyaron en el software WINSTEPS (Versión 4.0.1)<sup>107</sup>, y en el paquete de software estadístico IBM SPSS Statistics versión 25.0<sup>137</sup>.

### 8.2.5 RESULTADOS

En la Figura 8.28 se presentan algunos modelos de escala más reducidas que CISSve, parametrizados para los grupos de edad (9-17) y (18-30). En estos instrumentos, los ítems que las constituyen (en verde) han sido seleccionados aplicando el análisis de Rasch con ayuda del software Winsteps sobre la escala de partida CISSve, llevando a cabo una metodología de reducción de ítems similar a la explicada en el apartado 8.1, de forma que el ajuste al modelo y la fiabilidad quedaban maximizados para cada tamaño de escala (número de ítems de la subescala).

Por ejemplo, la última columna muestra una escala de 4 ítems optimizada para una muestra de adultos jóvenes (17-30) de una población general, o las dos columnas inmediatamente anteriores, muestran dos escalas diferentes de 3 ítems para adultos jóvenes (17-30) de una población general. La escala de la penúltima columna contendría los ítems 1, 2 y 10 de la escala CISSve, mientras que la antepenúltima columna estaría formada por los ítems 1,2 y 3 de la escala CISSve.



En el caso de los adultos, se aprecia cómo versiones con muy pocos ítems pueden alcanzar altos valores de fiabilidad, mientras que en el caso de los niños, para alcanzar valores de fiabilidad razonables es importante introducir ítems ligados al factor de “falta de concentración”. En concreto, se puede apreciar como los ítems, 5 y 6, añaden ruido en el caso de los adultos, pero hacen que el instrumento sea más fiable cuando se aplica a niños.

### 8.2.6 DISCUSIÓN

Esta aportación ofrece una fundamentación empírica para el desarrollo de futuras escalas o subescalas adaptadas a grupos etarios específicos, y pone de relieve la vital importancia que tiene el considerar la edad como variable clave en cualquier investigación dentro del ámbito de la Insuficiencia de convergencia

Como señalabamos, nuestra aportación en este trabajo se limita únicamente en abrir la línea de investigación que trate de optimizar subescalas parametrizadas por rangos de edad, de forma que con instrumentos más ligeros podamos ser más precisos separando los sujetos sintomáticos de los que no los son.

Gracias a ese enfoque podremos mejorar la precisión tanto en niños como en adultos, introduciendo o extrayendo ítems según el caso y ganando así especificidad. Además, el hecho de utilizar sólo los ítems más adecuados en cada caso ayudará a minimizar los tiempos de administración y pre-diagnóstico.

### 8.2.7 CONCLUSIONES

Los resultados sugieren que la parametrización de subescalas de CISSve adaptadas a distintos grupos de edad podría conducir a una optimización en la identificación de sintomatología asociada a IC. Este hallazgo sienta las bases para futuras investigaciones que busquen mejorar la eficiencia y especificidad de instrumentos de diagnóstico en este campo.

## 8.3 CONSTRUCCIÓN DE UN BANCO DE ÍTEMS

En el modelo de Rasch, como sabemos, los ítems se comportan de forma más precisa a la hora de medir cuando su nivel de dificultad está más próximo al nivel de habilidad de los sujetos que los contestan, o en el contexto clínico, al nivel de sintomatología que tienen los sujetos. La dificultad de un ítem no depende de la dificultad de otros ítems, ni de la muestra que se evalúa. Como repasamos en el apartado 1.2.5 de la introducción, el nivel de dificultad de un ítem puede ser comparado con el nivel de dificultad de otro y a su vez, con la habilidad de las personas que los responden, ya que tanto ítems como sujetos comparten escala, ofreciendo así la posibilidad de cuantificar la probabilidad de que un sujeto de una determinada habilidad conteste correctamente un ítem de determinada dificultad. Esto se hace posible, gracias las propiedades de medición conjunta y objetividad específica del modelo de Rasch.

### 8.3.1 OBJETIVOS

Con esta aportación se pretende comprobar la aplicabilidad práctica del modelo de Rasch en la construcción de un amplio banco de ítems adaptativos para evaluar la Insuficiencia de convergencia. Se enfoca específicamente en determinar el nivel de dificultad de nuevos ítems adicionados experimentalmente y en valorar su ajuste al modelo en comparación con los ítems preexistentes de la escala CISSve.

### 8.3.2 JUSTIFICACIÓN

La integración de nuevos ítems en una escala existente ofrece el potencial de mejorar la precisión del instrumento y adaptarlo a diferentes poblaciones objetivo. Este abordaje se fundamenta en las propiedades de medición conjunta y objetividad específica que caracterizan al modelo de Rasch.

### 8.3.3 HIPÓTESIS

Los ítems adicionales, una vez integrados en la escala CISSve, permitirán mejorar el 'targeting' y la precisión del instrumento para distintas poblaciones.

### 8.3.4 MÉTODOS

Se añadieron 8 ítems experimentales a la escala CISSve (anexo VI) durante las sesiones de toma de muestras. Posteriormente, se evaluó su nivel de dificultad y se posicionaron en la escala en unidades "logit" utilizando modelo de Rasch. Se aplicaron análisis factoriales utilizando el paquete de IBM SPSS Statistics versión 25.0 <sup>137</sup> y se representaron mapas de Wright para evaluar la compatibilidad de los nuevos ítems con los existentes y con las características de la muestra utilizando el paquete de software WINSTEPS (Versión 4.0.1) <sup>107</sup>.

### 8.3.5 RESULTADOS

Los nuevos ítems se ajustaron al modelo de Rasch en diferentes grados de calidad. No se observó una mejora significativa en las propiedades de la escala CISSve. Además, los ítems adicionales se ajustaron a factores previamente identificados, como fatiga visual, disfunción visual y problemas de concentración.

En la Figura 8.30 se puede observar una tabla que cuantifica los niveles de dificultad de los 8 nuevos ítems introducidos y sus características de ajuste al modelo de Rasch. Se puede apreciar, tanto la posición en cuanto al nivel de dificultad de cada uno de ellos respecto a los demás ítems de la escala CISSVe, como su calidad de ajuste al modelo, que como se comprueba, no es la misma en unos que en otros. En todos los casos, las medidas aportadas por los ítems adicionales se situaron dentro del rango que ya medía la escala CISSVe y no se encontró que ninguno mejorara significativamente las propiedades de la escala inicial.

INPUT: 458 Person 23 Item REPORTED: 439 Person 23 Item 5 CAIS WINSTEPS 4.1,0  
 Person: REAL SEP.: 2.48 REL.: .86 ... Item: REAL SEP.: 6.94 REL.: .98

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT [MISQ]	OUTFIT [MISQ]	PTMEASUR-AL	EXACT MATCH	Item				
7	202	437	.90	.07	1.07	.8	.95	-.44	.42	61.7	62.8	-Ves doble?	
8	219	438	.81	.07	1.00	.0	.79	-1.9	.52	.43	64.8	60.5	-Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
16	239	438	.72	.07	1.36	3.8	1.13	1.2	.43	.44	58.3	58.6	-Tienes la sensación de puente bizco?
9	311	439	.42	.06	1.01	-.2	1.03	.3	.48	.48	51.7	51.2	-Te parece que lees lento?
21	328	439	.35	.06	1.20	2.4	1.18	1.8	.46	.49	47.6	49.9	-Guitas los ojos?
17	351	438	.26	.06	.79	-3.1	.74	-2.9	.59	.50	50.7	47.6	-A veces se te emborronan las letras?
12	359	437	.24	.06	1.05	.6	1.06	.7	.51	.51	48.5	47.5	-Tienes sensación de "tirantez" alrededor de los ojos?
15	406	437	.08	.06	.84	-2.4	.85	-1.7	.56	.53	49.9	45.2	-Tienes que releer la misma línea de texto?
13	417	433	.03	.06	.80	-3.1	.80	-2.5	.61	.53	49.2	44.6	-Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
11	428	438	.02	.06	.89	-1.6	.85	-1.8	.57	.53	51.6	44.4	-Se te irritan los ojos?
18	442	437	-.03	.05	1.40	5.2	1.39	4.1	.44	.54	35.0	43.0	-Señalas con el dedo (o algo similar) la palabra que estas leyendo?
14	447	436	-.05	.05	.89	-1.6	.95	-.5	.53	.54	49.3	43.0	-Te pierdes de línea al leer?
6	456	439	-.06	.05	1.03	.4	1.05	.6	.53	.54	46.2	42.8	-Te cuesta recordar lo que has leído?
10	468	436	-.11	.05	.71	-4.8	.69	-4.3	.63	.55	45.6	42.6	-Te duelen los ojos?
2	484	439	-.14	.05	.70	-4.9	.68	-4.3	.62	.55	49.9	42.0	-Notas incomodidad en tus ojos?
23	509	439	-.21	.05	.92	-1.2	.89	-1.4	.62	.56	44.8	41.7	-Tras un tiempo leyendo o trabajando de cerca, ¿tienes que esforzarte para ver bien?
5	577	437	-.40	.05	.98	-3.1	1.00	.0	.57	.58	42.2	40.0	-Tienes que concentrarte?
19	577	437	-.40	.05	1.70	9.0	1.72	7.9	.45	.58	29.2	40.1	-Inclinas la cabeza hacia un lado?
20	586	439	-.42	.05	1.28	4.1	1.24	3.0	.50	.58	38.1	40.1	-Te acercas mucho a lo que estás mirando?
3	593	439	-.44	.05	1.04	.6	1.15	1.9	.53	.59	37.4	40.0	-Te duele la cabeza?
22	598	439	-.45	.05	1.18	2.6	1.11	1.5	.55	.59	37.0	39.4	-Tras un tiempo leyendo u observando una pantalla de cerca, ¿has notado que te molesten las luces?
1	631	439	-.53	.05	.52	-9.1	.54	-7.6	.69	.59	56.4	39.2	-Notas tus ojos cansados?
4	650	435	-.59	.05	1.14	2.1	1.13	1.7	.55	.60	33.8	38.3	-Te entra sueño?
MEAN	446.9	437.6	.00	.06	1.02	.0	1.00	-.2		46.9	45.4		
P.SD	129.5	1.5	.42	.01	.26	3.7	.25	3.1		8.8	6.8		

FIGURA 8.30 – CARACTERÍSTICAS DE LOS 8 NUEVOS ÍTEMES INTRODUCIDOS – COMPARATIVA CON LAS CARACTERÍSTICAS DE LOS 15 ÍTEMES PERTENECIENTES A LA ESCALA CISSVE

El disponer de amplios bancos de ítems y el conocer las características psicométricas de cada uno, facilita la posibilidad de poder elegir los más apropiados en función de la usabilidad que se desee dar al modelo. En la Figura 8.31 se aprecia un mapa de Wright que enfrenta el conjunto de los 23 ítems a la muestra de estudio y en la que se puede comprender de forma visual cuáles son los más apropiados para medir en cada intervalo de la escala en la muestra de estudio elegida. Si pretendiéramos examinar una población diferente, elegiríamos los ítems más adecuados en cuanto a nivel de dificultad para esa población de forma que se maximizara la precisión, adecuando al máximo el nivel de los ítems al de los sujetos para cada caso.

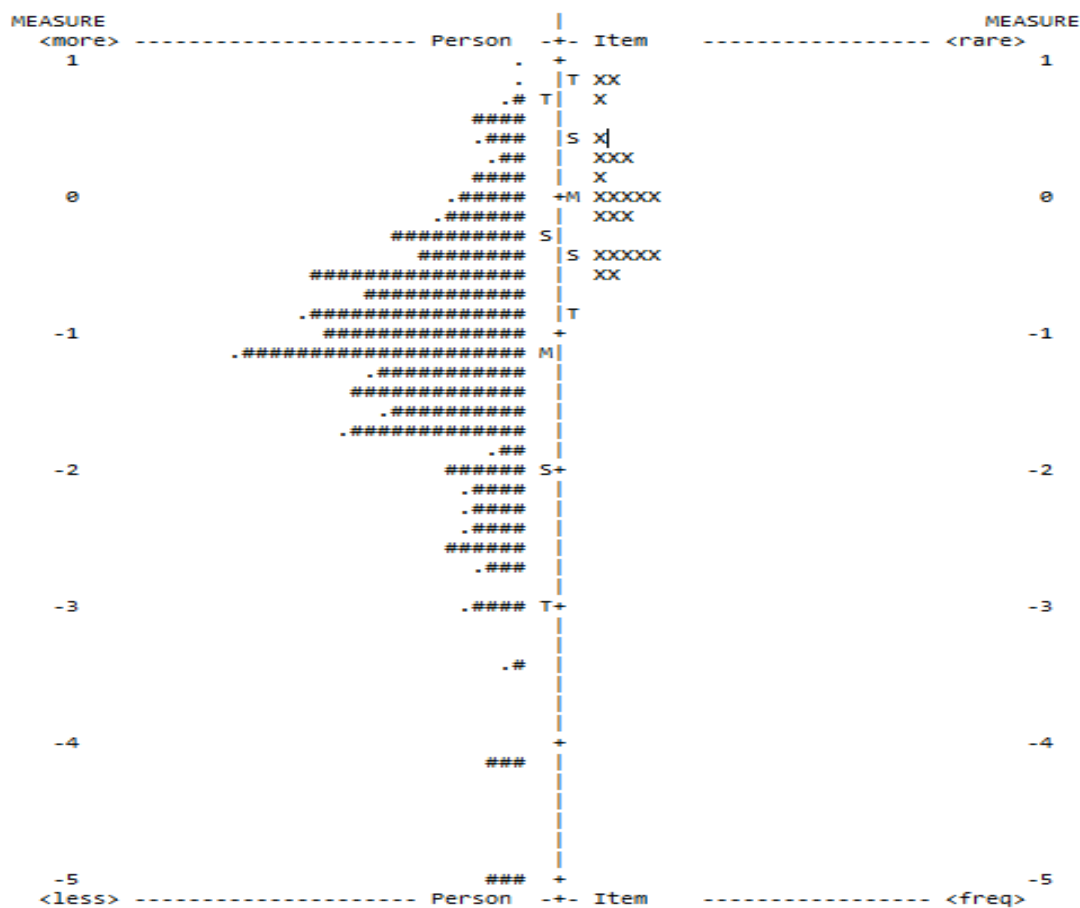


FIGURA 8.31 – MAPA DE WRIGHT QUE POSICIONA EN LA MISMA ESCALA LOS 23 ÍTEMES (15 DE LA ESCALA CISSVE + 8 ADICIONALES) Y LOS SUJETOS DE LA MUESTRA UTILIZADA EN ESTE TRABAJO

El objetivo de este trabajo queda limitado a iniciar una nueva línea de trabajo que se podrá continuar en futuros estudios de investigación y que tiene como meta el desarrollar un amplio banco de ítems organizados por nivel de dificultad, que posibilite la construcción de escalas mejor adaptadas a cada tipo de población con la que trabaje en cada caso concreto. Tiene sentido que puedan ser agrupados según los tipos de sintomatología que describan y de acuerdo con los factores encontrados durante el desarrollo de este trabajo que hemos podido observar, son los que resumen los síntomas asociados a la IC.

De esta forma y a modo de ejemplo, se puede observar cómo cada uno de los 8 ítems añadidos de forma experimental en nuestro trabajo, queda mapeado al factor correspondiente de los

que se encontraron al realizar el análisis factorial que se presentó en el apartado 5.3.3.3.7 cuando se analizaron los atributos que mide la escala CISSve. Así pues, realizando un análisis factorial de nuevo sobre los 23 ítems podemos comprobar como los ítems 20, 22 y 23 pasan a engrosar el grupo de síntomas relacionados con la fatiga visual tras trabajo prolongado de cerca, el 16, 17 y 21 claramente relacionados con síntomas asociados a disfunción visual pasan a ocupar su lugar en ese grupo y por último el 18 y 19 se sitúan entre los que describen síntomas relacionados con problemas de concentración a leer o trabajar de cerca.

En la Figura 8.32 se puede ver cómo un análisis factorial realizado sobre los 23 ítems traslada a cada nuevo ítem de entre 8 introducidos a su grupo correspondiente de entre los que se explicaron durante el estudio de la dimensionalidad de la escala CISSve en el cuerpo de este trabajo.

	Componente		
	1	2	3
ITEM1	,781		
ITEM2	,717		
ITEM10	,671		
ITEM11	,655		
ITEM23	,562		
ITEM12	,535		
ITEM3	,519		
ITEM22	,480		
ITEM20	,446		
ITEM8		,749	
ITEM7		,733	
ITEM17		,669	
ITEM13	,434	,624	
ITEM16		,537	
ITEM21		,423	
ITEM18			,667
ITEM15			,643
ITEM14			,638
ITEM5			,599
ITEM6			,552
ITEM9			,548
ITEM4	,455		,459
ITEM19			,427

**FATIGA TRAS TRABAJO PROLONGADO DE CERCA**

- 20 - ¿Te acercas mucho a lo que estás mirando?
- 22 - Tras un tiempo leyendo u observando una pantalla de cerca. ¿Has notado que te molesten las luces?
- 23 - Tras un tiempo leyendo o trabajando de cerca, ¿tienes que esforzarte para ver bien?

**DISFUNCIÓN VISUAL**

- 16 - ¿Tienes la sensación de ponerte bizco?
- 17- ¿A veces se te emborronan las letras?
- 21 - ¿Guiñas los ojos?

**FALTA DE CONCENTRACIÓN**

- 18 - ¿Señalas con el dedo (o algo similar) la palabra que estás leyendo?
- 19 - ¿Inclinas la cabeza hacia un lado?

FIGURA 8.32 – ANÁLISIS FACTORIAL SOBRE LOS 23 ÍTEMS (15 DE LA ESCALA CISSVE + 8 ADICIONALES)

8.3.6 DISCUSIÓN

La inclusión de nuevos ítems ilustra empíricamente la flexibilidad y adaptabilidad del modelo de Rasch. Sin embargo, el estudio sugiere que los ítems añadidos, aunque son compatibles en términos de nivel de dificultad, no necesariamente aportan mejoras significativas a las propiedades métricas del instrumento existente.

El potencial que proporciona el análisis de Rasch de cara a la aplicabilidad práctica, es la posibilidad, por un lado, el proporcionar una herramienta que habilita al profesional a poder comparar el nivel de dificultad de cada ítem (sintomatología que es capaz de discriminar) y por

lo tanto precisar la cantidad de información que cada ítem es capaz de aportar sobre la sintomatología de cada sujeto en función del grado de disfunción que padezca y por otro, el de poder adaptar el conjunto de preguntas a la muestra en función del tipo de población del que se trate.

Gracias a las propiedades del modelo de Rasch los ítems incorporados en el instrumento, pueden ser comparados en cuanto a nivel de dificultad con los que ya estaban disponibles, y así poder enriquecer los intervalos de la escala que se corresponden con grupos poblacionales determinados a los que se quiere medir, como por ejemplo poblaciones generales no sintomáticas, poblaciones clínicas sintomáticas, etc, consiguiendo mejorar así la precisión de medida en esas zonas.

Por lo tanto, la posibilidad de construir un amplio banco de ítems organizados por nivel de dificultad nos confiere la versatilidad de poder adaptar la escala a las diferentes muestras que se necesiten examinar, mejorando así el “targeting”, y por lo tanto la precisión en las medidas para esos grupos.

Como se explicó en el apartado 5.2.3.3.4, en el que se repasaban los niveles de síntomas que era capaz de distinguir CISSve, utilizando el método de Wright, se pudo comprobar que, con independencia de la muestra a evaluar, la escala era capaz de distinguir hasta 6.3 niveles de síntomas, mientras, que, para la muestra seleccionada en el caso de nuestro estudio, sólo discernía 3.3 niveles de síntomas. Esto se explica por el hecho de que, al tratarse de una población mayoritariamente asintomática, los sujetos de esa muestra sólo se movían entre los primeros niveles de síntomas detectables por la escala, y por lo tanto, los ítems que más información aportaban, para esa población, eran los que mejor mapeaban en logits con el nivel de sintomatología de los sujetos. Esta circunstancia se pudo comprobar nuevamente al ser capaces de crear versiones reducidas de la escala CISSve, manteniendo prácticamente intacta la fiabilidad (apartado 8.1).

Sin embargo, en el hipotético caso de que se quisiera utilizar la escala para poblaciones clínicas sintomáticas, es presumible que el añadir ítems más difíciles que pudieran mapear con los sujetos con ese tipo de clínicas, ayudaría a mejorar la precisión en esos intervalos de “dificultad” de la escala, contribuyendo así a mejorar la fiabilidad del instrumento para esa hipotética muestra, y así habilitar la posibilidad de ser capaces de distinguir mejor los diferentes niveles de sintomatología entre esos sujetos.

Esta línea podría ser objeto de interés para futuras líneas de investigación

### 8.3.7 CONCLUSIONES

El modelo de Rasch ofrece un marco robusto para la ampliación de bancos de ítems, permitiendo adaptar la escala a diferentes poblaciones. No obstante, la adición de nuevos ítems debe ser realizada con cautela, ya que no garantiza automáticamente un aumento en la precisión o la fiabilidad del instrumento.

## 8.4 UTILIZACIÓN DE METODOS PREDICTIVOS BASADOS EN ALGORITMOS DE CLASIFICACIÓN DE MACHINE LEARNING PARA REALIZAR CUESTIONARIOS ADAPTATIVOS

En este apartado, aprovechando el potencial de la “ciencia de datos” (data science) y utilizando técnicas de “minería de datos”<sup>184,185</sup> y algoritmos de “machine learning”<sup>186,187</sup>, pilares de la inteligencia artificial<sup>188</sup> tratamos de explotar el potente análisis predictivo que artifican y buscamos su aplicabilidad a la administración inteligente de escalas. La creación de tests adaptativos mediante árboles de decisión dinámicos<sup>187,189</sup> que consigan adaptar las preguntas de la escala a los sujetos en función de las respuestas a las anteriores es una línea digna de ser investigada en futuros trabajos y puede llevar a este tipo de instrumentos a un punto de evolución diferencial.

Recordemos que el modelo de Rasch sostiene que la dificultad de un ítem no depende de la muestra medida ni de los demás ítems y en base a estas premisas, un sujeto que contestara de determinada forma a un ítem podría ser dirigido hacia el siguiente en función de su respuesta, minimizando el camino crítico hasta conseguir clasificarlo.

### 8.4.1 OBJETIVOS

El propósito primordial de este estudio es explorar la aplicabilidad de técnicas de “machine learning”, concretamente árboles de decisión dinámicos, para el diseño y la administración de cuestionarios adaptativos. Se centra especialmente en la evaluación de las escalas CISSve y CISSve-R2 para determinar su eficacia en la clasificación de sujetos como "sintomáticos" o "no sintomáticos" conforme a criterios preestablecidos.

### 8.4.2 JUSTIFICACIÓN

La necesidad de este trabajo surge de la convergencia entre ciencia de datos y escalas de medición en psicología y medicina. Utilizar algoritmos de aprendizaje automático para adaptar las preguntas de la escala a los sujetos podría significar un avance diferencial en la evaluación psicométrica y médica, al minimizar el tiempo y maximizar la precisión en el diagnóstico.

### 8.4.3 HIPÓTESIS

Se postula que el uso de árboles de decisión dinámicos basados en algoritmos predictivos de “machine learning” podría aumentar la eficacia en la clasificación de individuos según escalas CISSve y CISSve-R2, al comparar la concordancia de los resultados con el análisis de Rasch previamente realizado (apartado 5.2.3.3).

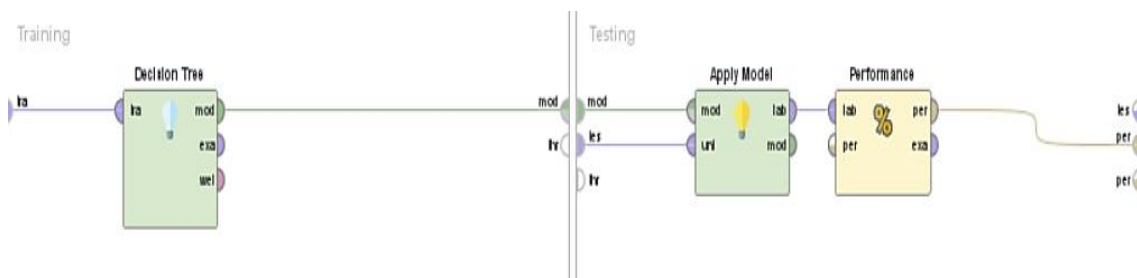
### 8.4.4 MÉTODOS

Para la evaluación de las escalas CISSve y CISSve-R2, se empleó un enfoque de modelo predictivo<sup>184</sup> basado en árboles de decisión binarios<sup>187,189</sup>. El algoritmo se entrenó con una parte de la muestra utilizada durante el estudio (n=458, M=16.02, DT=5.68 años), extraída al azar que contenía las respuestas a los ítems de los sujetos y su clasificación como sintomáticos o no sintomáticos. En el caso de la escala CISSve, la clasificación que se inyectó para el entrenamiento del algoritmo se obtuvo de acuerdo con el criterio establecido por los autores del cuestionario original, CISSv15, Borsting y Rouse<sup>59,60,74</sup> (Puntuación>16 para niños y Puntuación>21 para adultos). En el caso de la escala CISSve-R2 (versión de 2 ítems), la puntuación de corte se situó en 4, que como se pudo observar en el apartado 8.1.5 suponía el “0” de escala de dicho instrumento según el el análisis de Rasch que se llevó a cabo.

Adicionalmente, se llevó a cabo un tercer análisis, realizando una preparación previa de los datos, en la cuál se organizaron las puntuaciones en lugar de como una simple agregación de las puntuaciones de cada ítem, (tal y como fueron obtenidas originariamente), en este caso mediante la definición de nuevas variables asociadas a cada factor, es decir a la fatiga visual, a la concentración y a función visual, de acuerdo a los factores obtenidos durante el estudio (apartado 5.2.3.3.7). Se establecieron: Fatiga visual = (ítem 1 + ítem 2 + ítem 3 + ítem 10 + ítem 11 + ítem 12)/6; Concentración = (ítem 4 + ítem 5 + ítem 6 + ítem 15)/4 y Función visual = (ítem 7 + ítem 8 + ítem 9 + ítem 13 + ítem 14)/5.

### 8.4.5 RESULTADOS

En el diagrama de la Figura 8.33, se puede ver de forma esquemática, a la izquierda cómo se entrena al modelo, y a la derecha, como a continuación se aplica sobre los datos de la escala y se mide su rendimiento.



**FIGURA 8.33 – ESQUEMA DEL PROCESO PARA LA PREDICCIÓN DE SUJETOS “SINTOMÁTICOS” Y “NO SINTOMÁTICOS” EN LAS ESCALAS CISSVE Y CISSVE-R2, SIGUIENDO UN ALGORITMO DE CLASIFICACIÓN BASADO EN ÁRBOLES BINARIOS**



Observando la tabla de la

Figura 8.35 -, podemos ver un resumen del rendimiento en la clasificación de los sujetos de acuerdo a este algoritmo. Se puede comprobar qué porcentaje de “verdaderos positivos” (VP), pacientes son clasificados (pred true) como sintomáticos por el algoritmo y siendo realmente sintomáticos de acuerdo a los criterios predefinidos (true true), cuáles son “verdaderos negativos” (VN), clasificados como asintomáticos (pred false) y realmente lo son según el criterio definido (true false), e igualmente, se pueden observar los “falsos positivos” (FP), que enfrenta los (pred true) a los (true false) y los “falsos negativos” (FN) que en este caso enfrenta a los (pred false) con los (true true).

Es decir podemos obtener los parámetros de sensibilidad:  $(VP)/(VP+FN)=75\%$  y de especificidad,  $VN/(VN+FP)=85.7\%$ . Como también se observa en la

Figura 8.35 -, el modelo presenta una precisión del 82.6%, muy en línea con la fiabilidad de un 83% que daba el análisis de Rasch.

accuracy: 82.61% +/- 3.42% (micro average: 82.60%)

	true false	true true	class precision
pred. false	263	44	85.67%
pred. true	31	93	75.00%
class recall	89.46%	67.88%	

FIGURA 8.35 - RENDIMIENTO DEL ÁRBOL DE DECISIÓN BINARIO APLICADO A CISSVE

Haciendo el mismo ejercicio con la escala de 2 ítems (CISSve-R2), podemos observar en la

Figura 8.36 - el árbol de decisión obtenido, cuando se toma como criterio una puntuación de corte para sintomáticos de 4 puntos (“0” de escala en el modelo de Rasch).

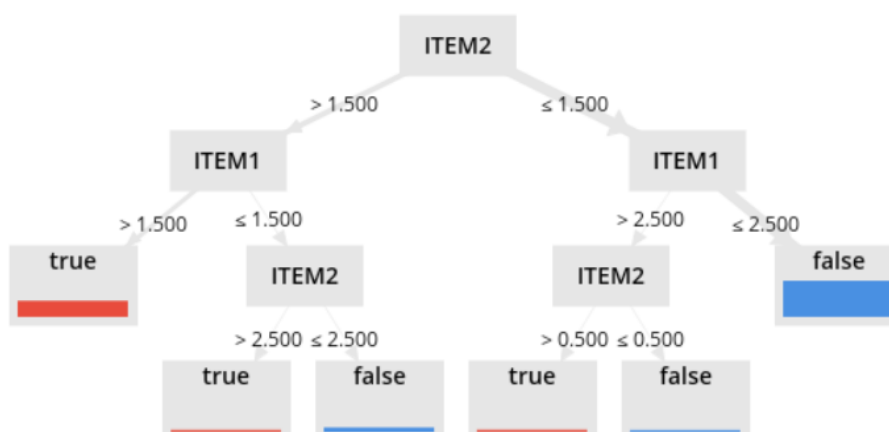


FIGURA 8.36 - ÁRBOL DE DECISIÓN BINARIO QUE CLASIFICA A LOS SUJETOS COMO “SINTOMÁTICOS” O NO SINTOMÁTICOS DE ACUERDO A LOS CRITERIOS DEFINIDOS PARA CISSVE-R2 (PUNTUACIÓN>=4), TOMANDO EN CUENTA SUS RESPUESTAS A LOS ÍTEMS DE LA ESCALA CISSVE-R2

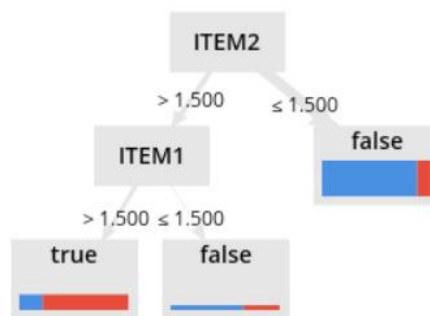
Este árbol arroja unos valores de rendimiento fabulosos clasificando a los sujetos de acuerdo al criterio de “sintomáticos” si su puntuación es mayor o igual que 4. Como se puede observar en la Figura 8.37.

accuracy: 98.37% +/- 2.70% (micro average: 98.38%)

	true false	true true	class precision
pred. false	298	0	100.00%
pred. true	7	126	94.74%
class recall	97.70%	100.00%	

**FIGURA 8.37 - RENDIMIENTO DEL ÁRBOL DE DECISIÓN BINARIO APLICADO A CISSVE-R2**

En la Figura 8.38 podemos observar un árbol binario que clasifica los sujetos que contestan la escala CISSve-R2 (2 ítems), pero esta vez de acuerdo a los criterios de sintomatología marcados para la escala CISSve (15 ítems) para esos mismos sujetos, es decir, representa el modelo para clasificar los sujetos como sintomáticos de la escala CISSve, contestando únicamente los 2 ítems de la escala CISSve-R2.



**FIGURA 8.38 - ÁRBOL DE DECISIÓN BINARIO QUE CLASIFICA A LOS SUJETOS COMO “SINTOMÁTICOS” O NO SINTOMÁTICOS DE ACUERDO A LOS CRITERIOS DEFINIDOS PARA CISSVE, TOMANDO EN CUENTA SUS RESPUESTAS A LOS ÍTEMS DE LA ESCALA CISSVE-R2**

Como vemos, el árbol comienza por el ítem más difícil de los 2, el ítem 2. Si la respuesta al ítem 2 es “0”, es decir “nunca” o “1”, es decir “casi nunca”, entonces se clasifica el sujeto como “no sintomático” directamente. Si la respuesta al ítem 2, es “2”, “3” o “4” el algoritmo mira la respuesta al ítem 1. Si fuera “1”, o “0”, el algoritmo clasificaría como “no sintomático” al sujeto, sin embargo, de ser “2”, “3”, o “4”, le marcaría como sintomático. En resumen. Este algoritmo clasifica como sintomáticos solamente a los sujetos que pasandoles por ese orden los ítems, reporten “2” o más puntos en el ítem 2 y seguidamente, “2” o más puntos en el ítem 1.

En la Figura 8.39 se puede observar que este algoritmo alcanza una precisión de casi un 83%, muy elevada considerando que está evaluando el porcentaje de pacientes que son correctamente clasificados como sintomáticos bajo el criterio de CISSve. Este alto rendimiento

va en consonancia con la alta fiabilidad que ya otorgaba a esta versión reducida de 2 ítems, el análisis de Rasch.

De la misma forma y a modo de colofón, construimos un nuevo modelo en el que las variables que introducimos como entrada al árbol de decisión, son además del grupo de edad al que pertenece el sujeto, 3 variables resumen que aglutinan a cada grupo de síntomas asociados con cada uno de los factores componentes de CISSve. La forma de construir estas variables, consiste en sumar las respuestas correspondientes a los ítems que constituyen un mismo factor, y dividir esa suma entre el número de ítems que componen el factor. Ese valor otorgará una puntuación a esa variable. Así pues se forman las variables, fatiga, concentración y función visual, siendo  $\text{fatiga} = (\text{item 1} + \text{item 2} + \text{item 3} + \text{item 10} + \text{item 11} + \text{item 12}) / 6$ ;  $\text{concentración} = (\text{item 4} + \text{item 5} + \text{item 6} + \text{item 15}) / 4$  y  $\text{función visual} = (\text{item 7} + \text{item 8} + \text{item 9} + \text{item 13} + \text{item 14}) / 5$ .

accuracy: 82.84% +/- 6.81% (micro average: 82.83%)

	true false	true true	class precision
pred. false	267	47	85.03%
pred. true	27	90	76.92%
class recall	90.82%	65.69%	

**FIGURA 8.39 - RENDIMIENTO DEL ÁRBOL DE DECISIÓN BINARIO APLICADO A CISSVE-R2 CON LOS CRITERIOS DE CLASIFICACIÓN COMO SINTOMÁTICO DEFINIDOS PARA CISSVE**

Este modelo, cuyo árbol de decisión se puede observar en la Figura 8.41, pretende aportar significado clínico a la clasificación como “sintomático” o “no sintomático” de acuerdo al tipo de síntomas reportados.

Como muestra la tabla de la Figura 8.40, la precisión es excelente, circunstancia que señala de nuevo a una posible vía de investigación que ya habíamos apuntado para futuros trabajos en el apartado 6.3 de la discusión, la de construir una escala más corta que sintetice en menos ítems los síntomas asociados a una misma tipología que nosotros hemos aglutinado en cada factor.

accuracy: 91.41% +/- 1.93% (micro average: 91.42%)

	true false	true true	class precision
pred. false	276	19	93.56%
pred. true	18	118	86.76%
class recall	93.88%	86.13%	

**FIGURA 8.40 - RENDIMIENTO DEL ÁRBOL DE DECISIÓN BINARIO CONSTRUIDO CON LAS VARIABLES QUE AGLUTINAN LOS 3 FACTORES COMPONENTES DE CISSVE (FATIGA VISUAL, PROBLEMAS DE VISIÓN BINOCULAR Y CAPACIDAD DE CONCENTRACIÓN)**



#### 8.4.6 DISCUSIÓN

Los resultados obtenidos corroboran la eficiencia de los modelos predictivos basados en árboles de decisión. Estos resultados están en línea con las métricas obtenidas a través del análisis de Rasch, lo que valida aún más la utilidad del método. Además, se han establecido modelos que incluyen variables sintomáticas específicas como fatiga, concentración y función visual.

Estos modelos son sólo una pequeña muestra del enorme potencial y la aplicabilidad de métodos predictivos basados en algoritmos de clasificación como éste. Por supuesto se pueden utilizar otros muchos tipos, como pueden ser, regresiones logísticas, redes neuronales o modelos bayesianos, pero, como habíamos introducido anteriormente, el único propósito de este ejercicio consiste por una parte en la verificación del correcto funcionamiento de las escalas estudiadas de acuerdo a las propiedades observadas durante el transcurso del trabajo investigativo, y por otra y en un ámbito distinto, el de introducir una nueva línea de trabajo para futuras investigaciones. Hemos hecho otras muchas simulaciones con diferentes modelos de una complejidad un poco más elevada a los presentados en este trabajo pero como decimos nuestro objetivo es simplemente dejar esta línea abierta a trabajos venideros.

La programación de un simple algoritmo capaz de mostrar a un sujeto un ítem distinto al que le mostraría a otro sujeto en función de su respuesta al ítem anterior, y siguiendo las directrices marcadas por el correspondiente algoritmo de clasificación preentrenado (en este caso el árbol de decisión), sería una implementación práctica sencilla y de enorme utilidad para el profesional clínico. Una herramienta así, minimizaría el tiempo invertido por el sujeto en contestar a la escala al ser clasificado como sintomático o no sintomático en el menor número de pasos, ya que la aplicación le guiaría por el camino crítico.

#### 8.4.7 CONCLUSIONES

El estudio demuestra que la utilización de métodos predictivos basados en algoritmos de “machine learning” es una línea de investigación viable y eficaz para el diseño de cuestionarios adaptativos. Las escalas CISSve y CISSve-R2 han mostrado eficiencia en la clasificación de sujetos, y los árboles de decisión binarios han demostrado ser una herramienta robusta para tal fin.





**COMUNICACIONES  
CIENTÍFICAS  
DERIVADAS  
DE LA TESIS**



## 9 COMUNICACIONES CIENTÍFICAS DERIVADAS DE LA TESIS

### 9.1 PUBLICACIONES EN REVISTAS INTERNACIONALES

- González-Pérez M, Pérez-Garmendia C, Barrio AR, García-Montero M, Antona B. Spanish Cross-Cultural Adaptation and Rasch Analysis of the Convergence Insufficiency Symptom Survey (CISS). *Transl Vis Sci Technol.* 2020;9(4):23.

### 9.2 COMUNICACIONES EN CONGRESOS NACIONALES E INTERNACIONALES

- Pérez Garmendia, Carlos; González Pérez, Mariano; Barrio, Ana Rosa; Antona Peñalba B. Traducción y adaptación a español del cuestionario CISS. 1ª Jornada PhDay Complutense. Madrid; 2017.
- Pérez Garmendia, Carlos; González Pérez, Mariano; Barrio, Ana Rosa; Antona Peñalba B. Translation and adaptation into Spanish of CISS. 2ª Jornada PhDay Complutense. Madrid; 2018.
- Pérez Garmendia, Carlos; González Pérez, Mariano; Barrio, Ana Rosa; Antona Peñalba B. Convergence Insufficiency Symptom Survey - Análisis de comportamiento de CISSve vs CISS. 3ª Jornada PhDay Complutense. Madrid. 2019.
- Pérez Garmendia, Carlos; González Pérez, Mariano; Barrio, Ana Rosa; Antona Peñalba B. Reingeniería de la escala Convergence Symptom Survey versión española (CISSve). In: IV Jornada Científica BiblioPro. Madrid; 2019.
- Pérez Garmendia, Carlos; González Pérez, Mariano; ; Barrio, Ana Rosa; Antona Peñalba B. The CISSve-R2 scale. A reduced versión of CISSve. 4ª Jornada PhDay Complutense. Madrid. 2020
- Pérez Garmendia, Carlos; González Pérez, Mariano; Barrio, Ana Rosa; Antona Peñalba B. Predictive Machine Learning adapted to the CISSve Survey. 5ª Jornada PhDay Complutense. Madrid. 2021
- Pérez Garmendia, Carlos; González Pérez, Mariano; Barrio, Ana Rosa; Antona Peñalba B. Analysis of the CISSve scores pre and post visual therapy in patients clinically diagnosed with convergence insufficiency. 6ª Jornada PhDay Complutense. Madrid 2022
- Pérez Garmendia, Carlos; González Pérez, Mariano; Barrio, Ana Rosa; Antona Peñalba B. Caracterización y análisis factorial de la versión adaptada a español (CISSve) de la escala "Convergence Insufficiency Symptom Survey" en sujetos entre 9 y 30 años. 26 Congreso Internacional de Optometría y Óptica Oftálmica. Madrid 2020.



# REFERENCIAS



## 10 REFERENCIAS

1. Unamuno EA. Recensión bibliográfica. *Gac Sanit.* 2013;27(6):570. doi:10.1016/j.gaceta.2013.04.002
2. González Aguilar H, Rubio Pachamango A. The structure of abstracts in original articles. *Educacion Medica.* 2018;19:222-223. doi:10.1016/j.edumed.2017.10.011
3. Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report. *Value in Health.* 2009;12(8):1075-1083. doi:10.1111/j.1524-4733.2009.00603.x
4. Muñiz J, Elosua P, Hambleton RK. Directrices para la traducción y adaptación de los tests: Segunda edición. *Psicothema.* 2013;25(2):151-157. doi:10.7334/psicothema2013.24
5. Fisher WP. A Practical Approach to Modeling Complex Adaptive Flows in Psychology and Social Science. In: *Procedia Computer Science.* Vol 114. Elsevier B.V.; 2017:165-174. doi:10.1016/j.procs.2017.09.027
6. Andrich D. A rating formulation for ordered response categories. *Psychometrika.* 1978;43(4):561-573. doi:10.1007/BF02293814
7. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Rasch analysis of visual function and quality of life questionnaires. *Optometry and Vision Science.* 2009;86(10):1160-1168. doi:10.1097/OPX.0b013e3181bab10c
8. Ventura-León J. Escalas, inventarios y cuestionarios: ¿son lo mismo? *Educación Médica.* Published online May 2019. doi:10.1016/j.edumed.2019.04.001
9. Supo. J, Cómo validar un instrumento – La guía para validar un instrumento en 10 pasos N ° 2012-04073, ISBN: 1492278904 ISBN-13: 978- 1492278900.2013.31-64pp.
10. Piedad Molina M N, Mora Forero C. Insuficiencia de convergencia. *Ciencia & Tecnologia para la Salud Visual Y Ocular.* 2010;8(2):91-102.
11. Scheiman M, Wick B. *Clinical Management of Binocular Vision: Heterophoric, Accommodative, and Eye Movement Disorders: Fifth Edition.* Lippincott Williams & Wilkins; 2020.
12. Perceiving in Depth, Volume 1 - Hardback - Ian P. Howard - Oxford University Press. Accessed May 22, 2020. <https://global.oup.com/academic/product/perceiving-in-depth-volume-1-9780199764143?cc=es&lang=en&>
13. García-Muñoz Á, Carbonell-Bonete S, Cacho-Martínez P. Symptomatology associated with accommodative and binocular vision anomalies. *J Optom.* 2014;7(4):178-192. doi:10.1016/j.optom.2014.06.005

## REFERENCIAS

14. Cacho-Martínez P, García-Muñoz Á, Ruiz-Cantero MT. Is there any evidence for the validity of diagnostic criteria used for accommodative and nonstrabismic binocular dysfunctions? *J Optom.* 2014;7(1):2-21. doi:10.1016/j.optom.2013.01.004
15. Nityananda V, Read JCA. Stereopsis in animals: Evolution, function and mechanisms. *Journal of Experimental Biology.* 2017;220(14):2502-2512. doi:10.1242/jeb.143883
16. Benjamin WJ. *Borish's Clinical Refraction.* Elsevier Inc.; 2006. doi:10.1016/B978-0-7506-7524-6.X5001-7
17. Katz BJ. Models of the Visual System. *Journal of Neuro-Ophthalmology.* 2004;24(1):83-84. doi:10.1097/00041327-200403000-00019
18. Cufflin MP, Mallen EAH. Dynamic accommodation responses following adaptation to defocus. *Optometry and Vision Science.* 2008;85(10):982-991. doi:10.1097/OPX.0b013e3181886fda
19. Adaptation of Tonic Accommodation Reduces Accommodative Lag and Is Masked in Darkness - PubMed. Accessed May 23, 2020. <https://pubmed.ncbi.nlm.nih.gov/3700031/>
20. Hung GK, Ciuffreda KJ. Models of Refractive Error Development. In: ; 2002:643-677. doi:10.1007/978-1-4757-5865-8\_18
21. Sen DK, Malik SR. Accommodative-convergence over accommodation (AC/A) ratio (in normal Indian subjects). *Indian J Ophthalmol.* 1972;20(4):153-157.
22. Hirani KJ, Firth AY. Convergence accommodation to convergence (CA/C) ratio: stability with different levels of convergence demand. *British and Irish Orthoptic Journal.* 2009;6(0):60. doi:10.22599/bioj.11
23. Cooper J, Selenow A, Ciuffreda KJ, et al. Reduction of asthenopia in patients with convergence insufficiency after fusional vergence training. *Optometry and Vision Science.* 1983;60(12):982-989. doi:10.1097/00006324-198312000-00007
24. Amos JF. Clinical Procedures in Primary Eye Care, 4th ed. *Optometry and Vision Science.* 2014;91(1):e22. doi:10.1097/opx.0000000000000176
25. Candia Zegarra E, Venegas Palomino J. Identificación de las principales competencias del personal de ventas de una empresa de consumo masivo y su relación con los resultados de alto desempeño. *Repositorio de la Universidad del Pacífico - UP.* Published online 2016.
26. Carbonell Bonete S. Prevalencia y sintomatología de las disfunciones acomodativas y binoculares en la población universitaria. *Tesis doctoral.* Published online 2014:1-5.
27. Antona Peñalba Beatriz, Barra Lázaro F, Barrio de Santos ARosa, González Díaz-Obregón Enrique, Sánchez Pérez Isabel. *Procedimientos Clínicos Para La Evaluación de La Visión Binocular.* Editorial Medica Panamericana; 2017.
28. Piñero Llorens. La disparidad de fijación: concepto y manejo clínico. *Gaceta de Optometría y Óptica Oftálmica.* Published online 2015:502. 5p.
29. Jiménez R, Jiménez JR, González R PAM. Importancia de la estereopsis en la vida cotidiana. *Gaceta Optica.* 2006;404:10-12.

30. Banks MS, Kim J, Shibata T. Insight into vergence/accommodation mismatch. In: *Head- and Helmet-Mounted Displays XVIII: Design and Applications*. Vol 8735. SPIE; 2013:873509. doi:10.1117/12.2019866
31. Junghans BM, Kiely PM, Crewther DP, Crewther SG. Referral rates for a functional vision screening among a large cosmopolitan sample of Australian children. *Ophthalmic and Physiological Optics*. 2002;22(1):10-25. doi:10.1046/j.1475-1313.2002.00010.x
32. Morris H, O'Connor AR, Stephenson MG, Mitxhell M, Price G AS. Clinical assessment of stereopsis and its functional significance. *Br Ir Orthopt J*. Published online 2005:2:8-14.
33. Kříž P, Skorkovská Š. Distance associated heterophoria measured with polarized cross test of MKH method and its relationship to refractive error and age. *Clin Optom (Auckl)*. 2017;9:55-65. doi:10.2147/OPTO.S123436
34. Grosvenor T. *Optometria de Atención Primaria*. Masson; 2005.
35. Cacho-Martínez P, García-Muñoz Á, Ruiz-Cantero MT. Do we really know the prevalence of accommodative and nonstrabismic binocular dysfunctions? *J Optom*. 2010;3(4):185-197. doi:10.1016/S1888-4296(10)70028-5
36. Onnes K. Huygens Institute - Royal Netherlands Academy of Arts and Sciences ( KNAW ). *Proceeding of Royal Netherlands Academy of Arts and Sciences (KNAW)*. 1911;13 II(September):1274-1276.
37. DAUM KM. Accommodative Insufficiency. *Optometry and Vision Science*. 1983;60(5):352-359. doi:10.1097/00006324-198305000-00002
38. Shin HS, Park SC, Park CM. Relationship between accommodative and vergence dysfunctions and academic achievement for primary school children. *Ophthalmic and Physiological Optics*. 2009;29(6):615-624. doi:10.1111/j.1475-1313.2009.00684.x
39. Ruiz M, Villasante F, León F, et al. Dyspepsia-related quality of life. Spanish adaptation and validation of the questionnaire «Dyspepsia-Related Health Scale». *Med Clin (Barc)*. 2001;117(15):567-573. doi:10.1016/S0025-7753(01)72182-3
40. Borsting E, Chase C, Tosha C, Ridder WIWH. Longitudinal study of visual discomfort symptoms in college students. *Optom Vis Sci*. 2008;85(10):992-998. doi:10.1097/OPX.0b013e31818883cd
41. FAAO PDBaLos. The prevalence of vergence accommodation disorders in a school-age population. *Clin Exp Optom*. 1992;75(1):10-18. doi:10.1111/j.1444-0938.1992.tb01010.x
42. Daum KM. Accommodative dysfunction. *Documenta Ophthalmologica*. 1983;55(3):177-198. doi:10.1007/BF00140808
43. Hyndman J. Spasm of the Near Reflex: Literature Review and Proposed Management Strategy. *J Binocul Vis Ocul Motil*. 2018;68(3):78-86. doi:10.1080/2576117X.2018.1486680
44. FAA PSDBaLos. Clinical criteria for vergence accommodation dysfunction. *Clin Exp Optom*. 1991;74(4):112-119. doi:10.1111/j.1444-0938.1991.tb04622.x

## REFERENCIAS

45. Antona B, Barrio AR, Gascó A, Pinar A, González-Pérez M, Puell MC. Symptoms associated with reading from a smartphone in conditions of light and dark. *Appl Ergon*. 2018;68. doi:10.1016/j.apergo.2017.10.014
46. Simons HD, Grisham JD. *Binocular Anomalies and Reading Problems*. Vol 58.; 1987:578-587.
47. Porcar E, Martinez-Palomera A. Prevalence of general binocular dysfunctions in a population of university students. *Optometry and Vision Science*. 1997;74(2):111-113. doi:10.1097/00006324-199702000-00023
48. Duane A. A new classification of the motor anomalies of the eye based upon physiological principles: together with their symptoms, diagnosis and treatment. Published online 1897.
49. Daum KM. Convergence insufficiency. *Optometry and Vision Science*. 1984;61(1):16-22. doi:10.1097/00006324-198401000-00003
50. Cooper J, Jamal N. *Convergence Insufficiency-a Major Review*. Vol 83.; 2012:137-158.
51. González-Pérez M, Susi R, Antona B, Barrio A, González E. The Computer-Vision Symptom Scale (CVSS17): Development and initial validation. *Invest Ophthalmol Vis Sci*. 2014;55(7):4504-4511. doi:10.1167/iovs.13-13818
52. Chen DL. *CE@Home*.
53. Martínez BA, González VBA, Sánchez LEG. Calibración del índice de hiperactividad de conners mediante el modelo de rasch. *Universitas Psychologica*. 2013;12(3):957-970. doi:10.11144/Javeriana.UPSY12-3.cihc
54. Borsting E, Rouse M, Chu R. Measuring ADHD behaviors in children with symptomatic accommodative dysfunction or convergence insufficiency: a preliminary study. *Optometry*. 2005;76(10):588-592. doi:10.1016/j.optm.2005.07.007
55. Karaca I, Demirkılıç Biler E, Palamar Onay M, Özbaran B, Üretmen Ö. Stereoacuity, Fusional Vergence Amplitudes, and Refractive Errors Prior to Treatment in Patients with Attention-Deficit Hyperactivity Disorder. *Turk J Ophthalmol*. 2020;50(1):15-19. doi:10.4274/tjo.galenos.2019.17802
56. Shin HS, Park SC, Maples WC. Effectiveness of vision therapy for convergence dysfunctions and long-term stability after vision therapy. *Ophthalmic and Physiological Optics*. 2011;31(2):180-189. doi:10.1111/j.1475-1313.2011.00821.x
57. Borsting E, Rouse M DLP. Prospective Insufficiency, comparison of convergence Insufficiency, and normal binocular children on CIRS symptom surveys. *Convergence*. *Optom Vis Sci*. Published online 1999:1999;76:221-8.
58. Lara F, Cacho P, García Á, Megías R. General binocular disorders: prevalence in a clinic population. *Ophthalmic and Physiological Optics*. 2001;21(1):70-74. doi:10.1046/j.1475-1313.2001.00540.x
59. Borsting E, Rouse M, Lynn M, Scheiman M, Cotter S CJ, Borsting EJEJ, Rouse MWM, et al. Validity and reliability of the revised convergence insufficiency symptom survey in

- children aged 9 to 18 years. *Optom Vis Sci.* 2003;80(12):832-838. doi:10.1097/00006324-200312000-00014
60. Rouse MW, Borsting EJ, Mitchell GL, et al. Validity and reliability of the revised convergence insufficiency symptom survey in adults. *Ophthalmic and Physiological Optics.* 2004;24(5):384-390. doi:10.1111/j.1475-1313.2004.00202.x
  61. Rouse M, Borsting E, Mitchell G, Cotter S, Kulp M SM. Validity of the convergence insufficiency symptom survey: A confirmatory study. *Optom Vis Sci.* 2009;86:357-363.
  62. Cooper J, Burns C, Cotter S, Griffin J, Scheiman M. *Care of the Patient with Accommodative and Vergence Dysfunction. Optometric Clinical Practice Guideline.;* 2006.
  63. León Álvarez A, Medrano Muñoz SM, Márquez Galvis MM. Rangos de referencia de heteroforias y reservas fusionales entre los 6 y los 60 años de edad. *Ciencia & Tecnología para la Salud Visual y Ocular.* 2017;15(1):47. doi:10.19052/sv.4038
  64. Ganz ML, Xuan Z, Hunter DG. Prevalence and Correlates of Children's Diagnosed Eye and Vision Conditions. *Ophthalmology.* 2006;113(12):2298-2306. doi:10.1016/j.ophtha.2006.06.015
  65. Weiner G. Dictionary of visual science. *Optometry and Vision Science.* 1960;37(10):547-548. doi:10.1097/00006324-196010000-00005
  66. Weissberg E. *Essentials of Clinical Binocular Vision.* Butterworth Heinemann; 2004.
  67. Kulp MT. *Anomalies of Binocular Vision: Diagnosis & Management.* Vol 76. (Mosby L, ed.); 1999. doi:10.1097/00006324-199901000-00013
  68. Hussaindeen JR, Rakshit A, Singh NK, et al. Prevalence of non-strabismic anomalies of binocular vision in Tamil Nadu: report 2 of BAND study. *Clin Exp Optom.* 2017;100(6):642-648. doi:10.1111/cxo.12496
  69. Ma MML, Long W, She Z, et al. Convergence insufficiency in Chinese high school students. 2019;102(2):166-171.
  70. Hashemi H, Nabovati P, Khabazkhoob M, et al. The prevalence of convergence insufficiency in Iran: a population-based study. *Clin Exp Optom.* 2017;100(6):704-709. doi:10.1111/CXO.12522
  71. Carla SC, Sara BF, JM CR, Catalina PÁ. Prevalence of convergence insufficiency among Spanish school children aged 6 to 14 years. *J Optom.* 2022;15(4):278-283. doi:10.1016/j.optom.2021.11.004
  72. Vaughn W, Maples WC, Hoenes R. The association between vision quality of life and academics as measured by the College of Optometrists in Vision Development Quality of Life questionnaire. *Optometry.* 2006;77(3):116-123. doi:10.1016/j.optm.2006.01.004
  73. Martínez PC, Muñoz ÁG, Ruiz-Cantero MT. Treatment of accommodative and nonstrabismic binocular dysfunctions: A systematic review. *Optometry.* 2009;80(12):702-716. doi:10.1016/j.optm.2009.06.011

## REFERENCIAS

74. Scheiman M, Mitchell GL, Cotter S, et al. The convergence insufficiency treatment trial: design, methods, and baseline data. *Ophthalmic Epidemiol.* 2008;15(1):24-36. doi:10.1080/09286580701772037
75. Hussaindeen JR, Rakshit A, Singh NK, et al. The minimum test battery to screen for binocular vision anomalies: report 3 of the BAND study. *Clin Exp Optom.* 2018;101(2):281-287. doi:10.1111/CXO.12628
76. Pickwell y Evans B. *Binocular Vision Anomalies Investigation and Treatment.* Butterworth-Heinemann; 2002.
77. Hamed MM, Goss David A, Marzieh E. The relationship between binocular vision symptoms and near point of convergence. *Indian J Ophthalmol.* 2013;61(7):325-328. doi:10.4103/0301-4738.97553
78. Blake R, Wilson H. Binocular vision. *Vision Res.* 2011;51(7):754-770. doi:10.1016/j.visres.2010.10.009
79. Ansons AM, Davis H. *Diagnosis and Management of Ocular Motility Disorders: Fourth Edition.* (Blackwell W, ed.); 2013. doi:10.1002/9781118712368
80. Ophthalmology OCA of, 2001 undefined. Diagnosis and Management of Ocular Motility Disorders. *jamanetwork.com.*
81. Teitelbaum B, Pang Y, Krall J. Effectiveness of base in prism for presbyopes with convergence insufficiency. *Optom Vis Sci.* 2009;86(2):153-156. doi:10.1097/OPX.0B013E318194E985
82. Pang Y, Teitelbaum B, Krall J. Factors associated with base-in prism treatment outcomes for convergence insufficiency in symptomatic presbyopes. *Clin Exp Optom.* 2012;95(2):192-197. doi:10.1111/j.1444-0938.2011.00693.x
83. Molina M. NP, Forero Mora C. Insuficiencia de convergencia. *Ciencia & Tecnología para la Salud Visual y Ocular.* 2010;8(2):91-102. doi:10.19052/sv.821
84. Santos LH, Santos LRH, Ruiz LVH, et al. Consideraciones actuales en la insuficiencia de convergencia. *Revista Cubana de Oftalmología.* 2013;26(0).
85. Scheiman M, Kulp MT, Cotter SA, Lawrenson JG, Wang L, Li T. Interventions for convergence insufficiency: a network meta-analysis. *Cochrane Database Syst Rev.* 2020;12(12). doi:10.1002/14651858.CD006768.PUB3
86. Schriesheim CA, Novelli L. A comparative test of the interval-scale properties of magnitude estimation and case iii scaling and recommendations for equal-interval frequency response anchors. *Educ Psychol Meas.* 1989;49(1):59-74. doi:10.1177/0013164489491007
87. Carlsson AM. *Assessment of Chronic Pain. I. Aspects of the Reliability and Validity of the Visual Analogue Scale.* Vol 16. Elsevier Biomedical Press; 1983.
88. educativa AMR electrónica de investigación, 2018 undefined. Diseño del formato de escalas tipo Likert: un estado de la cuestión. *scielo.org.mx.*
89. Fink A. *Survey Research Methods.* Butterwort.; 2010. doi:10.1016/B978-0-08-044894-7.00296-7

90. Prieto L, Badia X. Health questionnaire: concept and methodology. *Atencion primaria / Sociedad Española de Medicina de Familia y Comunitaria*. 2001;28(3):201-209. doi:10.1016/S0212-6567(01)78932-9
91. FDA, HHS. Guidance for Industry Use in Medical Product Development to Support Labeling Claims Guidance for Industry. *Clinical/Medical Federal Register*. 2009;(December):1-39.
92. Conlon EGE, Lovegrove WWJ, Chekaluk E, Pattison PPE. Measuring Visual Discomfort. *Vis cogn*. 1999;6(6):637-663. doi:10.1080/135062899394885
93. Marran LF, De Land PN, Nguyen AL. Accommodative insufficiency is the primary source of symptoms in children diagnosed with convergence insufficiency. *Optometry and Vision Science*. 2006;83(5). doi:10.1097/01.opx.0000216097.78951.7b
94. Pesudovs K, Burr JM, Harley C EDB, Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment, and selection of questionnaires. *Optom Vis Sci*. 2007;84(8):663-674. doi:10.1097/OPX.0b013e318141fe75
95. Wild D, Grove A, Martin M, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in Health*. 2005;8(2):94-104.
96. Terwee CB, Prinsen CA, Chiarotto A, et al. COSMIN methodology for assessing the content validity of PROMs: User manual. *Circulation*. 2018;120(9):0-70.
97. Khadka J, Gothwal VK, Mcalinden C, Lamoureux EL, Pesudovs K. The importance of rating scales in measuring patient-reported outcomes. *Health Qual Life Outcomes*. 2012;10(1):80. doi:10.1186/1477-7525-10-80
98. Massof RW. Likert and Guttman scaling of visual function rating scale questionnaires. *Ophthalmic Epidemiol*. 2004;11(5):381-399. doi:10.1080/09286580490888771
99. Guil Bozal M. *ESCALA MIXTA LIKERT-THURSTONE*.
100. Sangrador O, Arias M. Estadística. Tipos de variables. Escalas de medida. *Evid Pediatr*. 2018;14:29.
101. Wright BD LJM. Observations are always ordinal ; measurements, however, must be interval. *Arch Phys Med Rehabil*. 1989;70(12):857-860.
102. Linacre J. Optimizing rating scale category effectiveness. *J Appl Meas*. 2002;3:85-106.
103. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Expanded. (University of Chicago, ed.); 1980.
104. Comparing and Choosing between “Partial Credit Models” (PCM) and “Rating Scale Models” (RSM). Accessed May 13, 2020. <https://www.rasch.org/rmt/rmt143k.htm>
105. Adams RJ, Wilson M, Wang WC. The multidimensional random coefficients multinomial logit model. *Appl Psychol Meas*. 1997;21(1):1-23. doi:10.1177/0146621697211001

## REFERENCIAS

106. Embretson SE, McCollam KMS. Psychometric Approaches to Understanding and Measuring Intelligence. In: *Handbook of Intelligence*. Cambridge University Press; 2012:423-444. doi:10.1017/cbo9780511807947.020
107. Linacre J. A User's Guide to Winsteps: Rasch model Computer Programme. Published online 2008.
108. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences: Second Edition.*; 2007. doi:10.4324/9781410614575
109. Cantú González V, Rodríguez Macías JC. Examen al egreso de la educación básica en el Área de salud (EXEEBAS): Análisis de unidimensionalidad. *Revista Complutense de Educacion*. 2018;29(1):61-77. doi:10.5209/RCED.51935
110. Zhang B, Walker CM. Impact of missing data on person-model fit and person trait estimation. *Appl Psychol Meas*. 2008;32(6):466-479. doi:10.1177/0146621607307692
111. A Critique of Rasch Residual Fit Statistics - PubMed. Accessed May 27, 2020. <https://pubmed.ncbi.nlm.nih.gov/12029176/>
112. B Wright JL. Reasonable mean-square fit values. *Rasch Meas Trans*. 1994;8:370.
113. Smith RM. Fit analysis in latent trait measurement models. *J Appl Meas*. 2000;1(2):199-218.
114. Khadka J, McAlinden C, Pesudovs K. Quality assessment of ophthalmic questionnaires: Review and recommendations. *Optometry and Vision Science*. 2013;90(8):720-744. doi:10.1097/OPX.0000000000000001
115. Khadka J, Huang J, Mollazadegan K, et al. Translation, cultural adaptation, and rasch analysis of the visual function (VF-14) questionnaire. *Invest Ophthalmol Vis Sci*. 2014;55(7):4413-4420. doi:10.1167/iovs.14-14017
116. instrumento | Definición | Diccionario de la lengua española | RAE - ASALE. Accessed May 25, 2020. <https://dle.rae.es/instrumento>
117. González Pérez M. Desarrollo y validación de una escala para medir la sintomatología visual asociada al uso de videoterminals en el trabajo. *Universidad Complutense de Madrid*. Published online 2015:34-38. doi:ISBN: 978-84-693-1123-3
118. López MA, Gabilondo A, Codony M, et al. Adaptation into Spanish of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) and preliminary validation in a student sample. *Quality of Life Research*. 2013;22(5):1099-1104. doi:10.1007/s11136-012-0238-z
119. Maheswaran H, Welch S, Powell J, Stewart-Brown S. Evaluating the responsiveness of the Warwick-Edinburgh mental well being scale (WEMWBS): Group and individual level analysis. *Health Qual Life Outcomes*. 2012;10(1):1-8.
120. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med*. 1993;118(8):622-629. doi:10.7326/0003-4819-118-8-199304150-00009
121. Acquadro C, Conway K, Hareendran A, Aaronson N. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value in Health*. 2008;11(3):509-521. doi:10.1111/j.1524-4733.2007.00292.x

122. Gjersing L, Caplehorn JR, Clausen T. Cross-cultural adaptation of research instruments: Language, setting, time and statistical considerations. *BMC Med Res Methodol*. 2010;10. doi:10.1186/1471-2288-10-13
123. Lauffer A, Solé L, Bernstein S, Lopes MH, Francisconi CF. Practical aspects for minimizing errors in the cross-cultural adaptation and validation of quality of life questionnaires. *Rev Gastroenterol Mex*. 2013;78(3):159-176. doi:10.1016/j.rgmx.2013.01.008
124. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. *Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures*. Vol 25.; 2000. doi:10.1097/00007632-200012150-00014
125. Acquadro C, Conway K, Hareendran A, Aaronson N. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value in Health*. 2008;11(3):509-521. doi:10.1111/j.1524-4733.2007.00292.x
126. Adams R, Wu M. *Rasch Model-Multiple Choice Test Modelling a Dichotomously Scored Multiple Choice Test with the Rasch Model.*; 2010.
127. González-Pérez M, Susi R, Barrio A, Antona B. Five levels of performance and two subscales identified in the computer-vision symptom scale (CVSS17) by Rasch, factor, and discriminant analysis. *PLoS One*. 2018;13(8):1-16. doi:10.1371/journal.pone.0202173
128. Pesudovs K, Garamendi E, Elliot DB. The Quality of Life Impact of Refractive Error Correction (QIRC) Questionnaire; development and validation. *Optometry and Vision Science*. 2004;81(10):769-777.
129. Mallinson T. Why measurement matters for measuring patient vision outcomes. *Optom Vis Sci*. 2007;84(8):675-682. doi:10.1097/OPX.0b013e3181339f44
130. LERÍS LÓPEZ D, VEA MUNIESA F, VELAMAZÁN GIMENO Á. Aprendizaje adaptativo en Moodle: tres casos prácticos. *Education in the Knowledge Society (EKS)*. 2015;16(4):138. doi:10.14201/eks201516138157
131. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986;323(6088):533-536. doi:10.1038/323533a0
132. Vyas V, Uma V. An Extensive study of Sentiment Analysis tools and Binary Classification of tweets using Rapid Miner. In: *Procedia Computer Science*. Vol 125. Elsevier B.V.; 2018:329-335. doi:10.1016/j.procs.2017.12.044
133. Willett P. The Porter stemming algorithm: Then and now. *Program*. 2006;40(3):219-223. doi:10.1108/00330330610681295
134. Yang D, Alsadoon A, Prasad PWC, Singh AK, Elchouemi A. An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment. In: *Procedia Computer Science*. Vol 125. Elsevier B.V.; 2018:2-10. doi:10.1016/j.procs.2017.12.003
135. Sheedy JE, Saladin JJ. Association of symptoms with measures of oculomotor deficiencies. *Optometry and Vision Science*. 1978;55(10):670-676. doi:10.1097/00006324-197810000-00002

## REFERENCIAS

136. Garamendi E, Pesudovs K, Stevens MJ, Elliott DB. The Refractive Status and Vision Profile: Evaluation of psychometric properties and comparison of Rasch and summated Likert-scaling. *Vision Res.* 2006;46(8-9):1375-1383. doi:10.1016/j.visres.2005.07.007
137. IBM Corporation. IBM SPSS Statistics 25 Core User's Guide. Published online 2017:300.
138. Patrick A, Hatzinger R, Maier MJ, Rusch T, Debelak R, Mair MP. Package 'eRm.'. Published online 2020.
139. Patrick Meyer J. *Applied Measurement with Jmetrik.*; 2014. doi:10.4324/9780203115190
140. CCDC. Conquest User Guide and Tutorials 2016 CSDS Release. 2016;(800579).
141. Lamoureux EL, Pallant JF, Pesudovs K, Hassell JB, Keeffe JE. The impact of vision impairment questionnaire: An evaluation of its measurement properties using Rasch analysis. *Invest Ophthalmol Vis Sci.* 2006;47(11):4732-4741. doi:10.1167/iovs.06-0220
142. Pesudovs K, Garamendi E, Elliot DB. The Quality of Life Impact of Refractive Error Correction (QIRC) Questionnaire; development and validation. *Optometry and Vision Science.* 2004;81(10):769-777.
143. An Adjustment for Sample Size in DIF Analysis. Accessed February 3, 2024. <https://www.rasch.org/rmt/rmt203e.htm>
144. Differential Item Functioning DIF Sample Size Nomogram. Accessed February 3, 2024. <https://www.rasch.org/rmt/rmt264a.htm>
145. McAlinden C, Khadka J, Pesudovs K. Precision (repeatability and reproducibility) studies and sample-size calculation. *J Cataract Refract Surg.* 2015;41(12):2598-2604. doi:10.1016/j.jcrs.2015.06.029
146. Marrugat. GRANMO: sample size and power calculator. *Institut Municipal d'Investigació Mèdica, Barcelona, Spain.* Published online 2012.
147. Model selection: Rating Scale Model (RSM) or Partial Credit Model (PCM)? Accessed May 13, 2020. <https://www.rasch.org/rmt/rmt1231.htm>
148. Andrich D. Rating scales and Rasch measurement. *Expert Rev Pharmacoecon Outcomes Res.* 2011;11(5):571-585. doi:10.1586/erp.11.59
149. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16(3):297-334. doi:10.1007/BF02310555
150. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Rasch analysis of the quality of life and vision function questionnaire. *Optometry and Vision Science.* 2009;86(7):E836-E844. doi:10.1097/OPX.0b013e3181ae1ec7
151. BD Wright MM. Rasch models overview. *J Appl Meas.* 2000;1:83-106.
152. Reliability, Separation, Strata Statistics. Accessed May 18, 2020. <https://www.rasch.org/rmt/rmt63i.htm>
153. Number of Person or Item Strata:  $(4 * \text{Separation} + 1) / 3$ . Accessed May 17, 2020. <https://www.rasch.org/rmt/rmt163f.htm>

154. Reliability and Separation. Accessed May 18, 2020. <https://www.rasch.org/rmt/rmt94n.htm>
155. Reliability and Separation Nomograms. Accessed May 18, 2020. <https://www.rasch.org/rmt/rmt92a.htm>
156. Separation, Reliability and Skewed Distributions: Statistically Different Levels of Performance. Accessed May 18, 2020. <https://www.rasch.org/rmt/rmt144k.htm>
157. González-Pérez M, Pérez-Garmendia C, Barrio AR, García-Montero M, Antona B. Spanish Cross-Cultural Adaptation and Rasch Analysis of the Convergence Insufficiency Symptom Survey (CISS). *Transl Vis Sci Technol.* 2020;9(4):23. doi:10.1167/tvst.9.4.23
158. Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Eval Health Prof.* 2005;28(2):212-232. doi:10.1177/0163278705275342
159. Sperber AD. Translation and Validation of Study Instruments for Cross-Cultural Research. In: *Gastroenterology*. Vol 126. W.B. Saunders; 2004. doi:10.1053/j.gastro.2003.10.016
160. Silvia M, Ester M, Pablo FJ, et al. Análisis del DIF en una Escala de Voluntad de Trabajo mediante los procedimientos de Mantel- Haenszel y Breslow-Day , regresión logística y el criterio de Mantel-Haenszel y Breslow Day. Published online 2006.
161. Zwick R, Thayer DT, Wingersky M. A simulation study of methods assessing differential item functioning in Computer-Adaptive Tests. *ETS*. Published online 1993.
162. Cómo minimizar errores al realizar la adaptación transcultural y la validación de los cuestionarios sobre calidad de vida: aspectos prácticos - ScienceDirect. Accessed April 16, 2020. <https://www.sciencedirect.com/science/article/pii/S0375090613000529>
163. The Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS). Accessed May 28, 2020. <https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/>
164. Cardemil F. Análisis de comparación y aplicaciones del método de Bland-Altman: ¿concordancia o correlación? *Medwave.* 2017;17(01):e6852. doi:10.5867/MEDWAVE.2017.01.6852
165. Borsting EJ, Rouse MW, Mitchell GL, et al. Validity and Reliability of the Revised Convergence Insufficiency Symptom Survey in Children Aged 9 to 18 Years. *Optometry and Vision Science.* 2003;80(12):832-838. doi:10.1097/00006324-200312000-00014
166. Praena Crespo M, Lora Espinosa A, Aquino Linares N, Sánchez Sánchez AM, Jiménez Cortés A. Versión española del NAKQ. Adaptación transcultural y análisis de fiabilidad y validez. *An Pediatr (Engl Ed).* 2009;70(3):209-217. doi:10.1016/j.anpedi.2008.10.013
167. BOE.es - Documento consolidado BOE-A-1999-23750. Accessed May 29, 2020. <https://www.boe.es/buscar/act.php?id=BOE-A-1999-23750>
168. Ebel RL. The reliability of an index of item discrimination. *Educ Psychol Meas.* Published online 1951. doi:10.1177/001316445101100311

## REFERENCIAS

169. Arenas Mejía C, Tavera Perez IS. Validez del cuestionario CISS-V15 para el diagnóstico de la insuficiencia de convergencia. *Revista Investigación en Salud Universidad de Boyacá*. 2016;3(2):127. doi:10.24267/23897325.183
170. Gothwal VK, Wright T, Lamoureux EL, Pesudovs K. Psychometric properties of visual functioning index using Rasch analysis. *Acta Ophthalmol*. 2010;88(7):797-803. doi:10.1111/j.1755-3768.2009.01562.x
171. Scott NW, Fayers PM, Bottomley A, et al. Comparing translations of the EORTC QLQ-C30 using differential item functioning analyses. *Quality of Life Research*. 2006;15(6):1103-1115. doi:10.1007/s11136-006-0040-x
172. Robitzsch A, Rupp AA. Impact of Missing Data on the Detection of Differential Item Functioning. *Educ Psychol Meas*. 2009;69(1):18-34. doi:10.1177/0013164408318756
173. Siegel, S. & Castellan-Jr. NJ. *Nonparametric Statistics for the Behavioral Sciences*, International Edition. Published online 1988:262-272.
174. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2--assessing respondent understanding. *Value in Health*. 2011;14(8).
175. Gandek B, Ware JE. Methods for validating and norming translations of health status questionnaires: The IQOLA Project approach. *J Clin Epidemiol*. 1998;51(11):953-959. doi:10.1016/S0895-4356(98)00086-9
176. Bradley C, Massof RW. Validating Translations of Rating Scale Questionnaires Using Rasch Analysis. *Ophthalmic Epidemiol*. 2017;24(1):1-2. doi:10.1080/09286586.2016.1246667
177. Logan NS, Gilmartin B. School vision screening, ages 5-16 years: The evidence-base for content, provision and efficacy. *Ophthalmic and Physiological Optics*. 2004;24(6):481-492. doi:10.1111/j.1475-1313.2004.00247.x
178. Horan L, Ticho BH, Allen M. Is the convergence insufficiency symptoms survey specific for convergence insufficiency? *Journal of American Association for Pediatric Ophthalmology and Strabismus*. 2013;17(1):e17. doi:10.1016/j.jaapos.2012.12.061
179. Tavares C, Nunes AMMF, Nunes AJS, Pato MV, Monteiro PML. Translation and validation of Convergence Insufficiency Symptom Survey (CISS) to Portuguese - psychometric results. *Arq Bras Oftalmol*. 2014;77(1):21-24. doi:10.5935/0004-2749.20140007
180. Nunes ASAF, Monteiro PML, Ferreira FBP, Nunes ASAF. Convergence insufficiency and accommodative insufficiency in children. *BMC Ophthalmol*. 2019;19(1):1-8. doi:10.1186/s12886-019-1061-x
181. Nunes AFAS, Monteiro PL, Nunes AFAS. Factor structure of the convergence insufficiency symptom survey questionnaire. Awadein A, ed. *PLoS One*. 2020;15(2):e0229511. doi:10.1371/journal.pone.0229511

182. Horwood AMA, Toor S, Riddell PPM. Screening for convergence insufficiency using the CISS is not indicated in young adults. *British Journal of Ophthalmology*. 2014;98(5):679-683. doi:10.1136/bjophthalmol-2013-304533
183. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Activities of daily vision scale: What do the subscales measure? *Invest Ophthalmol Vis Sci*. 2010;51(2):694-700. doi:10.1167/iops.09-3448
184. Hongbo D. Data Mining Techniques and Applications: An Introduction. Published online 2010.
185. Sudrajat R, Irianingsih I, Krisnawan D. Analysis of data mining classification by comparison of C4.5 and ID algorithms. *IOP Conf Ser Mater Sci Eng*. 2017;166(1). doi:10.1088/1757-899X/166/1/012031
186. Trees D. Machine Learning , Machine Learning 10-601. *Mach Learn*. 2009;2012(April):1-16.
187. Mitchell TM (Tom M. Machine Learning (McGraw-Hill International Editions Computer Science Series): Tom M. Mitchell: 9780071154673. Published online 1997:414.
188. Briganti G, Le Moine O. Artificial Intelligence in Medicine: Today and Tomorrow. *Front Med (Lausanne)*. Published online 2020. doi:10.3389/fmed.2020.00027
189. Song YY, Lu Y. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*. Published online 2015. doi:10.11919/j.issn.1002-0829.215044



# ANEXOS





## 11 ANEXOS

## 11.1 ANEXO I – GUÍA DE KHADKA ADAPTADA

Propiedades psicométricas	Definición	Puntuación	Criterio de calidad
<b>Categorías de respuesta</b>	El grado en que las categorías de respuesta utilizadas para cuantificar los ítems siguen un orden lógico. La distancia entre los umbrales que separan estas categorías debe estar entre 1.4 y 5 logits (categorías equidistantes).	A B C	Las categorías aparecen ordenadas y equidistantes Las categorías aparecen ordenadas, pero no son equidistantes Las categorías están desordenadas
<b>Dimensionalidad</b>	El grado en que el instrumento mide un único concepto. Análisis de las componentes principales basado en dos parámetros: la cantidad de varianza explicada por la medida y el autovalor de la varianza desconocida en el primer contraste.	A B C	Varianza explicada por la medida > 60% y autovalor de primer contraste $\leq 2.0$ Varianza explicada por la medida entre 50% y 60% y autovalor de primer contraste $\leq 2.0$ Varianza explicada por la medida < 50%, autovalor > 2. Indicios de multidimensionalidad
<b>Precisión</b>	La capacidad del instrumento para distinguir distintos grupos de capacidad. Se cuantifica mediante el PSI o el coeficiente de fiabilidad $\alpha$ (valor mínimo aceptable: PSI=2.0 ó fiabilidad: 0.8)	A B C	PSI $\geq 2.50$ , $\alpha \geq 0.85$ $2.00 \leq \text{PSI} \leq 2.49$ , $0.80 \leq \alpha \leq 0.85$ PSI < 2.00, $\alpha < 0.80$
<b>Estadísticas de ajuste de los ítems</b>	El nivel de ajuste entre los datos recogidos por los ítems del instrumento y lo previsto por el modelo de Rasch. Se emplean dos estadísticos de ajuste: El ajuste interno (Infit) y el ajuste externo (Outfit. Un valor de alguno de ellos fuera del intervalo (0.5, 1.5) se considera no aceptable.	A B C	Todos los ítems con Infit y Outfit entre 0.7 y 1.3 Uno o dos ítems con Infit o Outfit fuera del intervalo (0.5,1,5) Más de dos ítems con Infit o Outfit fuera del intervalo (0.5,1,5)
<b>Funcionamiento diferencial de los ítems (DIF)</b>	Indica hasta que punto los niveles de capacidad de los diferentes subgrupos de la población de interés difieren para un determinado ítem (<0.5 logits, insignificante; entre 0.50 y 1.0, leve; >1.0, notable)	A B C	Todos los ítems con DIF < 0.5 logits Algunos ítems con DIF entre 0.5 y 1.0. Máximo 1 ítem con DIF>1.0 Más de 1 ítem con DIF>1.0
<b>Acoplamiento o Ajuste ítem-sujeto</b>	El grado en que la dificultad de los ítems se ajusta a las capacidades de los participantes. Una diferencia de más de un logit se considera significativa	A B C	$\leq 1$ logit Diferencia entre 1 y 2 logits $\geq 2$ logit
<b>Validez</b>			
<b>Validez convergente</b>	Correlación con otro instrumento que mida un constructo similar	A B C	Correlación con un instrumento relacionado entre 0.3 y 0.9 Correlación con un instrumento relacionado entre 0.2 y 0.3 Correlación con un instrumento relacionado < 0.2
<b>Validez divergente</b>	Correlación con otro instrumento que mide un constructo diferente	A B C	Correlación con otro instrumento < 0.1 Correlación con otro instrumento entre 0.1 y 0.2 Correlación con otro instrumento < 0.2
<b>Fiabilidad</b>			
<b>Repetitividad</b>	El grado de estabilidad temporal que ha demostrado el instrumento tras ser administrado en dos ocasiones. Se considera una buena fiabilidad test re-test un ICC $\geq 0.8$	A B C	

A=alto, B=medio, C=bajo.

## 11.2 ANEXO II – INFORME FAVORABLE DEL COMITÉ ÉTICO DE INVESTIGACIÓN CLÍNICA DEL HOSPITAL CLÍNICO SAN CARLOS

 Hospital Clínico San Carlos		<b>Informe Dictamen Protocolo Favorable</b> C.P. - C.I. 16/287-E 14 de julio de 2016
---	---	--

**CEIC Hospital Clínico San Carlos**

Dra. Mar García Arenillas  
Presidenta del CEIC Hospital Clínico San Carlos

**CERTIFICA**

Que el CEIC Hospital Clínico San Carlos en su reunión del día 06/07/2016, acta 7.1/16 ha evaluado la propuesta del promotor/investigador referida al estudio:


**Título: "Traducción al español y validación del Cuestionario Convergence Insufficiency Symptom Survey (CISS V-15)".**

Que en este estudio:

- o Se cumplen los requisitos necesarios de idoneidad del protocolo en relación con los objetivos del estudio y están justificados los riesgos y molestias previsibles para el sujeto.
- o Es adecuado el procedimiento para obtener el consentimiento informado.
- o La capacidad del investigador y los medios disponibles son adecuados para llevar a cabo el estudio.
- o El alcance de las compensaciones económicas previstas no interfiere con el respeto de los postulados éticos.
- o Se cumplen los preceptos éticos formulados en la Declaración de Helsinki de la Asociación Médica mundial sobre principios éticos para las investigaciones médicas en seres humanos y en sus posteriores revisiones, así como aquellos exigidos por la normativa legal aplicable en función de las características del estudio.

Es por ello que el Comité **informa favorablemente** sobre la realización de dicho proyecto por la Dra. Beatriz Antona Peñalba y el Dr. Carlos Pérez Garmendia, como investigadores en la Facultad de Óptica y Optometría de la Universidad Complutense de Madrid.

Lo que firmo en Madrid, a 14 de julio de 2016

  
**Dra. Mar García Arenillas**  
Presidenta del CEIC Hospital Clínico San Carlos


---

Hospital Clínico San Carlos  
Doctor Martín Lagos, s/n. Madrid 28040 Madrid España  
Tel. 91 330 34 13 Fax. 91 330 32 99 Correo electrónico ceic.hcsc@salud.madrid.org

Página 1 de 1



### 11.3.2. ASSENTIMIENTO INFORMADO DIRIGIDO A ADULTOS JÓVENES PARTICIPANTES



FACULTAD DE ÓPTICA Y OPTOMETRÍA  
UNIVERSIDAD COMPLUTENSE DE MADRID  
C/ ARCOS DE JALÓN 118 - 28037 MADRID - ESPAÑA

---

**ASSENTIMIENTO INFORMADO PARA PARTICIPANTES DE 12 A 17 AÑOS**

**Traducción al Español y validación del Cuestionario *Convergence Insufficiency Symptom Survey (CISS)*.**

**Descripción:**

Has sido invitado a participar en una investigación sobre los síntomas visuales asociados a tareas de visión próxima. Esta investigación es llevada a cabo por el Optico-Optometrista Carlos Pérez Garmendia como parte de la tesis doctoral que realiza en la Universidad Complutense de Madrid.

El propósito de esta investigación es realizar la traducción al español y adaptación cultural de un cuestionario llamado CISS que permita evaluar los síntomas visuales asociados a la realización de tareas en cerca. También se quiere evaluar la calidad del cuestionario traducido.

Si aceptas participar en esta investigación, tu tarea consistirá en rellenar un breve cuestionario y en el caso de que se te incluya en el estudio de falibilidad, volverías a rellenarlo una semana después. El cuestionario se te entregará impreso en tu centro de estudios o bien en formato electrónico accediendo a una dirección web preparada para este fin. Participar en este estudio le llevará aproximadamente 10 minutos.

**Riesgos y beneficios:**

Al no realizarse ninguna intervención ni emplearse ningún fármaco, no existen riesgos asociados a tu participación en el estudio. Tu participación en esta investigación puede tener como beneficio asociado la detección de una posible alteración de tu visión binocular. Si tu puntuación en el cuestionario CISS fuera elevada respecto a lo normal se te indicará la conveniencia de realizar un examen visual completo en la Clínica Universitaria de Optometría de la UCM.

**Confidencialidad:**

Tu identidad será protegida mediante un sistema de codificación que permitirá que sólo las personas que realizan el estudio tengan acceso a la información y, en ningún caso, se cederá a terceros cumpliendo en todo momento con lo estipulado por la Ley Orgánica 15/99 de 13 de diciembre de Protección de Datos de Carácter Personal. Debes saber que tiene derecho de acceso, rectificación y cancelación de los mismos en cualquier momento.

**Derechos:**

**Entiende que tu participación es completamente voluntaria y que tienes derecho a no querer participar o retirarte del estudio en cualquier momento, sin ninguna penalización. Tienes derecho a no contestar alguna pregunta en particular.**

Tu firma en este documento significa que quieres participar en esta investigación después de haber leído y discutido la información presentada en esta hoja de consentimiento.

Entiendo y acepto mi participación en este estudio de investigación.

_____	_____	_____
(Nombre del participante)	(Firma)	(Fecha)

He explicado los contenidos del presente consentimiento, atendiendo a sus dudas y explicando los riesgos y beneficios que conlleva su participación.

_____	_____	_____
Carlos Pérez Garmendia	(Firma)	(Fecha)

1

## 11.4 ANEXO IV – TABLAS Y GRÁFICOS AUXILIARES. DESCRIPTIVOS DE LA ESCALA CISSve

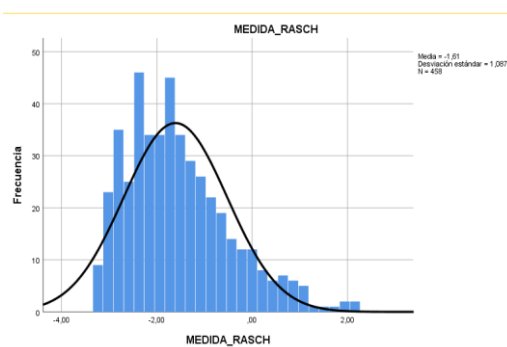
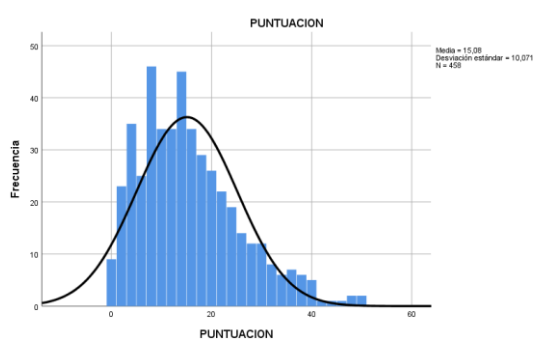
### 11.4.1 DESCRIPTIVOS DE LA VARIABLE PUNTUACIÓN EN LA ESCALA CISSve

Estadísticos			MEDIDA_RASCH	PUNTUACION
N	Válido		458	458
	Perdidos		0	0
Media			-1,6100	15,08
Error estándar de la media			,05077	,471
Mediana			-1,7789 <sup>a</sup>	13,51 <sup>a</sup>
Moda			-2,91	3
Desv. Desviación			1,08663	10,071
Varianza			1,181	101,419
Asimetría			,832	,832
Error estándar de asimetría			,114	,114
Curtosis			,447	,447
Error estándar de curtosis			,228	,228
Rango			5,40	50
Mínimo			-3,24	0
Máximo			2,16	50
Percentiles	10		-2,9019 <sup>b</sup>	3,10 <sup>b</sup>
	20		-2,5592	6,28
	30		-2,3082	8,61
	40		-2,0231	11,25
	50		-1,7789	13,51
	60		-1,5179	15,93
	70		-1,1897	18,97
	80		-,7707	22,85
	90		-,0770	29,28

a. Se ha calculado a partir de datos agrupados.

b. Los percentiles se calculan a partir de datos agrupados.

Estadísticos descriptivos de la variable “puntuación”. Unidades en “logits” y en “puntos”



Representación gráfica de la distribución de la variable “puntuación”. Unidades en “puntos” y “logits”

Prueba de Kolmogorov-Smirnov para una muestra		
		MEDIDA_RAS CH
N		458
Parámetros normales <sup>a,b</sup>	Media	-1,6100
	Desv. Desviación	1,08663
Máximas diferencias extremas	Absoluto	,091
	Positivo	,091
	Negativo	-,067
Estadístico de prueba		,091
Sig. asintótica(bilateral)		,000 <sup>c</sup>
Significación exacta (bilateral)		,001
Probabilidad en el punto		,000

a. La distribución de prueba es normal.  
 b. Se calcula a partir de datos.  
 c. Corrección de significación de Lilliefors.

Resultados de la prueba de Kolmogorov-Smirnov realizados con SPSS v25

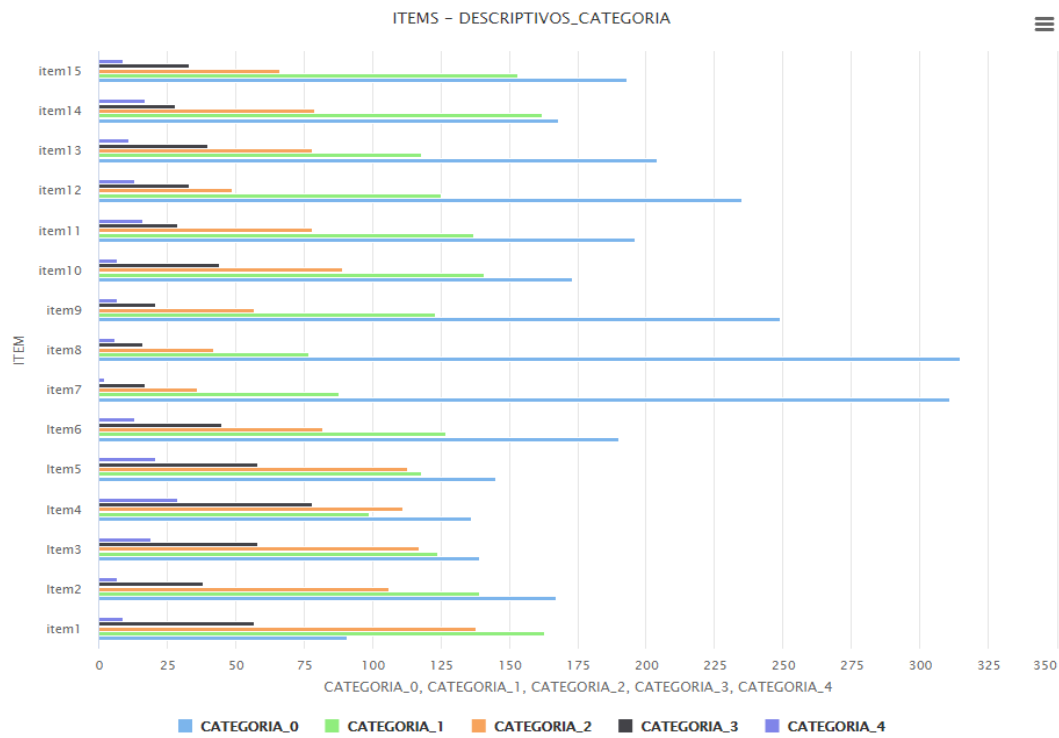
### 11.4.2 DESCRIPTIVOS DE LOS ITEMS DE LA ESCALA CISSv6

		Estadísticos														
		ITEM1	ITEM2	ITEM3	ITEM4	ITEM5	ITEM6	ITEM7	ITEM8	ITEM9	ITEM10	ITEM11	ITEM12	ITEM13	ITEM14	ITEM15
N	Válido	458	457	457	453	455	457	454	456	457	454	456	455	451	454	454
	Perdidos	0	1	1	5	3	1	4	2	1	4	2	3	7	4	4
Media		1,41	1,08	1,33	1,48	1,32	1,05	,48	,51	,72	1,06	,97	,82	,97	1,04	,93
Error estándar de la media		,047	,048	,054	,059	,055	,052	,039	,042	,045	,049	,051	,050	,052	,050	,048
Mediana		1,38 <sup>a</sup>	,95 <sup>a</sup>	1,23 <sup>a</sup>	1,39 <sup>a</sup>	1,20 <sup>a</sup>	,84 <sup>a</sup>	,36 <sup>a</sup>	,36 <sup>a</sup>	,56 <sup>a</sup>	,89 <sup>a</sup>	,78 <sup>a</sup>	,61 <sup>a</sup>	,77 <sup>a</sup>	,87 <sup>a</sup>	,75 <sup>a</sup>
Moda		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Desv. Desviación		1,004	1,031	1,156	1,258	1,179	1,116	,829	,897	,956	1,050	1,083	1,069	1,096	1,062	1,018
Varianza		1,008	1,064	1,336	1,582	1,391	1,246	,687	,804	,914	1,103	1,173	1,142	1,201	1,129	1,036
Asimetría		,319	,636	,478	,334	,501	,831	1,796	1,858	1,332	,717	1,022	1,284	,912	,986	1,021
Error estándar de asimetría		,114	,114	,114	,115	,114	,114	,115	,114	,114	,115	,114	,114	,115	,115	,115
Curtosis		-,512	-,409	-,683	-1,004	-,705	-,240	2,678	2,955	1,248	-,369	,369	,864	-,099	,461	,408
Error estándar de curtosis		,228	,228	,228	,229	,228	,228	,229	,228	,228	,229	,228	,228	,229	,229	,229
Rango		4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Mínimo		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Máximo		4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Percentiles	10	,00 <sup>b</sup>	,.b,c	,.b,c	,.b,c	,.b,c	,.b,c	,.b,c	,.b,c	,.b,c	,.b,c	,.b,c	,.b,c	,.b,c	,.b,c	,.b,c
	20	,.36	,.05	,.17	,.19	,.14	.	.	.	.	,.03	.	.	.	,.04	.
	30	,.72	,.35	,.51	,.58	,.49	,.27	.	.	,.07	,.32	,.23	,.11	,.21	,.32	,.23
	40	1,07	,.65	,.86	,.96	,.83	,.55	,.13	,.13	,.31	,.61	,.51	,.36	,.49	,.59	,.49
	50	1,38	,.95	1,23	1,39	1,20	,.84	,.36	,.36	,.56	,.89	,.78	,.61	,.77	,.87	,.75
	60	1,68	1,31	1,61	1,82	1,60	1,20	,.59	,.59	,.80	1,25	1,08	,.86	1,08	1,19	1,03
	70	1,98	1,68	1,99	2,28	1,99	1,64	,.81	,.83	1,10	1,65	1,51	1,24	1,54	1,57	1,44
	80	2,45	2,09	2,50	2,76	2,52	2,12	1,13	1,19	1,61	2,07	1,93	1,76	2,00	1,95	1,86
	90	2,91	2,73	3,06	3,42	3,11	2,84	1,86	1,96	2,28	2,75	2,72	2,61	2,76	2,73	2,60

a. Se ha calculado a partir de datos agrupados.  
 b. Los percentiles se calculan a partir de datos agrupados.  
 c. El límite inferior del primer intervalo o el límite superior del último intervalo no se conoce. Algunos percentiles no están definidos.

Estadísticos descriptivos de los ítems expresados en “logits”

### 11.4.3 COMPORTAMIENTO DE LAS CATEGORÍAS DE LA ESCALA CISSVe



Representación de las respuestas observadas a cada categoría en cada ítem

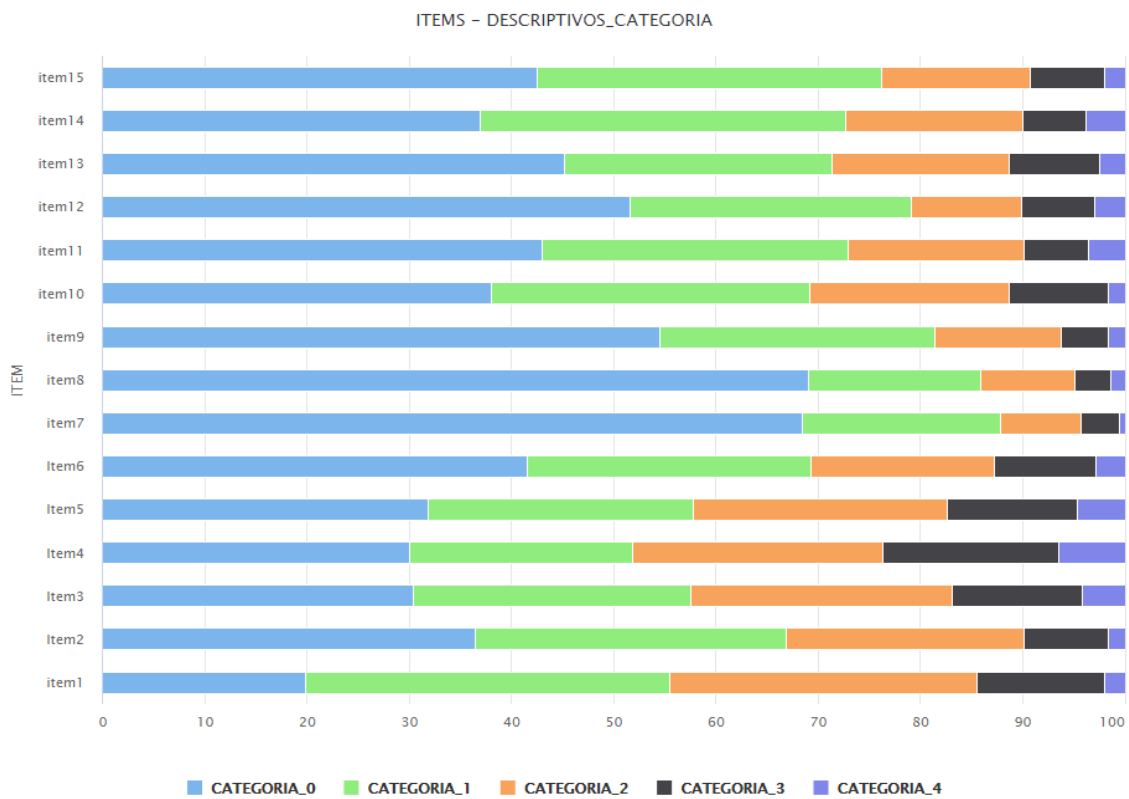


Gráfico de barras que muestra la distribución porcentual de respuestas en cada categoría para cada ítem

ANEXOS

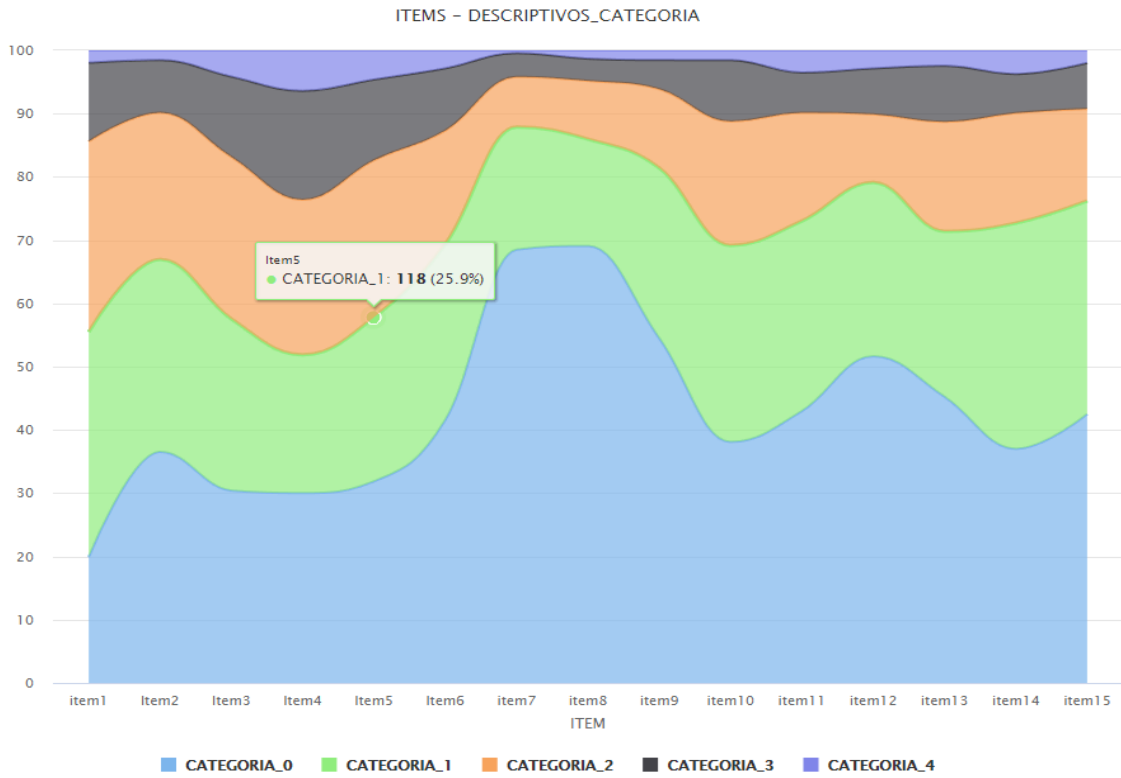


Gráfico de áreas apiladas que muestra la distribución porcentual de respuestas en cada categoría para cada ítem

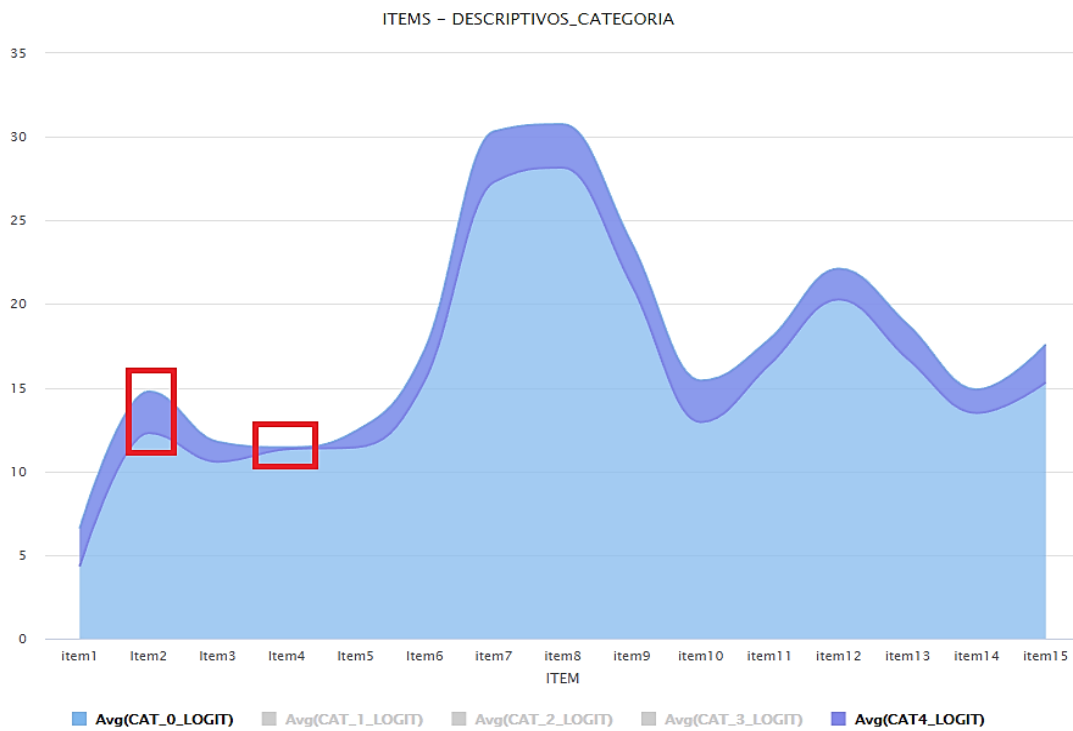


Gráfico que muestra la diferencia en la distribución de respuestas entre la categoría más alta y la más baja para cada ítem

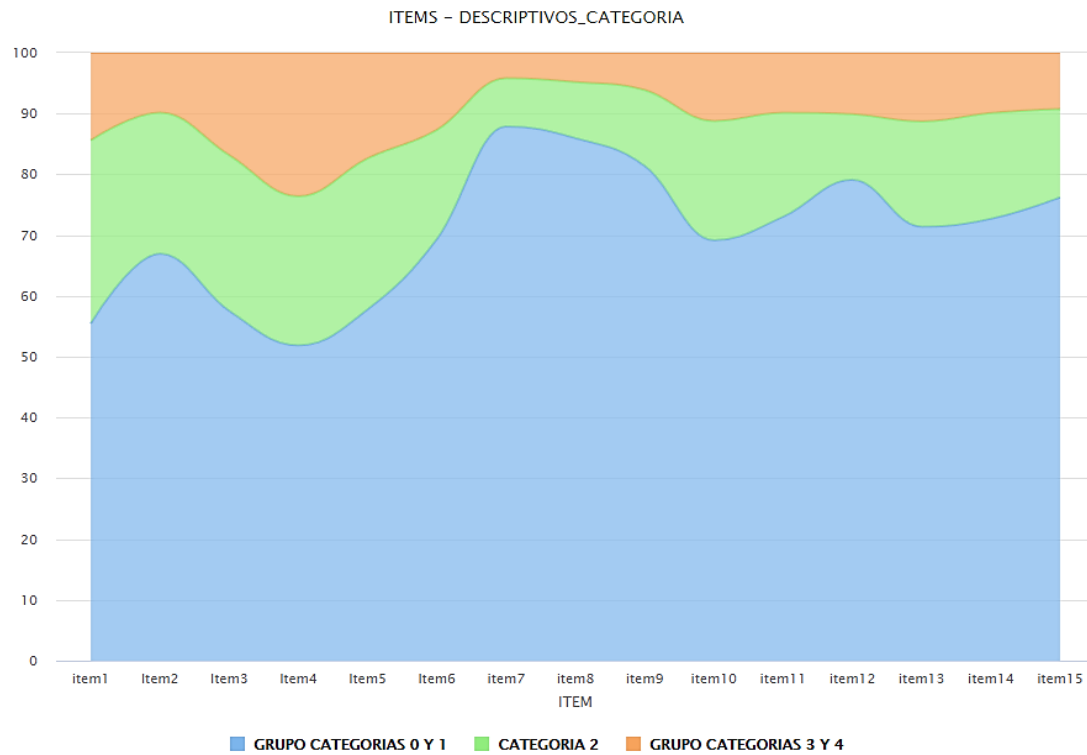
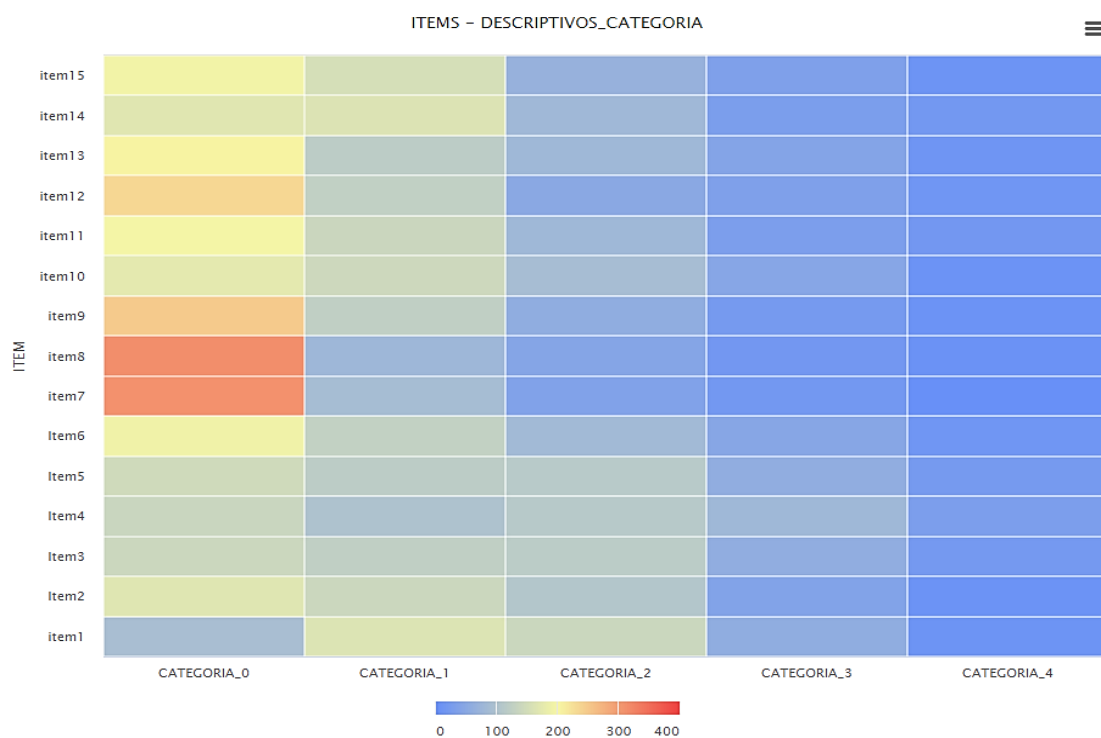
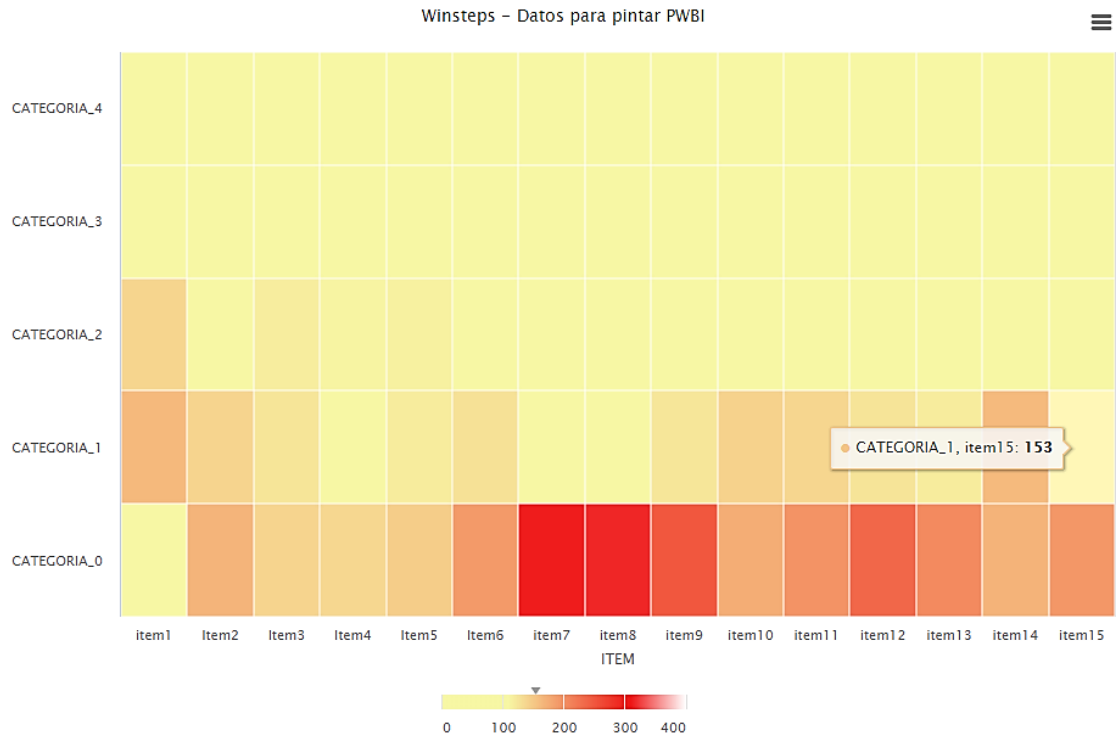


Gráfico que agrupa las dos categorías más bajas (menos sintomatología) y las dos más altas (más sintomatología) y muestra la distribución de respuestas para cada ítem en esas categorías redefinidas



Representación gráfica de la densidad de respuestas en cada categoría para cada ítem



Representación gráfica de la densidad de respuestas a cada ítem en cada categoría

## 11.4.4 DISCRIMINACIÓN DE LOS ÍTEMS DE la ESCALA CISSve

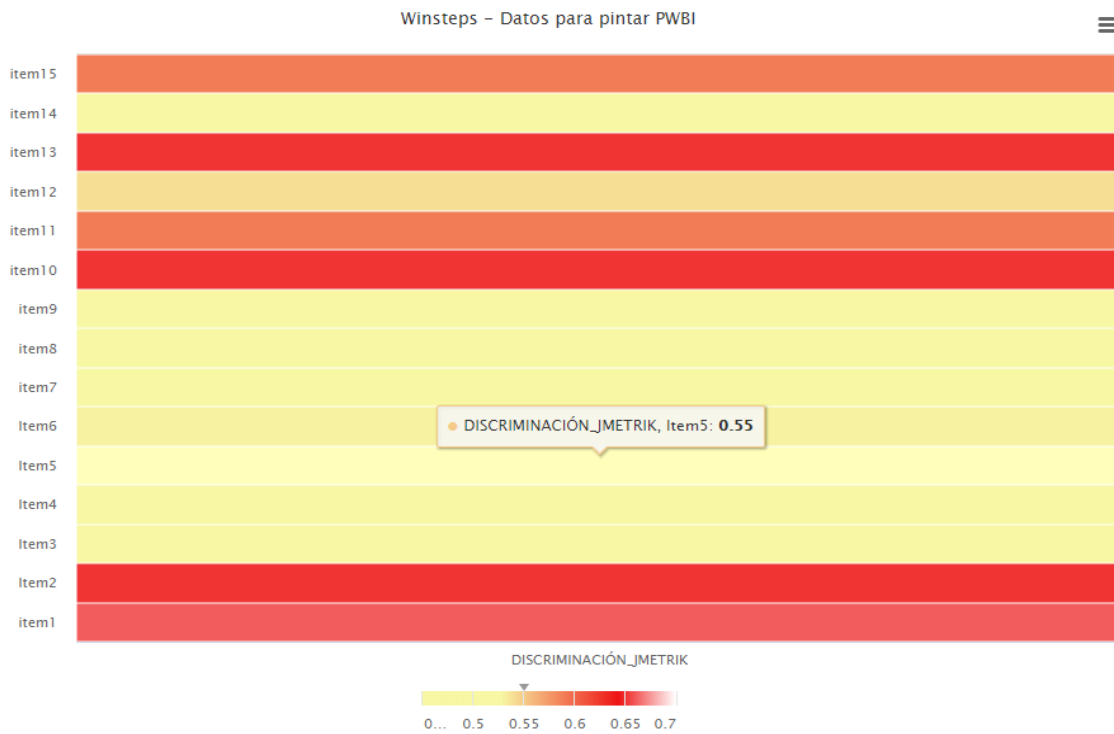


Gráfico que muestra el poder discriminatorio de cada ítem (cálculos realizados con Jmetrik)

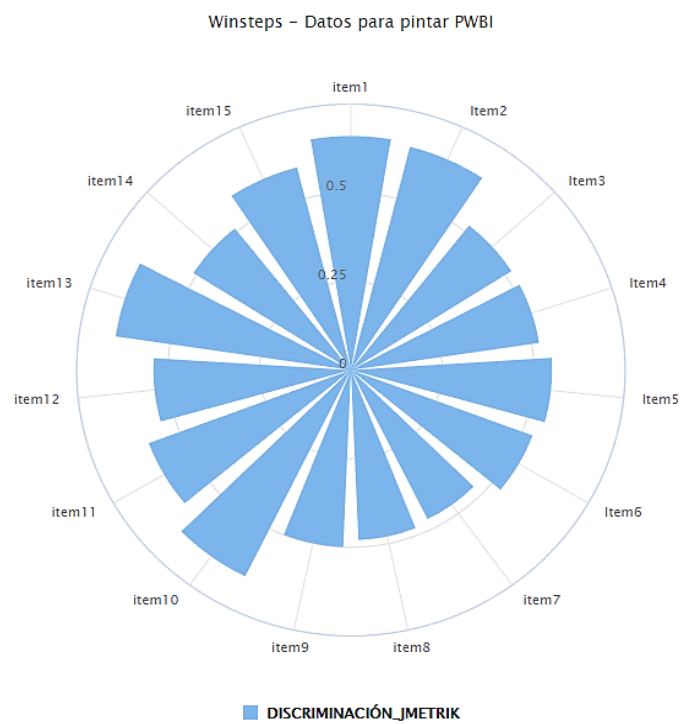


Gráfico radial (centro 0 y radio 0,75) que muestra el poder discriminatorio de cada ítem (cálculos realizados con Jmetrik)

Winsteps - Datos para pintar PWBI

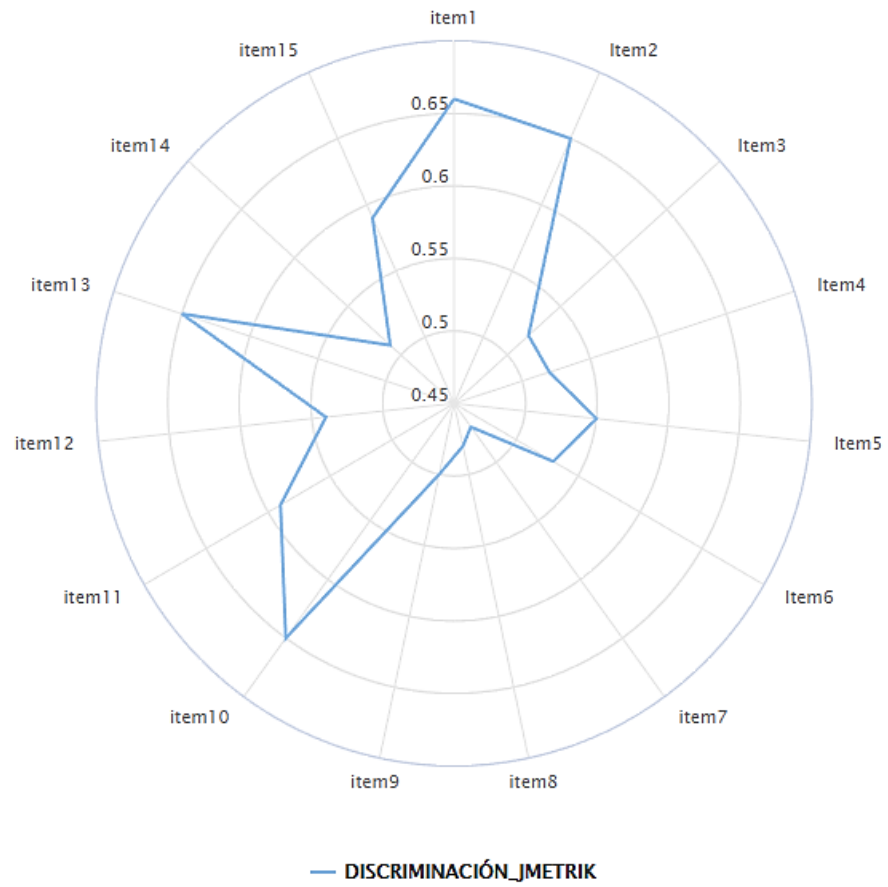


Gráfico radial (centro 0.45 y radio 0.7) que muestra el poder discriminatorio de cada ítem (cálculos realizados con Jmetrik)

## 11.4.5 DIFICULTAD DE LOS ÍTEMS DE LA ESCALA CISSVe

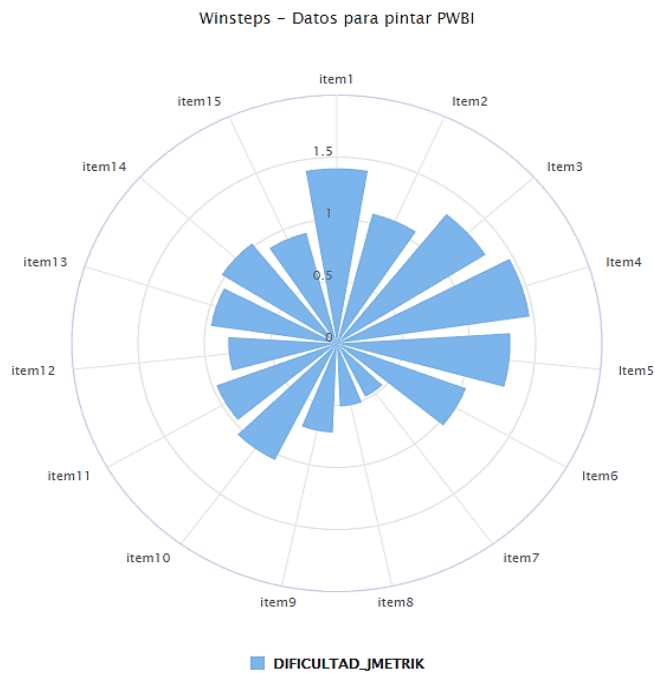
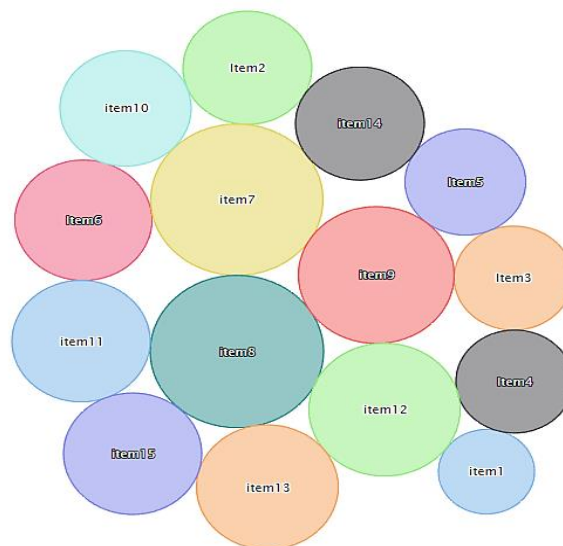


Gráfico radial (centro 0 y radio -2) que muestra la dificultad de cada ítem (cálculos realizados con Jmetrik) – Nótese que los valores de dificultad son tanto más grandes cuánto más fácil es que los sujetos contesten en categorías altas (ítems más fáciles en el modelo de Rasch)

ITEMS – DESCRIPTIVOS\_CATEGORIA



● item1   ● Item2   ● Item3   ● Item4   ● Item5   ● Item6   ● item7   ● item8   ● item9   ● item10  
 ● item11   ● item12   ● item13   ● item14   ● item15

Gráfico de burbujas que muestra la dificultad de cada ítem (cálculos realizados con Winstep). En el gráfico un tamaño de burbuja mayor representa mayor dificultad (más difícil que los sujetos repondan en las categorías altas, es decir sólo los muy sintomáticos lo harán)

### 11.4.6 INFIT Y OUTFIT DE LOS ÍTEMS DE LA ESCALA CISSVe

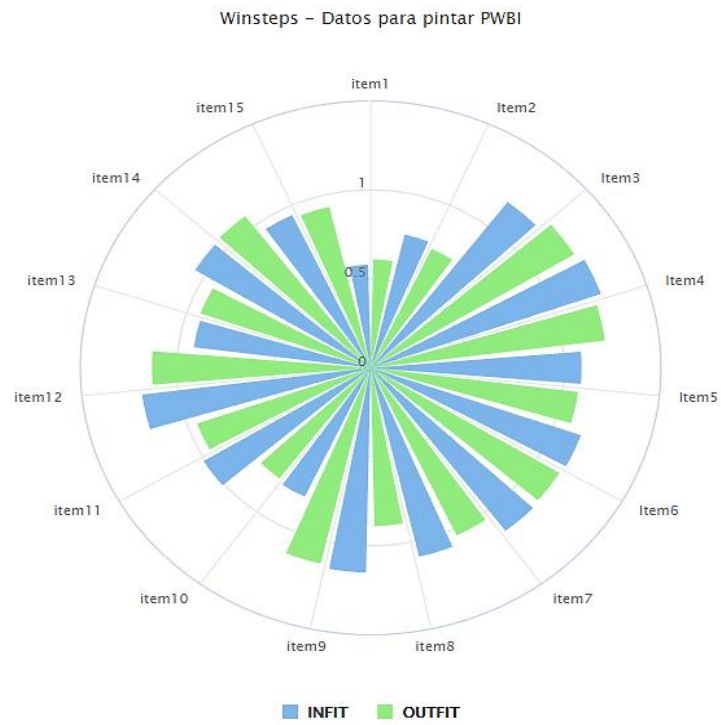


Gráfico radial (centro 0 y radio 1.5) que muestra la calidad de ajuste de cada ítem al modelo según los parámetros de infit y outfit calculados con Winstep

### 11.4.7 MEDIDAS ESPERADAS vs MEDIDAS OBSERVADAS EN LA ESCALA CISSve

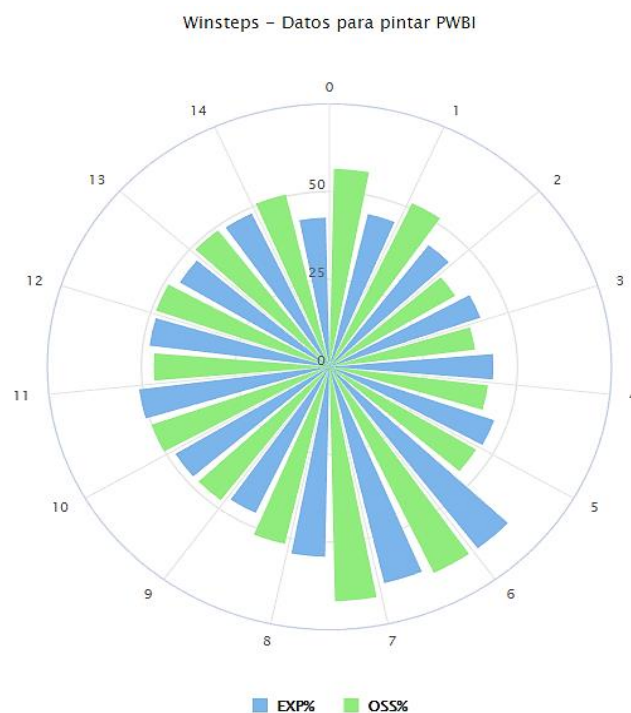


Gráfico radial (centro 0 y radio 75) que muestra en términos porcentuales una comparativa entre las respuestas observadas vs las esperadas por el modelo de Rasch según los cálculos realizados con Winstep

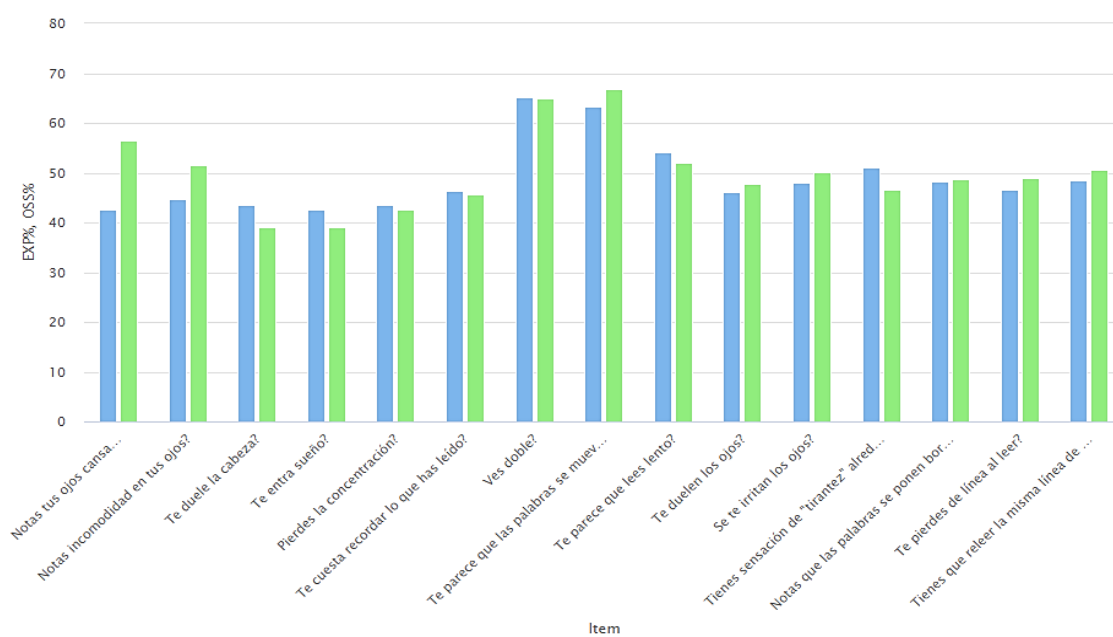


Gráfico de barras que muestra en términos porcentuales una comparativa entre las respuestas observadas vs las esperadas por el modelo de Rasch según los cálculos realizados con Winstep

## 11.5 ANEXO V – ESCALA CISSv15

### 11.5.1 PERMISO DE UTILIZACIÓN DE CISSv15

De: **Eric Borsting** <[eborsting@ketchum.edu](mailto:eborsting@ketchum.edu)>  
Date: mar., 8 mar. 2016 a las 20:45  
Subject: RE: The Convergence Insufficiency Symptom Survey (CISS)  
To: Scheiman, Mitchell <[mscheiman@salus.edu](mailto:mscheiman@salus.edu)>, Mariano González Pérez <[marian06@ucm.es](mailto:marian06@ucm.es)>

Hi

Sorry about the delay in responding. I have been traveling and then got the flu.

This sounds like a great project and feel free to adapt the CISS for your use. If you need any help please feel free to contact me.

Eric

Eric Borsting, OD, MS

Professor

Marshall B. Ketchum University

## 11.5.2 ESCALA CISSv15

## Convergence Insufficiency Symptom Survey

Name \_\_\_\_\_ DATE \_\_\_/\_\_\_/\_\_\_


**Clinician instructions:** Read the following subject instructions and then each item exactly as written. If subject responds with "yes" - please qualify with frequency choices. Do not give examples.

**Subject instructions:** Please answer the following questions about how your eyes feel when reading or doing close work.


		Never	(not very often) Infrequently	Sometimes	Fairly often	Always
1.	Do your eyes feel tired when reading or doing close work?					
2.	Do your eyes feel uncomfortable when reading or doing close work?					
3.	Do you have headaches when reading or doing close work?					
4.	Do you feel sleepy when reading or doing close work?					
5.	Do you lose concentration when reading or doing close work?					
6.	Do you have trouble remembering what you have read?					
7.	Do you have double vision when reading or doing close work?					
8.	Do you see the words move, jump, swim or appear to float on the page when reading or doing close work?					
9.	Do you feel like you read slowly?					
10.	Do your eyes ever hurt when reading or doing close work?					
11.	Do your eyes ever feel sore when reading or doing close work?					
12.	Do you feel a "pulling" feeling around your eyes when reading or doing close work?					
13.	Do you notice the words blurring or coming in and out of focus when reading or doing close work?					
14.	Do you lose your place while reading or doing close work?					
15.	Do you have to re-read the same line of words when reading?					
		__ x 0	__ x 1	__ x 2	__ x 3	__ x 4

TOTAL SCORE \_\_\_\_\_

## 11.6 ANEXO VI – ESCALA CISSve CON ÍTEMS ADICIONALES

		<b>FACULTAD DE ÓPTICA Y OPTOMETRÍA</b> <b>UNIVERSIDAD COMPLUTENSE DE MADRID</b> <small>C/ ARCOBDE JALÓN 118 - 28037 MADRID- ESPAÑA</small>	
<b>CUESTIONARIO - Síntomas de Insuficiencia de Convergencia</b>			
<b>NOMBRE</b>		<b>USUARIO DE GAFAS</b>	
<b>SEXO</b>		<b>USUARIO DE LENTILLAS</b>	
<b>EDAD</b>		<b>HORAS DE TRABAJO DE CERCA</b> (MOVIL, TELE, ORDENADOR, LECTURA)	
<b>Instrucciones</b>			
- Las siguientes preguntas se refieren a cómo te sientes mientras lees o trabajas de cerca			
- Valora de 1 a 5 cada una de ellas de acuerdo a las categorías siguientes			
<b>CATEGORIAS DE RESPUESTA</b>			
1. Nunca			
2. Muy pocas veces			
3. Algunas veces			
4. Muchas veces			
5. Siempre			
<b>PREGUNTA</b>			<b>RESPUESTA</b>
1. ¿Notas tus ojos cansados?			
2. ¿Notas incomodidad en tus ojos?			
3. ¿Te duele la cabeza?			
4. ¿Te entra sueño?			
5. ¿Pierdes la concentración?			
6. ¿Te cuesta recordar lo que has leído?			
7. ¿Ves doble?			
8. ¿Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?			
9. ¿Te parece que lees lento?			
10. ¿Te duelen los ojos?			
11. ¿Se te irritan los ojos?			
12. ¿Tienes sensación de "tirantez" alrededor de los ojos?			
13. ¿Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?			
14. ¿Te pierdes de línea al leer?			
15. ¿Tienes que releer la misma línea de texto?			
16. ¿Tienes la sensación de ponerte bizco?			
17. ¿A veces se te emborronan las letras?			
18. ¿Señalas con el dedo (o algo similar) la palabra que estás leyendo?			
19. ¿Inclinas la cabeza hacia un lado?			
20. ¿Te acercas mucho a lo que estás mirando?			
21. ¿Guiñas los ojos?			
22. Tras un tiempo leyendo u observando una pantalla de cerca ¿Has notado que te molesten las luces?			
23. Tras un tiempo leyendo o trabajando de cerca ¿Tienes que esforzarte para ver bien?			

## 11.7 ANEXO VII – ESCALA CVSS17

	<p><b>FACULTAD DE ÓPTICA Y OPTOMETRÍA</b>  <b>UNIVERSIDAD COMPLUTENSE DE MADRID</b>          C/ ARCOS DE JALÓN 118 - 28037 MADRID - ESPAÑA</p>		
<hr/>			
<b>CVSS17</b>			
Apellidos, Nombre: _____			
Edad: _____	Fecha: _____		
<ul style="list-style-type: none"> <li>• Las preguntas se refieren a <b>cómo se ha sentido durante sus últimas cuatro semanas</b></li> <li>• Entendemos por <b>dispositivos con pantalla</b>, aparatos como el ordenador, el móvil, la tablet etc.</li> <li>• Si usa gafas o lentes de contacto habitualmente para mirar en cerca, por favor responda a todas las preguntas pensando en cómo se siente si las lleva puestas.</li> </ul>			
POR FAVOR, MARQUE CON UNA X SU OPCIÓN PREFERIDA EN CADA PREGUNTA			
<b>¿Ha notado que a veces se le emborronan las letras mientras usa dispositivos con pantalla?</b>			
No, nada	Sí, muy poco	Sí, un poco	
Sí, moderadamente	Sí, mucho	Sí, muchísimo	
<b>¿Nota sus ojos cansados durante o después del uso prolongado de dispositivos con pantalla?</b>			
Nunca	Casi nunca	Poco tiempo	Parte del tiempo
Mucho tiempo	Casi siempre	Siempre	
<b>¿Ha notado que le duelan los ojos mientras usa dispositivos con pantalla?</b>			
Constantemente	Frecuentemente	Raramente	Nunca
<b>¿Ha notado los ojos pesados tras un tiempo de uso de dispositivos con pantalla?</b>			
Constantemente	Frecuentemente	Raramente	Nunca
<b>¿Ha notado que cuando usa dispositivos con pantalla tenga que parpadear mucho?</b>			
Nunca	Raramente	Frecuentemente	Constantemente



FACULTAD DE ÓPTICA Y OPTOMETRÍA  
UNIVERSIDAD COMPLUTENSE DE MADRID

C/ ARCOS DE JALÓN 118 - 28037 MADRID- ESPAÑA

**¿Ha notado sensación de ardor en sus ojos?**

Constantemente                      Frecuentemente                      Raramente                      Nunca

**¿Ha notado que, tras un tiempo de uso de dispositivos con pantalla, tiene que esforzarse para poder conseguir ver bien?**

Sí, muchísimo                      Sí, mucho                      Sí, moderadamente  
Sí, un poco                      Sí, muy poco                      No, nada

**Mientras lee o escribe en dispositivos con pantalla ¿tiene la sensación de que se ponga bizco?**

Constantemente                      Frecuentemente                      Raramente                      Nunca

**¿Ha notado que cuando pasa mucho tiempo con dispositivos con pantalla llega un momento en que acaba viendo las letras dobles?**

Sí, muchísimo                      Sí, mucho                      Sí, moderadamente  
Sí, un poco                      Sí, muy poco                      No, nada

**¿Con qué frecuencia ha notado escozor en la vista mientras está usando dispositivos con pantalla?**

Nunca                      Raramente                      Frecuentemente                      Constantemente

**¿Ha notado que tras un tiempo de uso con dispositivos con pantalla le molesten las luces?**

Nunca                      Casi Nunca                      Unas Pocas Veces  
Varias Veces                      Muchas Veces                      Muchísimas Veces



FACULTAD DE ÓPTICA Y OPTOMETRÍA  
UNIVERSIDAD COMPLUTENSE DE MADRID

C/ ARCOS DE JALÓN 118 - 28037 MADRID - ESPAÑA

A CONTINUACIÓN, TENIENDO EN CUENTA SUS SENSACIONES DURANTE LAS CUATRO ÚLTIMAS SEMANAS, INDIQUE HASTA QUÉ PUNTO HA EXPERIMENTADO LAS SIGUIENTES MOLESTIAS:

	Nada	Muy Poco	Un poco	Moderadamente	Mucho	Muchísimo
Ojos llorosos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ojos Rojos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

POR FAVOR, INDIQUE HASTA QUÉ PUNTO LE PARECE CIERTA O FALSA CADA UNA DE LAS SIGUIENTES FRASES. Si usa gafas o lentes de contacto habitualmente para mirar en cerca, por favor responda a todas las preguntas pensando en cómo se siente cuando si los lleva puestos.

**Al final del día noto que me pesan los ojos**

Bastante falsa

Totalmente falsa

Bastante cierta

Totalmente cierta

**Tras un tiempo de uso con dispositivos con pantalla, noto que tengo que esforzarme para ver bien**

Totalmente cierta

Bastante cierta

Bastante falsa

Totalmente falsa

**Durante el uso de dispositivos con pantalla, tengo que cerrar los ojos para aliviar la sequedad que noto en los ojos**

Totalmente cierta

Bastante cierta

Bastante falsa

Totalmente falsa

**Tras un tiempo de uso con dispositivos con pantalla, me molestan las luces**

Bastante falsa


Totalmente falsa

Bastante cierta

Totalmente cierta

## 11.8 ANEXO VIII – ESCALA WEMWBS

### 11.8.1 PERMISO DE UTILIZACIÓN DE WEMWBS (versión española)



Perez, Carlos  
Para Daniel Serrani

Lu. 29/02/2016 10:28

Responder Responder a todos Reenviar

Se han quitado los saltos de línea adicionales de este mensaje.

Muchas gracias, Daniel  
Así haremos,

Un placer haber contactado con usted  
Enhorabuena por sus excelentes trabajos  
Carlos

-----Original Message-----  
From: Daniel Serrani [<mailto:danielserrani@argentina.com>]  
Sent: sábado, 27 de febrero de 2016 2:02  
To: Perez, Carlos  
Subject: Re: Colaboración

Estimado Profesor Carlos Perez, el cuestionario traducido y validado es de uso publico por lo cual puede usted disponer del mismo libremente, citando a la fuente original.  
Saludos  
Daniel Serrani MD, PhD

## 11.8.2 ESCALA WEMWBS

Equivalencia de traducciones entre la versión inglesa y la versión argentina

Versión español	Versión inglés
3. Me he sentido aliviado	I've been feeling relaxed
6. He resuelto bien los problemas	I've been dealing with problems well
10. Me he sentido con confianza en mí mismo	I've been feeling confident
11. He sido capaz de tomar mis propias decisiones con respecto a las cosas que me suceden	I've been able to make up my own mind about things
12. Me he sentido querido	I've been feeling loved

#### LA ESCALA DE BIENESTAR MENTAL DE WARWICK-EDINBURGH (EBMWE)

A continuación encontrará algunas afirmaciones acerca de ideas y emociones.

Tenga a bien tildar el casillero que mejor describa lo que pensó o sintió durante las últimas 2 semanas

AFIRMACIONES	En ningún momento	Raramente	Algunas veces	Con frecuencia	Todo el tiempo
Me he sentido optimista con relación al futuro	1	2	3	4	5
Me he sentido útil	1	2	3	4	5
Me he sentido aliviado	1	2	3	4	5
Me he sentido interesado por las demás personas	1	2	3	4	5
He tenido suficiente energía de reserva	1	2	3	4	5
He resuelto bien los problemas	1	2	3	4	5
He estado pensando con claridad	1	2	3	4	5
Me he sentido bien conmigo mismo	1	2	3	4	5
Me he sentido cercano a las demás personas	1	2	3	4	5
Me he sentido con confianza en mí mismo	1	2	3	4	5
He sido capaz de tomar mis propias decisiones con respecto a las cosas que me suceden	1	2	3	4	5
Me he sentido querido	1	2	3	4	5
He estado interesado en cosas nuevas	1	2	3	4	5
Me he sentido alegre	1	2	3	4	5

Warwick-Edinburgh Mental Well-Being Scale (WEMWBS)

© NHS Health Scotland, University of Warwick and University of Edinburgh, 2006, all rights reserved.

## 11.9 ANEXO IX – PUBLICACIONES EN REVISTAS INTERNACIONALES

tvst

Article

### Spanish Cross-Cultural Adaptation and Rasch Analysis of the Convergence Insufficiency Symptom Survey (CISS)

Mariano González-Pérez<sup>1</sup>, Carlos Pérez-Garmendia<sup>1</sup>, Ana Rosa Barrio<sup>1,2</sup>, María García-Montero<sup>1</sup>, and Beatriz Antona<sup>1,2</sup>

<sup>1</sup> Optics and Optometry Department, Faculty of Optics and Optometry, Complutense University of Madrid, Madrid, Spain

<sup>2</sup> Applied Vision Research Group, Faculty of Optics and Optometry, Complutense University of Madrid, Madrid, Spain

**Correspondence:** Mariano González-Pérez, Faculty of Optics and Optometry, Complutense University of Madrid, Calle de Arcos de Jalón, 118, 28037 Madrid, Spain. e-mail: marian06@ucm.es

**Received:** January 22, 2019

**Accepted:** January 15, 2020

**Published:** March 23, 2020

**Keywords:** convergence insufficiency; patient-reported outcomes; CISS; Rasch analysis; cross-cultural validation

**Citation:** González-Pérez M, Pérez-Garmendia C, Barrio AR, García-Montero M, Antona B. Spanish cross-cultural adaptation and Rasch analysis of the Convergence Insufficiency Symptom Survey (CISS). *Trans Vis Sci Tech.* 2020;9(4):23. <https://doi.org/10.1167/tvst.9.4.23>

**Purpose:** To culturally and linguistically adapt the Convergence Insufficiency Symptom Survey (CISS) to Spanish and assess the psychometric performance of the new version through Rasch analysis and classical test theory methods.

**Methods:** The Spanish version of the CISS (CISS<sub>VE</sub>) was completed by 449 subjects (9–30 years old) from the general population. The validity and reliability of CISS<sub>VE</sub> were assessed through Rasch statistics (precision, targeting, item fit, unidimensionality, and differential item functioning). To test construct validity, we calculated the coefficients of correlation between the CISS<sub>VE</sub> and the Computer-Vision Symptom Scale (CVSS17) or Warwick-Edinburgh Mental Well-Being Scale (WEMWBS). We determined test-retest reliability in a subset of 229 subjects. We used differential item functioning (DIF) to compare the CISS<sub>VE</sub> and the CISS after administering the CISS to 216 English children.

**Results:** After applying exclusion criteria, the responses of 420 participants (mean age, 18.62 years; female, 54.95%) revealed good Rasch model fit, good precision (person separation = 2.33), and suboptimal targeting (–1.37). There was some evidence of multidimensionality, but disattenuated correlations between the Rasch dimension and a possible secondary dimension were high, suggesting they were measuring similar constructs. No item bias according to gender or age was detected. Spearman's correlation was 0.34 ( $P < 0.001$ ) for CISS<sub>VE</sub>–CVSS17 and non-significant for CISS<sub>VE</sub>–WEMWBS. The limits of agreement for test-retest reliability were 9.67 and –8.71. Rasch analysis results indicated no difference between CISS and CISS<sub>VE</sub>.

**Conclusions:** According to our results, CISS<sub>VE</sub> is a valid and reliable tool for measuring the symptoms assessed by CISS in Spanish people 9 to 30 years of age.

**Translational Relevance:** CISS<sub>VE</sub> can measure convergence insufficiency symptoms in Spanish-speaking subjects.

#### Introduction

Convergence insufficiency is one of the most common abnormalities of binocular vision. It is usually associated with symptoms such as visual fatigue, headaches, blur, and double vision.<sup>1,2</sup> To measure these symptoms, the Convergence Insufficiency Symptom Survey (CISS) was developed in 1999 by the Convergence Insufficiency and Reading Study Group (CIRS).<sup>3</sup>

The first version of this questionnaire consisted of 13 items and assessed the frequency of each symptom using a four-option response scale.<sup>3</sup> In 2003, a revised

version was introduced<sup>1</sup> that included two more items and a new response scale with five choices: never, infrequently, sometimes, fairly often, and always. This new version made the tracking of changes during therapeutic interventions more sensitive.<sup>1</sup> The 15-item version of the CISS (hereafter CISS) is a frequently used outcome measure in binocular vision research and has been used to assess convergence insufficiency (CI) symptoms in various clinical groups from the ages of 8 to 30 years,<sup>1,2,4</sup> where subjects with symptomatic CI had a significantly higher CISS score than others with normal binocular vision. However, to our knowledge, no data have been reported regarding its

translational vision science & technology

Copyright 2020 The Authors  
tvst.arvojournals.org | ISSN: 2164-2591

This work is licensed under a Creative Commons Attribution 4.0 International License.



1

Downloaded from tvst.arvojournals.org on 03/27/2020

psychometric properties apart from its repeatability<sup>2</sup> and known-group validity.<sup>3,4</sup> This last variable reflects the ability of a questionnaire to discriminate between two groups known to differ a priori.

The CISS is not a condition-specific instrument for convergence insufficiency; rather, it is useful for measuring the symptoms associated with visual discomfort caused by different factors. Accordingly, it considers the most common symptoms regarding near-vision problems<sup>5</sup> and provides similar scores in children with accommodative insufficiency and convergence insufficiency.<sup>4</sup> In addition, as described by Horan et al.,<sup>6</sup> some patients with normal sensorimotor exam results were found to score high (i.e., showed a higher level of the assessed symptoms) on the CISS, while others with convergence insufficiency had relatively low scores.

The CISS was developed for English speakers. As there are 442 million native Spanish speakers worldwide, there is currently a need for a Spanish version. We generated a Spanish version of the CISS (CISS<sub>VE</sub>) following well-known guidelines<sup>7-11</sup> used for other recent cross-cultural adaptations<sup>12,13</sup> to ensure content and operational equivalence between the original CISS and the CISS<sub>VE</sub>. Most cross-cultural adaptation studies are based on modern psychometric models such as the Rasch item response theory (IRT) model. This model is recommended for the quality assessment of health questionnaires because (1) it generates a more precise measure, overcoming the limitations of traditional summary scoring through the transformation of ordinal raw scores into interval linear scales<sup>14-17</sup>; and (2) it provides insight into the psychometric properties of the scale and is able to match item difficulty to user skills.<sup>15</sup> The Rasch approach also provides data, such as person and item reliability, reflecting the overall performance of the instrument.<sup>16</sup> The objective of this study was to culturally and linguistically translate the Spanish version and assess its psychometric performance using Rasch analysis and classical test theory methods.

## Methods

Before the study outset, the authors of the CISS gave us their consent to develop a Spanish version of their instrument. The CISS questionnaire consists of 15 items. In reply to each question, the subject indicates the frequency of each symptom on a Likert scale, with scores ranging from 0 to 4: never (0), infrequently (1), sometimes (2), often (3), or always (4). The scores of every item are added to determine the final score,

which ranges from 0 (least symptomatic) to 60 (most symptomatic). The recommended cut-off is  $\geq 21$  for adults<sup>2</sup> and  $\geq 16$  for children 9 to 18 years of age.<sup>1</sup> The study was conducted in two stages. The questionnaire was first translated and adapted to Spanish (May 2016 to April 2017), and then the validity and repeatability of the Spanish version were assessed (May 2017 to January 2018).

Because the CISS assesses near-vision-related symptoms and not only convergence insufficiency symptoms, we enrolled subjects from the general population from four different institutions in Madrid, Spain: a primary school (CEIP Vargas Llosa, Madrid), a secondary school (IES Juan Rodríguez Villanueva), a university faculty (Optics and Optometry Faculty of the Universidad Complutense de Madrid), and a technology company (DXC). In addition, over the period from January 2019 to March 2019, we performed a psychometric analysis of the English CISS on subjects from the personal network of one of our researchers (CP-G) in Swindon, Wiltshire, UK. Subjects 9 to 30 years of age were recruited by convenience. Participants received no compensation for their cooperation. Exclusion criteria were mother tongue different from the questionnaire language, prior visual surgery (not refractive), active visual or neurologic disease, any medication that could affect vision, or any kind of disability preventing the subject from reading or understanding the instrument's questions. Out of 665 subjects enrolled, 449 (mean age, 18.62 years; range, 9-30 years; female, 54.95%) completed the Spanish version of the CISS (CISS<sub>VE</sub>), and 216 (mean age, 15.81 years; range, 12-20 years; female, 37.61%) completed the original CISS.

The study was approved by the Research Ethics Committee of the Hospital Clínico San Carlos (Madrid, Spain), and its protocol adhered to the tenets of the Declaration of Helsinki. All participants gave their written informed consent prior to participation. For participants younger than 18 years, consent was obtained from a parent or guardian. All children older than 12 years also provided their consent before any testing was done. Other than responses to the questionnaires, no other clinical data were collected in any subject.

## Translation and Transcultural Adaptation of CISS

Translation and adaptation were performed according to previously published guidelines<sup>7,18,19</sup> as a five-step process:

1. *Direct translation*—Two bilingual translators with Spanish as their mother tongue, one member of the research group (CP) and another person working in a different area of knowledge (banking), independently translated the original CISS version. CP provided clinical equivalence with the original CISS, and the other translator offered a vernacular version.
2. *Consensus version of the direct translation*—The two bilingual translators ensured that the translation was fully comprehensible.
3. *Back translation*—A further two bilingual translators, this time with the source language (English) as their mother tongue and who were blind to the original version, independently translated the consensus version back into Spanish. Both translators were naive about the concepts explored to avoid information bias.<sup>7</sup>
4. *Expert committee review*—The expert panel included all of the translators, one who was a professional translator involved in the research, two who were experts in binocular vision (AB and BA), and a team member who had a background in patient-reported outcome instrument development (MG-P). The panel held a meeting to consolidate the four previous translations and created the pre-final version of the Spanish CISS, designated *CISS-versión española*, or *CISS<sub>VE</sub>*. The panel solved discrepancies by consensus, and a written summary of this meeting is provided as Supplementary Table S1. As an example, after discussing whether the response option “fairly often” should be translated as “casi siempre” or “bastante menudo,” the latter translation was finally adopted by consensus.
5. *Pre-testing of the consensus version*—Cognitive interviews were conducted using a verbal probing technique with 48 native speakers between the ages of 9 and 30 years to ensure patient comprehension of the *CISS<sub>VE</sub>*; no new issues emerged in this pre-test.

### Analysis Strategy

For descriptive data generation and repeatability assessment, we used SPSS Statistics 22.0 (IBM Corp., Armonk, NY, USA).

### Rasch Analysis

The package Winsteps 4.0.1 (Winsteps.com, Beaverton, OR, USA)<sup>7</sup> was used for Rasch analysis. The Rasch model is an IRT that transforms raw scores to express the person ability and the difficulty of items on

the same scale, so that the difference between the ability of two people does not depend on the specific items with which their ability is estimated. The main IRT concept is that a mathematical model is used to predict the probability of a person successfully replying to an item according to person ability and item difficulty.<sup>20</sup> For the analysis, we chose the Andrich rating scale model (RSM), which assumes equal category thresholds across items, as all items share the same response option structure.<sup>21</sup> Respondents with a greater level of symptoms and items of greater difficulty were located on the negative side of the continuum scale and vice versa. The results of the Rasch method were then used to determine the following.

### Rating Scale Structure

The performance of the rating scale structure was assessed by examining the category threshold order. Disordering of categories occurs when the response options do not follow expected hierarchical ordering.<sup>13,22</sup>

### Item Fit Statistics

Both infit and outfit mean square fit statistics show the extent to which the items in the domain comply with Rasch model expectations.<sup>14</sup>

### Dimensionality

The scale is considered unidimensional when there is one latent variable of interest, and the level of this latent variable is the focus of measurement.<sup>20</sup> To assess multidimensionality, we used the results of the Rasch principal component analysis (PCA) of standardized residuals, which looks for patterns in the part of the data that does not agree with the Rasch measures (unexpected data). When groups of items share the same patterns of unexpected data, those items probably also share a substantive attribute in common, which we refer to as a “secondary dimension” or “contrast.”<sup>23</sup> When variance explained by the Rasch measures is  $\leq 50\%$  and/or the eigenvalue in the first contrast is  $\geq 2.0$ , this is an indication of subsets of items that suggest multidimensionality.<sup>14</sup> Next, we looked at the disattenuated coefficient of correlation between the first and second contrasts obtained in the PCA analysis. Disattenuated correlation approximates the correlation between the two contrasts without measurement error. According to Linacre,<sup>24</sup> 0.82 may be used as the cut-off to consider that the two contrasts measure the same variable.

### Person Separation Index and Levels of Performance

The Rasch-based Person Separation Index (PSI) is a reliability indicator, analogous to Cronbach’s  $\alpha$  of

traditional test theory in both its values and construction.<sup>25</sup> This index was obtained using Winsteps. The number of different levels of performance was computed according to the method described by Wright.<sup>21,26</sup>

#### Targeting

The extent to which item difficulty, defined as the point on the latent variable at which the highest and lowest category of an item have equal probability of being observed, matches the level of a participant's visual abilities was defined as the difference between the average difficulty of the items and the subject's mean level of symptoms.<sup>14</sup>

#### Differential Item Functioning by Gender, Age Group, and CISS Version

We examined each item to determine if there was any difference in the way subgroups (male vs. female; children under 18 years of age vs. young adults older than 18) answered each item—that is, no differential item functioning (DIF). In addition, because testing for DIF is a useful way to validate questionnaire translations,<sup>27</sup> we performed this analysis to test whether the CISS<sub>VE</sub> items were equivalent to those of the original survey. Accordingly, the DIF for an item was considered a cross-cultural or translational issue for that particular translation.<sup>28</sup>

The DIF analysis implemented in Winsteps is based on two methods:

1. Mantel-Haenszel method to estimate the log odds of DIF size and significance from cross-tabs of observations in the two groups
2. Logit-difference (logistic regression) method to estimate the difference between Rasch item difficulties for the two groups, maintaining everything else constant<sup>24</sup>

DIF contrast (i.e., difference in difficulty of the item between the two groups) was defined as no-DIF for  $<0.50$  logits, minimal for 0.50 to 1.0 logits, and notable for  $>1.0$  logits.<sup>13</sup>

The overall quality of the psychometric data obtained in this stage (except levels of performance) was assessed according to the criteria of the guidelines proposed by Khadka et al.<sup>14</sup> for quality assessment of ophthalmologic questionnaires.

#### Validity and Repeatability

Paper versions of the CISS<sub>VE</sub>, the Computer-Vision Symptom Scale (CVSS17),<sup>26</sup> and the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS)<sup>29</sup>

were administered to all participants except the primary-school children. Seven days later or longer, subjects again completed the CISS<sub>VE</sub> in a second session. Convergent validity was assessed by estimating the coefficient of correlation between the subjects' CISS<sub>VE</sub> and CVSS scores and divergent validity through the coefficient of correlation between CISS<sub>VE</sub> and WEMWBS scores. According to Khadka et al.,<sup>14</sup> a coefficient of correlation between CISS<sub>VE</sub> and CVSS greater than 0.3 would be considered as proof of convergent validity. The Kolmogorov-Smirnov test used on the CISS<sub>VE</sub>, CVSS, and WEMWBS scores indicated a non-normal distribution of all measures, so we calculated the Spearman's rho coefficient of correlation. The repeatability of the CISS<sub>VE</sub> was examined via the intraclass correlation coefficient (ICC) with the confidence interval set at 95%. In addition, Bland-Altman limits of agreement were determined to calculate the coefficient of repeatability (CoR) by subtracting the mean difference in scores between the two CISS<sub>VE</sub> sessions from the upper 95% limit.<sup>30</sup>

## Results

### Spanish Version of the CISS

Table 1 shows the CISS<sub>VE</sub> items and response descriptors emerging from the pre-test administered to 48 subjects (33.33% female) and the corresponding items and descriptors taken from the original CISS.

### Rasch Analysis

Of the questionnaires completed by 449 participants, responses to 429 questionnaires (mean age,  $15.92 \pm 5.59$  years; 55.2% female) were used in the Andrich's rating scale model (RSM) analysis implemented in Winsteps. The reasons for excluding 20 of the completed questionnaires from analysis were that more than 33% of the items were not answered by one participant, and outfit  $> 2.5$ <sup>26,31</sup> in the responses provided by 19 participants (outfit is sensitive to unexpected observations by persons on items<sup>24</sup>). The mean CISS<sub>VE</sub> score obtained was  $15.10 \pm 10.13$ , and the range was 1 to 50.

### Rating Scale Structure

There was no disordering of response categories (Fig. 1).

### Item Fit Statistics

Item fit statistics and item measure (difficulty, in logits) for the CISS<sub>VE</sub> are provided in Table 2. All items

**Table 1.** Items and Response Options for the Original Convergence Insufficiency Symptom Survey and the Spanish Version

	Convergence Insufficiency Symptom Survey (CISS) Original Version (US English) Items	Escala Sobre Síntomas de Insuficiencia de Convergencia (CISS <sub>VE</sub> ) Translated (European Spanish) Items
Instructions	Please answer the following questions about how your eyes feel when reading or doing close work.	Las siguientes preguntas se refieren a cómo se siente mientras lee o trabaja de cerca.
Questions		
1	Do your eyes feel tired when reading or doing close work?	¿Notas tus ojos cansados?
2	Do your eyes feel uncomfortable when reading or doing close work?	¿Notas incomodidad en tus ojos?
3	Do you have headaches when reading or doing close work?	¿Te duele la cabeza?
4	Do you feel sleepy when reading or doing close work?	¿Te entra sueño?
5	Do you lose concentration when reading or doing close work?	¿Pierdes la concentración?
6	Do you have trouble remembering what you have read?	¿Te cuesta recordar lo que has leído?
7	Do you have double vision when reading or doing close work?	¿Ves doble?
8	Do you see the words move, jump, swim, or appear to float on the page when reading or doing close work?	¿Te parece que las palabras se mueven, se mezclan o flotan sobre el texto?
9	Do you feel like you read slowly?	¿Te parece que lees lento?
10	Do your eyes ever hurt when reading or doing close work?	¿Te duelen los ojos?
11	Do your eyes ever feel sore when reading or doing close work?	¿Se te irritan los ojos?
12	Do you feel a "pulling" feeling around your eyes when reading or doing close work?	¿Tienes sensación de "tirantez" alrededor de los ojos?
13	Do you notice the words blurring or coming in and out of focus when reading or doing close work?	¿Notas que las palabras se ponen borrosas o que se enfocan y desenfocan?
14	Do you lose your place while reading or doing close work?	¿Te pierdes de línea al leer?
15	Do you have to re-read the same line of words when reading?	¿Tienes que releer la misma línea de texto?
Translation of response categories	0, never; 1, infrequently; 2, sometimes; 3, fairly often; 4, always	0, nunca; 1, muy pocas veces; 2, algunas veces; 3, muchas veces; 4, siempre

showed values inside the interval considered productive for measurement.<sup>13,17</sup> Only the infit and outfit of item 1 were outside the more stringent criterion (0.7–1.3) proposed by Pesudovs et al.<sup>16</sup> and Khadka et al.<sup>14</sup>

#### Dimensionality

Our PCA analysis of the CISS<sub>VE</sub> revealed that 46.3% of the raw variance was explained by the CISS<sub>VE</sub> measures, and an eigenvalue of the first contrast

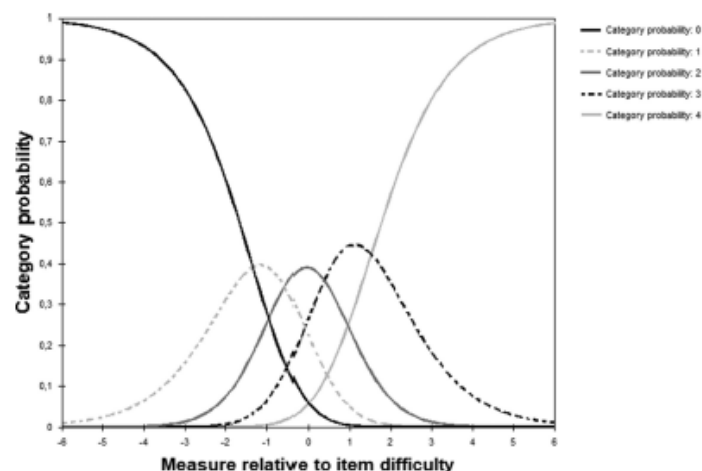


Figure 1. Category probability curves of the CISS<sub>VE</sub>. The figure shows the performance of the five response categories of the CISS<sub>VE</sub>, which asked about the frequency for each of the assessed symptoms. The curve at the extreme left represents "never," and the curve at the extreme right represents "always."

of 2.19. All other contrasts had eigenvalues below 2.00. Thus, in our analysis, the secondary dimension was noticeable because it was bigger than 2.0, indicating that the CISS<sub>VE</sub> measures two different

latent traits. Table 3 shows the items covering the secondary dimension. The disattenuated coefficient of correlation between the first and second contrasts was 0.84, indicating that both dimensions share about 67%

Table 2. Rasch Fit Statistics and Item Measure for CISS<sub>VE</sub>

Item No.	Item Descriptor (in US English)	Infit (MNSQ)	Outfit (MNSQ)	Measure (Logits)
1	Do your eyes feel tired when reading or doing close work?	0.59	0.6	-0.68
2	Do your eyes feel uncomfortable when reading or doing close work?	0.76	0.74	-0.19
3	Do you have headaches when reading or doing close work?	1.2	1.26	-0.55
4	Do you feel sleepy when reading or doing close work?	1.19	1.14	-0.8
5	Do you lose concentration when reading or doing close work?	1.08	1.07	-0.53
6	Do you have trouble remembering what you have read?	1.18	1.15	-0.07
7	Do you have double vision when reading or doing close work?	1.1	1.01	1.17
8	Do you see the words move, jump, swim, or appear to float on the page when reading or doing close work?	1.13	0.9	0.95
9	Do you feel like you read slowly?	1.13	1.13	0.47
10	Do your eyes ever hurt when reading or doing close work?	0.79	0.77	-0.15
11	Do your eyes ever feel sore when reading or doing close work?	1.03	0.96	0.03
12	Do you feel a "pulling" feeling around your eyes when reading or doing close work?	1.17	1.12	0.23
13	Do you notice the words blurring or coming in and out of focus when reading or doing close work?	0.93	0.96	0.07
14	Do you lose your place while reading or doing close work?	1.09	1.09	-0.05
15	Do you have to re-read the same line of words when reading?	0.95	0.94	0.09

MNSQ, infit/outfit mean square.

Spanish Translation and Rasch Analysis of CISS

**Table 3.** Items Comprising the Second Dimension

	Item Descriptor	Standardized Residual Loading
5	Do you lose concentration when reading or doing close work?	0.53
6	Do you have trouble remembering what you have read?	0.47
15	Do you have to re-read the same line of words when reading?	0.45
14	Do you lose your place while reading or doing close work?	0.42
9	Do you feel like you read slowly?	0.35
4	Do you feel sleepy when reading or doing close work?	0.28

Item descriptors (in English) are ordered by first contrast loading as defined by Rasch-residual-based PCA analysis.

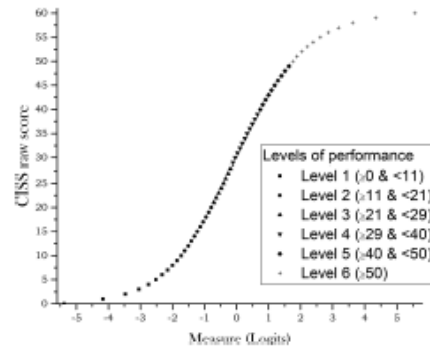
**Table 4.** Summary of the Multidimensionality Test Results and Their Interpretation

Multidimensionality Tests	Result	Interpretation
PCA analysis: percent of raw variance explained by the measure	46.3	Suggests multidimensionality
PCA analysis: eigenvalue of the unexplained variance in the first contrast	2.19	Suggests multidimensionality; secondary dimension was noticeable
Disattenuated coefficient of correlation between the first and second contrasts obtained from PCA analysis	0.84	Rules out dimensionality

of the person measure variance. Table 4 summarizes the results of the tests used to assess unidimensionality and how we used these data to decide whether the CISS<sub>VE</sub> could be considered unidimensional. According to these results, the CISS<sub>VE</sub> can be considered a unidimensional instrument.

**Person Separation Index and Performance Levels**

The PSI for CISS<sub>VE</sub> was 2.33, indicating a reliability of 0.85 and meaning that the CISS<sub>VE</sub> was able to distinguish 3.44 strata of scores. Using the Wright



**Figure 2.** Plot of estimated measure (x-axis) for any CISS<sub>VE</sub> raw score (y-axis). Different symbols represent distinct levels of performance as indicated in the figure inset.

method (a sample-independent method suitable for clinical samples) to determine the number of performance levels across the CISS<sub>VE</sub> score range, we found that the CISS<sub>VE</sub> could distinguish 6.3 levels of symptoms. Figure 2 shows the estimated measure for any CISS<sub>VE</sub> raw score and the correspondence between the raw score and level of performance. Cronbach's  $\alpha$  was 0.90. Table 5 shows the distribution of the CISS<sub>VE</sub> and CISS scores obtained by the subjects included in this study according to performance level.

**Targeting**

The targeting value was -1.37 logits. The item-person map (Figure 3) shows that the items were too difficult for the ability level in this sample, because we assessed a population-based sample in which most were not expected to have near-vision symptoms.

**Differential Item Functioning by Gender and Age**

The results of DIF by gender revealed neither notable DIF nor minimal DIF for any of the CISS<sub>VE</sub> items. Just one item (item 14) showed minimal DIF (0.78) according to age group, as this item was more difficult for young adults than for children. To assess the psychometric properties of CISS<sub>VE</sub>, we compared our Rasch analysis results against the Rasch model expectation<sup>22</sup> using the quality criteria proposed by Khadka et al.<sup>14</sup> (Table 6).

**English Version Versus Spanish Version**

For the English version analysis, after applying the exclusion criteria, we used the responses for 216 questionnaires but excluded those of four with outfit > 2.5. As in Rasch theory, an extreme score (0 or 60

translational vision science & technology

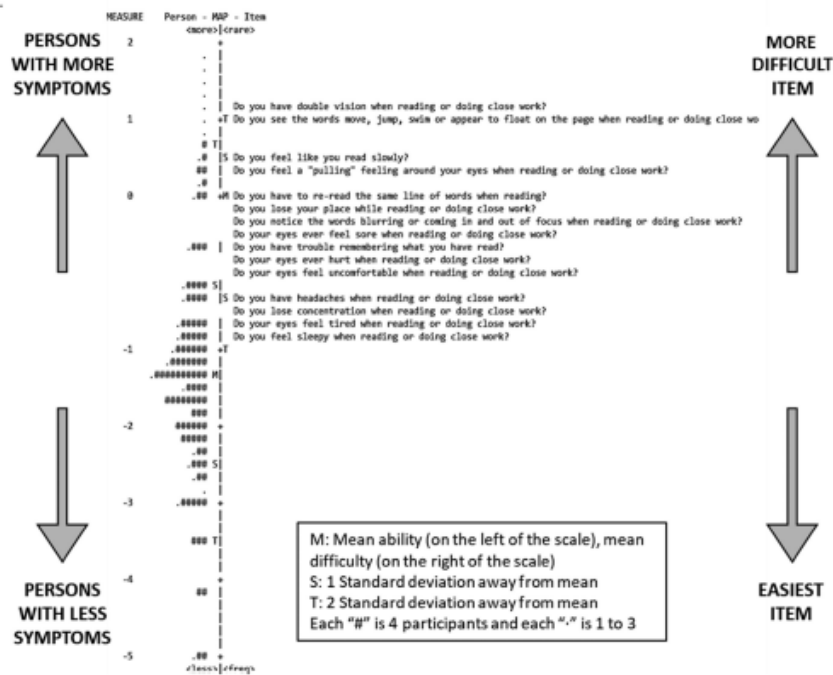
**Table 5.** Demographic and Score Distributions for the 637 Sets of Responses Analyzed in Our Study

	Total (N = 637)	CISS <sub>VE</sub> (n = 429)	CISS (n = 208)
Mean age ± SD	15.90 ± 4.71	15.92 ± 5.59	15.86 ± 1.61
Female proportion	49.8%	55.2%	38%
Score range	1–50	1–50	1–49
Score, mean ± SD	15.86 ± 9.74	15.10 ± 10.13	16.10 ± 9.50
Score median, IQR	15.8–29	14.8–30	16.1–28.1
Level of symptoms, %			
Level 1	33.12	35.20	28.85
Level 2	38.93	38.23	40.38
Level 3	17.11	15.15	21.15
Level 4	8.95	9.32	8.17
Level 5	1.73	1.86	1.44
Level 6	0.16	0.23	0

SD, standard deviation; IQR, interquartile range.

for the CISS) on a questionnaire corresponds to an infinite ability measure (ability measure = symptoms level when using the CISS), which is impractical and also misleading in most situations.<sup>24</sup> Winsteps excluded from the analysis four more completed questionnaires with scores of 0. Finally, the results of 208

translational vision science & technology



**Figure 3.** Item-person map for the CISS<sub>VE</sub>. The Rasch item-person map orders the self-reported level of symptoms of the patients in our study (left side) and the item difficulty (right side).

**Table 6.** Psychometric Properties of the CISS<sub>VE</sub>

Parameter	Rasch Model Expectation	CISS <sub>VE</sub>
Number of items	–	15
Response categories ordering	Ordered	Ordered
Person separation index (reliability)	>2.0 (>0.80)	2.33 (0.85)
PCA		
Raw variance explained by measure, %	>50	45.8
Eigenvalue of the first contrast, logits	<2.0	2.18
Number of items with infit outside 0.7 to 1.3	0	1
Number of items with outfit outside 0.7 to 1.3	0	1
Number of items with DIF for gender of <0.5 logits and $P > 0.05$	None	None
Number of items with DIF for age group of <0.5 logits and $P > 0.05$	None	1
Targeting	$\geq -1.0$	-1.37

PCA, principal component analysis; DIF, differential item functioning.

**Table 7.** Comparison Between the Spanish and English Versions

Parameter	CISS <sub>VE</sub> (n = 429)	CISS (n = 208)
Number of items	15	15
Response categories ordering	Ordered	Ordered
Person separation index (reliability)	2.33 (0.85)	2.31 (0.84)
PCA		
Raw variance explained by measure, %	45.8	46.2
Eigenvalue of the first contrast, logits	2.18	2.13
Number of items with infit outside 0.5 to 1.5	0	0
Number of items with outfit outside 0.5 to 1.5	0	1
Targeting	-1.37	-1.16

PCA, principal component analysis of the residuals; DIF, differential item functioning.

questionnaires (38.0% female; mean age,  $15.86 \pm 1.61$  years) were used in the Andrich's RSM analysis; this sample size is larger than the minimum recommended for DIF assessment. The mean CISS<sub>VE</sub> score was  $16.10 \pm 9.50$ , and the range was 1 to 49. Table 7 compares the main psychometric properties of the two CISS versions.

#### The DIF Contrast Was Below 0.50 for Every Item When Comparing CISS and CISS<sub>VE</sub>

Because the Kolmogorov-Smirnov test indicated a non-normal distribution of the CISS<sub>VE</sub> scores, we ran a Kruskal-

Wallis test followed by Dunn's multiple comparisons on the 429 completed questionnaires. This was designed to examine differences in CISS<sub>VE</sub> performance according to gender and age: males from 9 to 17 years (boys), females from 9 to 17 years (girls), males from 18 to 30 years (young men), and females from 18 to 30 years (young women). The Kruskal-Wallis  $H$  test detected a significant among between the CISS<sub>VE</sub> groups examined ( $H$ , 46.14;  $P < 0.001$ ). Descriptive statistics are provided in Supplementary Table S2 and the significant differences detected by the Dunn's test are shown in Supplementary Table S3.

#### Convergent Validity, Divergent Validity, and Repeatability

We calculated the Spearman rho correlation index between the CISS<sub>VE</sub> and the CVSS at 0.34 ( $P < 0.001$ ). No significant association was detected between the CISS<sub>VE</sub> and the seven items covered in the shortened version of WEMWBS (Fig. 4), which indicates evidence of divergent validity. Correlation between the CISS<sub>VE</sub> and the CVSS was weak (0.34)<sup>32</sup> yet may be considered proof of convergent validity. For the subjects who completed the CISS<sub>VE</sub> twice (test-retest time interval:  $10.23 \pm 3.40$  days), the two-way single-measure ICC for test-retest repeatability was 0.878 (95% confidence interval, 0.845-0.905), and the CoR was 9.22. Figure 5 provides the Bland-Altman plot for the CISS<sub>VE</sub>. The mean difference between sessions was 0.48, and the limits of agreement including 95% of the differences were 9.67 and -8.71, so the CoR was 9.22. For four subjects, the difference in CISS score between sessions was over 10, which was considered by Rouse et al.<sup>2</sup> as "significant and outside the range of normal variability." By considering these four subjects

## Spanish Translation and Rasch Analysis of CISS

TVST | March 2020 | Vol. 9 | No. 4 | Article 23 | 10

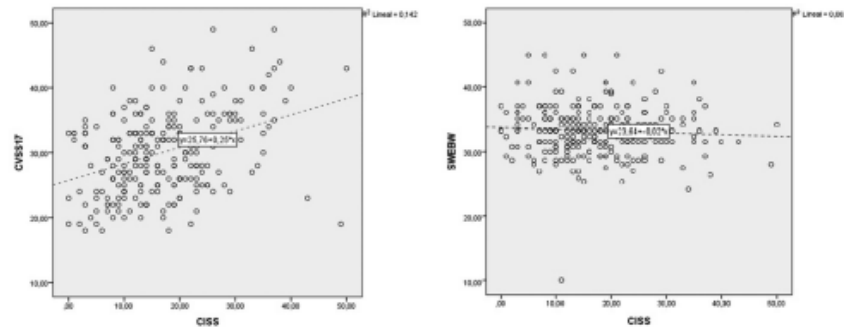


Figure 4. Scatterplots of correlations between the CISS<sub>VE</sub> and the CVSS and between the CISS<sub>VE</sub> and the shortened version of WEMWBS. The regression line is shown as a dotted line.

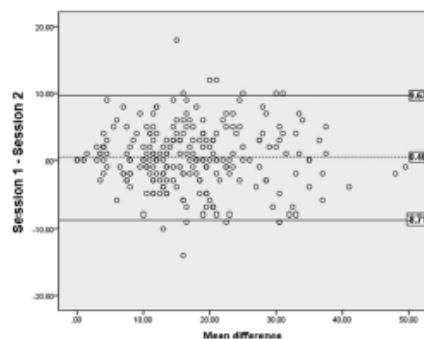


Figure 5. Bland-Altman plot for the CISS<sub>VE</sub>. The dotted line indicates the mean difference (MD) between scores obtained when completing the questionnaire on two occasions. The solid lines indicate the lower and the upper 95% limits of agreement ( $MD \pm 1.96$ ).

as outliers and excluding them from the analysis, the CoR would improve to 8.51.

## Discussion

In this study, we present the Spanish version of the CISS, which shows psychometric properties similar to those for the English version. Our Rasch analysis also confirmed that the overall performance of the instrument is acceptable. A good-quality translation is an essential part of cross-cultural adaptation, but this does not mean that the translated version retains the psychometric properties of the original tool. Other authors

propose a three-step process in which translation is followed by formal assessment of psychometric properties and validity and reliability testing.<sup>8,11</sup>

As recommended by Bradley and Massof,<sup>33</sup> we directly compared item psychometric properties between the CISS<sub>VE</sub> and the CISS to determine whether both tests worked in a similar way. According to their almost identical reliability and residual PCA results (Table 7), the psychometric performance of the CISS<sub>VE</sub> proved similar to that of the CISS. A small difference was noted in targeting ( $-1.37$  CISS<sub>VE</sub> vs.  $-1.16$  CISS), as the mean symptoms score in the English sample was one point higher than in the Spanish subjects. When comparing both versions, just one item (item 12: Do you feel a “pulling” feeling around your eyes when reading or doing close work?) showed an outfit value (1.86) far from Rasch model expectation (1.50). This was attributed to six respondents who scored 0 on every item except item 12, which they awarded a score of 1. Exclusion of these six subjects yielded an outfit value of 0.90. This indicates that the poor performance of item 12 may be attributed to the composition of our English sample. To confirm the equivalence between both versions, our DIF analysis confirmed that the CISS items had been optimally translated into Spanish (European).

We also compared the psychometric properties of the CISS<sub>VE</sub> arising from Rasch analysis through Rasch model expectation<sup>34</sup> (Table 6). Measurement precision was high, and more than six levels of convergence insufficiency symptoms could be distinguished in the study population. Further, although the CIRS group did not use Rasch analysis to develop the original questionnaire, just one item of our Spanish version (item 1: Do your eyes feel tired when reading or

doing close work?) showed Rasch infit and outfit values under the minimum suggested by Khadka et al.<sup>14</sup> (0.7). This indicates that this item's responses are too predictable despite being within the interval of 0.5 to 1.5, considered productive for measurement.<sup>35</sup> Because the presence of one or two items with infit or outfit between 0.5 and 0.7 is deemed acceptable,<sup>14,16</sup> item 1 was retained in the questionnaire without any modification.

Our results revealed that the CISS<sub>VE</sub> is an instrument without DIF for gender and that minimal DIF<sup>13</sup> according to age exists for only one item (item 14: Do you lose your place while reading or doing close work?). According to Khadka et al.,<sup>14</sup> a notable DIF is  $>1.0$  logits, so we could directly compare CISS<sub>VE</sub> scores across these subgroups. Our analysis revealed higher CISS<sub>VE</sub> scores in the young adults than in children (Supplementary Table S3). These differences may be due to a greater cognitive load associated with near vision in the young adults group exacerbating the symptoms normally induced by visual stressors.<sup>36</sup>

We also examined convergent validity by comparing the CISS<sub>VE</sub> and the CVSS17. As predicted, a significant association emerged between them with a coefficient of correlation higher than 0.3 (Fig. 4). This value is the minimum recommended by Khadka et al.<sup>14</sup> when assessing convergence validity. Further, it is the minimum suggested by the COSMIN guideline for systematic reviews of patient-reported outcome measures<sup>37</sup> when evaluating construct validity by studying correlations with instruments measuring related but dissimilar constructs (like we did here). In addition, there was no correlation between the CISS<sub>VE</sub> and the Spanish version of the WEMWBS,<sup>29</sup> so our study provides some evidence of CISS<sub>VE</sub> divergent validity, as it works in the expected manner.

Our CISS<sub>VE</sub> showed a mean score ( $15.10 \pm 10.13$ ) that was comparable to those reported in studies conducted in similar populations, such as the adolescents assessed by Horan et al.<sup>6</sup> ( $16.3 \pm 11.4$ ) and the university students used to develop the Portuguese version (CISS<sub>VP</sub>;  $15.56 \pm 8.86$ ).<sup>3</sup> These two studies provided mean scores for the entire group of participants instead of separating scores according to subjects' visual problems. Furthermore, the mean score obtained in our sample (from a general population) is higher than those reported by Rouse et al.<sup>2,38</sup> in adults and children with normal binocular vision recruited from a clinical population ( $11.3 \pm 8.1$  for adults and  $10.4 \pm 8.1$  for children). As expected, our sample's mean score was lower than the values obtained in children<sup>7</sup> and adults<sup>38</sup> with symptomatic CI ( $37.3 \pm 9.3$  for adults and  $29.8 \pm 8.1$  for children).

A main strength of our study was that we used Rasch analysis to analyze the psychometric properties of both the CISS<sub>VE</sub> and the CISS and used data from the DIF analysis to compare the items of the two versions. However, our study also has the limitation of suboptimal targeting due to the use of the general population instead of a purposive sample.

Good targeting determines higher person reliability, so tests with poor targeting are worse at distinguishing between high and low performers. Thus, this could be a limitation of the CISS<sub>VE</sub> because the targeting value ( $-1.37$  logits) was lower than recommended ( $<-1.0$ ) by Khadka et al.<sup>14</sup> and Pesudovs et al.<sup>16</sup> This suboptimal targeting is a typical issue of scales designed to measure symptoms<sup>26</sup> when administered to the general population, as we did, and could indicate less measurement precision in subjects scoring far from the items' distribution mean (i.e., subjects with fewer symptoms). However, the number of levels of performance (5.8) determined by the Wright method, which is a sample-independent technique derived from Rasch analysis,<sup>39</sup> suggests the high reliability of the CISS<sub>VE</sub>. Given this sample-independent reliability along with the fact that clinicians and/or researchers usually focus on persons with scores closer to the items' mean, we consider this CISS<sub>VE</sub> mistargeting acceptable for its purpose.

Rasch analysis of the CISS<sub>VE</sub> and the CISS suggested multidimensionality, as the variance explained by its measures was under 50% and the eigenvalue of the first strength was above 2.0. When we examined the six items included in the putative dimension arising from the PCA analysis (Table 3), we noted that they were those items exploring complaints other than visual and ocular symptoms. Accordingly, subjects provided answers about reading consciousness or reading performance differently than they did about symptoms. However, because this secondary dimension is closely related to the Rasch dimension, as shown by their disattenuated correlation coefficient (0.84), we can consider that the two dimensions are two different categories of the same trait (e.g., calculus and trigonometry items on a math test), so we may consider the CISS<sub>VE</sub> a unidimensional tool for statistical purposes.<sup>24</sup>

In addition, to compare 95% limits of agreement, we selected the study by Rouse et al.,<sup>2</sup> as it has a mean test-retest interval similar to ours ( $10.50 \pm 7.50$  days vs.  $10.23 \pm 3.40$  days). As expected, limits of agreement were similar in our study ( $9.67$  and  $-8.71$ ) and in the study by Rouse et al.<sup>2</sup> ( $9.0$  and  $-7.6$ ). According to Khadka et al.<sup>14</sup> and Pesudovs et al.,<sup>16</sup> the ICC reported in our sample (0.878) indicates high test-retest reliability. The CoR was 9.22 (8.51 without the four outliers), showing that, in the test-retest data, the probabil-

ity of detecting a test-retest change in CISS score greater than 9.22 in the test population is 2.5%. These results imply that a clinician can be sure that a treatment has a significant impact on a patient's symptoms when finding a change of 10 points, the same value provided by Rouse et al.<sup>2</sup> for the original CISS. Test-retest reliability is optimal when the limits of agreement are lower than the minimal clinically important difference (MCID) values, although lower values of a similar magnitude are considered positive.<sup>16</sup> To sum up, the CoR of the CISS<sub>VE</sub> would be good if we consider valid the value given for the CISS (10 points), but further studies are needed to define precise MCID values for this questionnaire.

Rasch analysis could be used to reengineer the CISS to enhance areas in which it here showed lower performance such as dimensionality or targeting. There are several options for this, such as collapsing some categories and/or deleting the second dimension found in the residual PCA. As several options are available, the clinical relevance of any change should be considered to assess the effectiveness of these improvements. For example, we could consider deleting the second dimension if the instrument becomes more sensitive to clinically meaningful changes.

## Conclusions

We developed a Spanish version of the CISS showing performance similar to that of the original version in English. We also identified some psychometric properties of the CISS that should be addressed in future studies to improve this instrument as a measure of near-vision-related symptoms.

## Acknowledgments

The authors thank the parents, pupils, and teachers of the schools CEIP Vargas Llosa and IES Juan Rodríguez Villanueva; members of the Faculty of Optics and Optometry of the Complutense University; and the workers of DXC for their cooperation. We also thank Anahí González Bergaz for help with distributing the questionnaires.

This study was partially supported by Project Santander-UCM PR26/16-20302.

Disclosure: **M. González-Pérez**, None; **C. Pérez-Garmendia**, None; **A.R. Barrio**, None; **M. García-Montero**, None; **B. Antona**, None

## References

1. Borsting EJ, Rouse MW, Mitchell GL, et al. Validity and reliability of the revised convergence insufficiency symptom survey in children aged 9 to 18 years. *Optom Vis Sci*. 2003;80:832-838.
2. Rouse MW, Borsting EJ, Mitchell GL, et al. Validity and reliability of the revised convergence insufficiency symptom survey in adults. *Ophthalmic Physiol Opt*. 2004;24:384-390.
3. Borsting E, Rouse MW, De Land PN. Prospective comparison of convergence insufficiency and normal binocular children on CIRS symptom surveys. Convergence Insufficiency and Reading Study (CIRS) group. *Optom Vis Sci*. 1999;76:221-228.
4. Borsting E, Rouse MW, Deland PN, et al. Association of symptoms and convergence and accommodative insufficiency in school-age children. *Optometry*. 2003;74:25-34.
5. Tavares C, Nunes AMMF, Nunes AJS, Pato MV, Monteiro PML. Translation and validation of Convergence Insufficiency Symptom Survey (CISS) to Portuguese - psychometric results. *Arg Bras Ophthalmol*. 2014;77:21-24.
6. Horan LA, Ticho BH, Khammar AJ, Allen MS, Shah BA. Is the convergence insufficiency symptom survey specific for convergence insufficiency? A prospective, randomized study. *Am Orthopt J*. 2015;65:99-103.
7. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine J*. 2000;25:3186-3191.
8. Gandek B, Ware JE, Jr. Methods for validating and norming translations of health status questionnaires: the IQOLA Project approach. International Quality of Life Assessment. *J Clin Epidemiol*. 1998;51:953-959.
9. Gjersing L, Caplehorn JR, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Med Res Methodol*. 2010;10:13.
10. Muñoz J, Elosua P, Hambleton RK. Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*. 2013;25:151-157.
11. Ware JE Jr, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA Project approach. International Quality of Life Assessment. *J Clin Epidemiol*. 1998;51:945-952.
12. Adnan TH, Mohamed Apandi M, Kamaruddin H, et al. Catquest-9SF questionnaire: validation of

- Malay and Chinese-language versions using Rasch analysis. *Health Qual Life Outcomes*. 2018;16:5.
13. Khadka J, Huang J, Mollazadegan K, et al. Translation, cultural adaptation, and Rasch analysis of the visual function (VF-14) questionnaire. *Invest Ophthalmol Vis Sci*. 2014;55:4413–4420.
  14. Khadka J, McAlinden C, Pesudovs K. Quality assessment of ophthalmic questionnaires: review and recommendation. *Optom Vis Sci*. 2013;90:720–744.
  15. McAlinden C, Pesudovs K, Moore J. The development of an instrument to measure quality of vision: the Quality of Vision (QoV) questionnaire. *Invest Ophthalmol Vis Sci*. 2010;51:5537–5545.
  16. Pesudovs K, Burr J, Harley C, Elliott D. The development, assessment, and selection of questionnaires. *Optom Vis Sci*. 2007;84:663–674.
  17. Pesudovs K, Garamendi E, Elliott D. The Quality of Life Impact of Refractive Correction (QIRC) Questionnaire: development and validation. *Optom Vis Sci*. 2004;81:769–777.
  18. WHO. Management of substance abuse: process of translation and adaptation of instruments. Available at: [https://www.who.int/substance\\_abuse/research\\_tools/translation/en/](https://www.who.int/substance_abuse/research_tools/translation/en/). Accessed February 17, 2020.
  19. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcome (PRO) measures: report of the ISPOR task force for translating adaptation. *Value Health*. 2005;8:94–104.
  20. Wu M, Adams R. *Applying the Rasch Model to Psycho-Social Measurement: A Practical Approach*. Melbourne, Australia: Educational Measurement Solutions; 2007:87.
  21. Wright BD. Model selection: rating scale model (RSM) or partial credit model (PCM)? *Rasch Meas Trans*. 1998;12:641–642.
  22. Khadka J, Gothwal VK, McAlinden C, Lamoureux EL, Pesudovs K. The importance of rating scales in measuring patient-reported outcomes. *Health Qual Life Outcomes*. 2012;10:80.
  23. Linacre JM, Wright B. Winsteps help for Rasch analysis. Available at: URL: <http://www.winsteps.com/index.htm>. Accessed June 27, 2019.
  24. Linacre J. *Winsteps Rasch Measurement Computer Program User's Guide*. Beaverton, OR: Winsteps.com; 2014.
  25. Marais I, Andrich D. Formalizing dimension and response violations of local independence in the unidimensional Rasch model. *J Appl Meas*. 2008;9:200–215.
  26. González-Pérez M, Susi R, Antona B, Barrio A, González E. The Computer-Vision Symptom Scale (CVSS17): development and initial validation. *Invest Ophthalmol Vis Sci*. 2014;55:4504–4511.
  27. Petersen MA, Groenbold M, Bjorner JB, et al. Use of differential item functioning analysis to assess the equivalence of translations of a questionnaire. *Qual Life Res*. 2003;12:373–385.
  28. Watt T, Barbesino G, Bjorner JB, et al. Cross-cultural validity of the thyroid-specific quality-of-life patient-reported outcome measure, ThyPRO. *Qual Life Res*. 2015;24:769–780.
  29. Castellvi P, Forero CG, Codony M, et al. The Spanish version of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) is valid for use in the general population. *Qual Life Res*. 2014;23:857–868.
  30. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–310.
  31. Gonzalez-Perez M, Susi R, Barrio A, Antona B. Five levels of performance and two subscales identified in the computer-vision symptom scale (CVSS17) by Rasch, factor, and discriminant analysis. *PLoS One*. 2018;13:e0202173.
  32. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg*. 2018;126:1763–1768.
  33. Bradley C, Massof RW. Validating translations of rating scale questionnaires using Rasch analysis. *Ophthalmic Epidemiol*. 2017;24:1–2.
  34. Khadka J, Gao R, Chen H, et al. Re-engineering the Hong Kong Quality of Life Questionnaire to assess cataract surgery outcomes. *J Refract Surg*. 2018;34:413–418.
  35. Linacre J. What do infit and outfit, mean-square and standardized mean? *Rasch Meas Trans*. 2002;16:878.
  36. Gowrisankaran S, Nahar NK, Hayes JR, Sheedy JE. Asthenopia and blink rate under visual and cognitive loads. *Optom Vis Sci*. 2012;89:97–104.
  37. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27:1147–1157.
  38. Rouse M, Borsting E, Mitchell GL, et al. Validity of the Convergence Insufficiency Symptom Survey: a confirmatory study. *Optom Vision Sci*. 2009;86:357–363.
  39. Wright B. Separation, reliability and skewed distributions: statistically different levels of performance. *Rasch Meas Trans*. 2001;14:786.